See *Addendum* on page 27 (November 2001)

The Endangered Medical Record

Both patient care and biomedical knowledge are threatened

Vergil Slee, MD

The honesty, accuracy, and completeness of the medical record are threatened by today's healthcare system and crippled by our "code-dependent" technology. The result is declining quality of care, increased costs, and a public policy often built on sand.

Summary

Three Basic Points

- The medical record is the key document for personal health care. We assume that our medical records provide faithful and complete memory for our physicians and for communication among all the caregivers. Increasingly this assumption may not be correct.

- Medical record information has many "secondary" uses – billing for our care and developing statistics for many purposes, ranging from clinical research to national health policy development. Each use requires the proper information.

- The entire information system is code-dependent. Information is captured as codes and retrieved via codes. The codes must be able to give us what we need.

The Medical Record System Has Four Correctable Flaws

Flaw 1. "The System" Encourages Distortion of Information. Powerful forces – reimbursement, regulations, fears, technology – distort the information going into the medical record. The effects of public policies and other factors which tend to bias the clinical content of medical records should be carefully considered and brought to national attention for correction.

Flaw 2. The System Can't Capture Exact Diagnoses. We do not capture detailed diagnoses. This can only be corrected by adding to the medical record system a new "front end" which captures and gives permanent codes to these diagnoses (these "clinical entities"). This is called "entity coding."

Flaw 3. We Can't Tell For Sure What Any of the Codes Mean. Our present diagnosis coding, called "category coding," tells only where the diagnoses go in *ICD-9-CM*, which is changed periodically, with the codes often changing their meanings. The only correction for this problem is to "tag" these and all other codes so that we can tell exactly what each one means.

Flaw 4. Our "Single-Classification" System of Statistics Can't Serve All Needs: We have to do everything – from billing to developing national policy – using the same groupings of patients. This can't be the best way to manage our health system. One size can't fit all. This flaw can be corrected only after our medical records provide entity codes to be used as the building blocks for classifications "tailored" to their purposes, ranging from reimbursement to public policy decisions to evidence-based medicine.

The Endangered Medical Record[1]

Your Personal Medical Record is a crucial element of your personal healthcare. Its primary function is to assist healthcare providers in caring for you, and for this it must contain a complete and accurate record of your health problems, their management, and your response. And our national health information system begins with the medical record. Information collected from individual records provides the basis for most of our knowledge about our nation's healthcare – who gets sick, what they get, how they're treated, how they respond. These records – several billion each year – contain priceless information, which could provide a depth to our collective medical knowledge impossible before now. But we can't get at all of this information, nor can we rely on what we do get. There have been problems for decades with our "system" of recording and retrieving health information, and it's time to give these problems immediate attention. We hope to convince you why it's imperative to fix these problems now – and how it can be done.

The Medical Record

The medical record is an irreplaceable tool for medical care

Medical records were invented by physicians and nurses in order to make sure that they remembered all the facts about each individual patient, and so that all the people concerned in the care – consultants, other physicians, nurses on other shifts, specialized technicians, referral institutions – would have the same information.

Your personal, individual medical record (sometimes you have more than one) should contain information about your past and present problems, diagnoses, care provided, and the results. At any time you come into contact with healthcare providers, information should be added by all the individuals involved (including you). Your record also collects information from the laboratory, X-ray department, pharmacy, and other sources.

Your medical record is a living, growing resource for your health and healthcare. As, over time, more and more caregivers become involved in your care, it is increasingly important that your record be able to function as the source of memory and the center of communications.

You probably assume that all essential information has been recorded in your medical record, and that as you visit your providers for care, that information will be available to them. Unfortunately, this assumption is increasingly shaky. There are elements of our present healthcare system, and of our information technology, which interfere with both the accurate recording of clinical information and its retrieval. Let's imagine some real-life scenarios:

[1] Extracted from *The Endangered Medical Record: Ensuring Its Integrity in the Age of Informatics.* Slee, VN, Slee, DA, and Schmidt, HJ. Saint Paul: Tringa Press, 2000. Revised November 2001.

Three Scenarios About Patient Care

What if – your physician went to your medical record to help his memory with your physical problem, which he had felt was partially due to stress, but the details were vague in his mind. Then he remembered that he had left out critical facts about a family problem because he was afraid that the computer would not keep them confidential – although he had been assured that security was higher in the electronic system than in his old paper-based records. This sometimes happens.

What if – your physician consulted your medical record about your illness last year, and found only that you had "some other disease of the respiratory system." Your former doctor is not available, so your new doctor looks up the code for "some other..." She finds 30 pages of things you didn't have, but is still no closer to what you actually had. This happens today.

What if – you had served in the Gulf War and when you returned you thought your symptoms were the result of something that happened over there. So did your doctor, and he tried to record Gulf War syndrome. He was instructed to put in your record codes for "late effects of war" and "fatigue" – there was no way to put down what he actually wanted to. This is still true in October 2000.

Secondary Uses of Medical Record Information

As our healthcare system has become more complex, medical records have been called upon to serve many "secondary" purposes, beyond direct patient care. These range from individual case billing, to quality management, research, formulating national health policy, and exchange of international health statistics. Many of these uses are vital.

The First "Secondary" Uses Of Medical Records

Early on, hospitals began using medical record information internally, for statistics on groups of patients. Hospitals wanted to know what their case loads were, and the trends, so they could intelligently plan their facilities. Physicians wanted to know the rates of success and failure with the care their hospitals provided and the success they achieved as individuals (quality measurement). Each statistic was, of course, developed by aggregating data from individual records.

The next demand for detail on the individual patient was external, and it began when healthcare insurance appeared on the scene. Payment was made (or denied) based on the clinical information (diagnoses and treatments) in the individual medical record.

Imagine again some scenarios, and what might result if we can't rely on our data:

Scenarios About Biomedical Information

What if – you were writing a newspaper story and tried to back-track on the AIDS epidemic as it had been treated in your state. You got figures, and for some reason you inquired as to just where they came from. You learned that they came from the coding in medical records and on bills submitted from hospitals and doctors. You also found that there had been a half dozen or more ways that AIDS cases might have been coded since the first cases in 1981, but that no one could tell you how they actually had been coded. Sometimes they were coded as diseases of the immune system and sometimes as infectious diseases. And no one was really sure whether or not all the codes that could have been used were actually looked up in compiling the statistics. And your state was not unique – all states had the same fuzzy data.

What if – you were doing research on the history of heart attacks (acute myocardial infarction (AMI)). Of course you had to look it up by finding its code, but you found that its codes had changed in mysterious ways. Before 1948 AMI was lumped

into a group of different kinds of arteriosclerotic heart disease. In 1962 it was separated out pretty well with its own code, but in 1965 in one coding system you couldn't separate it from "arteriosclerotic heart disease of over 8 weeks duration," while in another you had to look in two codes, "AMI with hypertension" and "AMI without hypertension." In still another system, you could also find out which artery was involved. Of course, all these codes had different numbers. Hopeless. Medicine knew a lot about AMI before 1948 – it was well known and diagnosed, including which artery. But the codes – the only kind we still use today – get between us and the detail we are looking for.

What if – you wanted to look at healthcare data to help in managing a hospital, and then at another time wanted to look at the same data as the base on which to develop public policy. You'd want quite different views, neither of which could be met very well from data whose primary use is to code for the bills submitted for care.

The Four Flaws

The problems illustrated in the scenarios above are the result of four "flaws" in our health information system:

1. **"The System" encourages distortion of information**
2. **The System can't capture exact diagnoses**
3. **We can't tell for sure what any of the codes mean**
4. **Our "Single Classification" System of statistics can't serve all needs**

Underlying the last three of the flaws, and part of the first flaw as well, is the fact that our system for retrieval of information, and increasingly for putting information into the medical record, is based on coding – our system is "code-dependent." We must learn to live with this code-dependency. So before we talk specifically about the flaws, we need some background on coding and classification in health information.

Code-Dependency in Health Information

Since the first "secondary usages" of medical records, their information has been coded. Hospitals and the larger clinics, which have greater demands on medical record information, found it more convenient and trustworthy to file and retrieve the records and the information in them by the use of codes, rather than by using narrative terms or phrases. This was simply the adoption of the methods used elsewhere in society. One doesn't order a car part by giving a long description of the part; one asks for it by giving its unique part number. Hospitals gave patients numbers and filed their medical records ("charts") by these numbers. Physicians were also given numbers, both for convenience and confidentiality. Diagnoses, operations, and other treatments were given code numbers.

It is a given today that retrieval of diagnoses and procedures is on the basis of their codes – not the language in which the diagnostic and procedure terms are expressed in the record itself prior to coding. This "fact of life" has resulted in some problems, but that doesn't mean we can't live with it. In fact, we need codes – they serve a number of essential purposes.

Where We Got Our First Clinical Codes

When coding of diagnoses and procedures first began, a source was needed for the codes themselves. The most prominent source of codes in the 1930s was the American Medical Association's *Standard Nomenclature of Diseases and Operations (SNDO),* which attempted to introduce a standardized, "controlled" language for physicians to use to describe what the patient presented, what they found, and what treatments they employed. *SNDO* was a "Nomenclature," not a coding system. But each term was preceded by a number used for sorting the terms, and quite naturally, this sorting number was immediately used as a code to be exchanged for the term.

"Category Codes" – and Why They're Essential

Not only do we need codes, but we need codes which will place patients into groups. Collecting diagnoses into categories is essential. No one could dispute the necessity, for some purposes, of classifying cases according to their demographic characteristics or on the basis of their problems, diagnoses, management, and outcomes. We need this information for:

Billing & healthcare financing – a million different diagnoses would be unmanageable

Quality review – displaying patterns of care

Administration, including staffing & facilities planning

Clinical research

Epidemiological studies

Evidence-based medicine

International statistics

Public policy, planning

Actuarial analyses

Coding diagnoses according to their categories (where they belong in a classification), rather then their identity (what the diagnoses are) is called "category coding."

Our diagnosis codes today originate from the *International Classification of Diseases*[2] *(ICD),* used universally in this country since the mid-1960s for management of medical record information. Each revision of *ICD* has required modification for use in the United States in order to give it additional specificity for clinical use. The current edition is *International Classification of Disease, 9th Revision, Clinical Modification (ICD-9-CM).*

The codes of *ICD-9-CM* are for some 13,000 groups of diagnoses – not for individual diagnoses. These groups are collected into categories, and it is the rare category, indeed, which contains only one diagnosis. Decoding only gives the name of the *ICD-9-CM* category. The code doesn't know what the original diagnosis was, only that it was one of the group of diagnoses within that category.

For hospital reimbursement, we further aggregate the various diagnoses (after they are coded to *ICD-9-CM*) into about 500 "*Diagnosis Related Groups (DRGs).*" That has been proven to be a large enough list of "products" for which to develop prices.

How We Got Into Category Coding

There are just entirely too many diagnoses to always deal with individually. *ICD-9-CM* lists the names for roughly 100,000 different diagnoses. Indexes to medical textbooks and publications add several hundred thousand more. And the language actually used by both caregivers and patients adds another few hundred thousand. With so many diagnoses, there are likely to be many diagnoses (or diagnostic terms) for which there are only a few patients. We could never make significant statistics or develop a reimbursement scheme without grouping the patients.

SNDO was abandoned in the 1950s as our source of diagnosis and procedure codes because it was unable to retrieve medical records in the fashion that they were called for. The demand made on the indexing system was for groups of patients – the diagnostic picture of the patients in each clinical service of the hospital, all patients having appendectomies, death rates for patients with heart attacks, and so forth. The ultra-specificity of the *SNDO* approach was worse than no coding at all; it actually interfered with assembly of the desired groups.

The codes of the *International Classification of Diseases (ICD),* which we began using in the 1950s, automatically grouped patients in clinically useful ways. They also used language much closer to that of medicine rather than the artificial, imposed set of modular terms found in *SNDO*.

Category coding evolved from the need to group cases, and has simply grown, over time, without serious analysis. The information system began with the manual systems used within individual hospitals, where individual medical records were always available if the exact diagnoses were required. As computers entered the picture, they simply picked up what was being done manually.

And, almost imperceptibly, we transferred our retrieval of information to the computer as well – and it uses the codes instead of the original narrative. Now we use codes for input. We are now beginning to realize the handicap we have created for ourselves by using only category coding in our information system.

[2] *The International Classification of Diseases* is a classification with periodic, numbered Revisions, published by the World Health Organization (WHO).

The "Four Flaws" – Where the Problems Are

With this background, let's take a closer look at the "flaws" in the system.

Flaw 1: "The System" Encourages Distortion of Information

The information that caregivers put in the medical record is being influenced more and more by the healthcare world. There are several elements in our system today which undermine, rather than support, truthful recording of clinical information. These include factors which tend to bias the information, and factors which interfere with the accuracy of information.

"Biasing" factors include reimbursement, federal regulations, and security fears.

"Interfering" factors include limitations on and inhibitions to getting accurate information into the medical record. Specifically, the emerging electronic medical record can seriously constrain what gets recorded.

Reimbursement

Medicare and other payers only cover specific diagnoses, procedures, and services. The provider must submit the correct codes to obtain payment. Software is available to assist in optimal coding. Sometimes, the software will warn the provider against using "unacceptable" codes (codes which may set off alarms in the payment system) – even if they are the accurate ones.

"Gaming" the System

Medical record information needed for billing is submitted in the form of codes. At first, billing coding was only detailed enough to find out if the case was entitled to benefits. Initially this was determined by asking for codes from a limited list (even though many codes were for things not covered).

The list contained, for example, an item entitled "mental illness" which, if present (i.e., if it was recorded or coded), could be denied. At this point the first "gaming" of the information system appeared. An acceptable primary diagnosis such as pneumonia was listed first, and the unacceptable but "actual" diagnosis, perhaps "alcoholism," was listed as secondary. This gaming worked as long as the billing system was simple and payment was determined only on the basis of the first diagnosis listed.

Federal regulations

Medicare requires specific documentation of the service for which payment is claimed by the provider. Software to handle this, and to achieve "bullet-proof bills," is available. This software tends to influence what is recorded.

Security fears

Concerns about confidentiality, exacerbated by the advent of electronic medical records, are – with or without justification – inhibiting the recording of essential information.

Technological Constraints

The electronic medical record is becoming a reality. Over 200 vendors are offering EMR products (sometimes called the computer-based medical record (CPR)). While the EMR offers the potential of carrying a vast amount of – and easily accessible – information, it can also serve to constrain the information which it "absorbs."

The Electronic Medical Record (EMR) – Are We There Yet?

There is no agreed-upon definition of an EMR, and the products range along a continuum of attributes. At the "low" end of the spectrum are what Peter Waegemann calls "automated medical records," which are essentially paper- based records (though roughly 50% of the information has been placed in the records by computers). At the other end are far more sophisticated products which have much more information, including laboratory reports, physicians' orders, drugs administered and known interactions, images of such record elements as electrocardiographic tracings, and other data. Some EMRs can collect the information from multiple sites, such as several hospitals and physicians' offices.

Soon the EMR will span a lifetime, and become known simply as the "electronic health record (EHR)." Everyone will probably carry a card, similar to a credit card (usually called a "smart card"), which carries their entire health history in electronic format.

The EMR can present information at the touch of a finger, with legibility and helpful display techniques. Time series of laboratory values, for example, can be shown graphically, abnormal values can be color-coded – features not practical with paper-based records. Patient education can be both triggered and provided within the sophisticated systems. Patient and physician reminder notices – health check points – can be generated automatically. Medical care can be made much more efficient and less error-prone.

Despite its many real advantages, adoption of the EMR has been slow, largely because it has required changes in the habits of physicians and other caregivers, and because the input methods have slowed down the care process (at least initially). Physicians have been loath to devote learning time. Time saved later in care – the promise of "delayed gratification" – has not been enough to overcome these problems.

For example, in the interests of making the EMR "user friendly," its designers are coming up with fast and easy ways for clinicians to record information. Perhaps the most common device is the "pick list," the computer-age version of the pencil-and-paper checklist. A pick list has a drop-down menu of "possible" terms. The user selects the appropriate term with a click of the mouse, and that term is then placed into the medical record. The pick list, however, has two inherent problems: (1) the terms which are on the list, and (2) the terms which are not on the list.

What is offered on the pick list. The only practical sources of terms for such lists of diagnoses have been the categories in *ICD-9-CM*, the "Disease" module of *SNOMED*,[3] and some proprietary "nomenclatures." Since it is *ICD-9-CM* which is involved in the payment system (because its codes are used in allocating cases to their *DRG*s for payment), all EMRs on the market appear to offer the *ICD-9-CM* list.

What is not on the pick list. It is impossible for any pick list to include *everything* a patient may present with. And, of course, it is almost inevitable that EMR vendors will in fact prune the list so that users are not offered codes which may give trouble in the payment system (whether they would be useful in describing the patient or not).

The physician can, of course, record what each diagnosis really is in a narrative note (by typing it in, or perhaps dictation). But this is more time consuming, and perhaps impossible to retrieve. It's much easier (perhaps, in some cases, mandatory) to use the list, even if it means selecting only the "closest thing."

[3] *SNOMED International: Systematized Nomenclature of Human and Veterinary Medicine*, College of American Pathologists, 1993.

Also, lawyers advise that nothing in medical record comments should be in conflict with the coded entries, because nothing is more damning than an internally "inconsistent" document.

Prognosis

If nothing is done to reverse this trend toward distortion of information, the medical record will lose more and more of its value for the physician and patient in the delivery of care to the individual, and will also lose more and more value as the source for many information needs in our society. We might even see the physician try to keep "two sets of books" – one for clinical care and one for reimbursement.

Flaw 2: The System Can't Capture Exact Diagnoses

A problem arose silently. No one called what we did "category coding" – it was simply called "coding" of diagnoses and procedures. We thought we were asking of a diagnosis, "What is its code?" – when we were really asking, "Where does it go?"

A category code tells us only where the case goes:

(1) its category, in

(2) a specific classification (e.g., *ICD-9-CM*)

It can never tell us more than that. It can only, very rarely, tell us what the diagnosis was.

Detail is lost, because the label (and code for) a category in the classification is substituted for the diagnosis which was recorded in the original record. Decoding brings back the category label; it can never bring back the detailed term. More than 10,000 diagnoses, for example, are crammed into 100 categories in *ICD-9-CM*. The total number of diagnoses in the universe of diagnoses is unknown, but the alphabetical index to *ICD-9-CM* has about 100,000 terms. Valid synonyms, local variations in usage, and expanding knowledge would increase that number greatly and continuously, perhaps by an order or more of magnitude.

The Guillain-Barré Story, Part I: What Caused the Epidemic?

In the 1970s, there was a sudden, sharp increase in the U.S. in the incidence of Guillain-Barré Syndrome (GBS), a rare type of paralysis.

Naturally, a cause was sought. Swine Flu vaccine had been recently developed and a large-scale immunization program was carried out. Investigators suspected the vaccine, and began searching for connections.

They ran into immediate trouble because Guillain-Barré syndrome had lost its identity during the coding process, when it was dropped into the category called "Polyneuritis and polyradiculitis." Since retrieval was via the code, the investigators could only find out whether there were or were not any cases in this group, not their specific diagnoses.

The needed specificity could only be determined by consultation of the original medical records, a task which would have to be carried out wherever the records were filed – that is, in each of thousands of individual hospitals, a task so enormous that it was never undertaken.

As a result of this experience, Guillain-Barré syndrome thereafter was given its own category, 357.0, in *ICD-9-CM* (and in *ICD-9* , its parent volume from WHO). But most diseases and diagnoses have not been so fortunate. And of course, this measure was taken only after the pig had left the pen.

The lack of specificity in coding is surprising, an aberration from the detail we demand elsewhere in clinical records. Drug administration, for example, is recorded as to the specific drug (not a group, such as antibiotics), along with its dosage, administration schedule, and duration of treatment. Under some circumstances, we can even find the manufacturing batch of the drug.

With patient age, we record the date and time of birth. Immediately after birth, age must be known to the minute, a bit later to the hour, then to the day, and only considerably later, to the year. It would be almost useless to record only the decade of a person's age.

The Gulf War Syndrome

In 1992, following the Gulf War, veterans complained of symptoms which they attributed to their war service. Much effort has been expended in an effort to decide whether or not there is actually such a condition, commonly called the Gulf War syndrome (or Persian Gulf illness).

Because of the medical debate on the issue, and because of the limitations of the coding system, no code has yet been offered for it (as of October 2000). Instead, coders are directed to transform the term Gulf War syndrome to the code for "Late effects of war" plus a second code to record the symptoms, e.g., "fatigue."

It will never be possible to reconstruct the cohort of persons who have presented with this possibility.

Lack of detail presents problems for both patient care and secondary uses.

Category Coding and Patient Care

In our traditional paper-based records, the exact diagnoses are recorded by the physician, either by writing them directly or by dictating and having them transcribed. The paper record can readily be consulted.

When electronic medical records replace the actual, detailed diagnosis with a category code, only that code's label is retrievable. And often the EMR offers only codes "immune" to challenge in the payment process. In these EMRs even the original medical record would not have the actual diagnosis – the clinical detail would be lost forever.

Category Coding "Cyst of the Pleura"

As an example of the this loss of information by category coding, "cyst of the pleura," a diagnostic term which labels a "clinical entity," is put into a category which transforms it into "Other diseases of the respiratory system, not elsewhere classified" (code 519.8). If this code were the total diagnostic information in the individual patient's record, as it might be in some electronic medical record systems, it would be of almost no help to the physician.

There are over 200 other categories of respiratory diagnoses – the physician would have to consult nearly 30 pages in *ICD-9-CM* to know what it (code 519.8) wasn't, and he would still not know what the patient's diagnosis was. And if someone wanted to search records to find patients with cyst of the pleura, the category code would return only whether or not there were any patients which fell into the "waste-basket" group where it would have been filed.

The developing electronic medical record (EMR) systems typically record only category codes. This diagnosis would, therefore, be absolutely unretrievable in a typical EMR, unless the physician had taken the time to record additional "comments" – and these would have to be laboriously searched.

Category Coding and the Bigger Picture

Every conclusion we reach from statistics developed from medical records will be futile if we gather only untrustworthy or useless information. We must make sure that the records tell the truth about the exact problems presented by the patients and give adequate detail about the care given. Without correct input, it will be meaningless to attempt to assess the successes and failures of our care, to accurately picture health and illness in the population, or to determine the resources needed to maximize health in the community.

Flaw 3: We Can't Tell For Sure What Any of the Codes Mean

Classifications change, and so do their categories and the contents of the categories. Biomedical progress is such that these changes cannot be avoided. Causes of diseases are elucidated. New diseases appear.

Progress in our knowledge of genetic abnormalities is an excellent example. In 1975, *ICD-9* provided 186 categories, nine of them for chromosomal abnormalities. In *ICD-10* (published in 1992, prior even to the vast explosion in knowledge in this field), the category number had grown to 709 categories, with 77 allocated for chromosomal abnormalities. No one has attempted to estimate what the counts would be if the classification were to be revised in the year 2000.

Why Category Codes Have No "Fixed Meaning"

As categories change, their "code" numbers often change with them. The categories are numbered (are given category codes) for the purpose of putting the them in the sequence demanded to make the classification orderly; they are actually "sorting" numbers. Naturally, they change when new sorting and sequencing demands it – it is no surprise that the same category codes are used for categories with different labels, different contents, or both.

How Classifications Are Created

The creation of a classification, say of diseases, is done by an "author" following a scheme. For example, the author decides that the desired sequence is: (1) infectious diseases, (2) cancers, (3) metabolic diseases, and (4) injuries. Here, these numbers in parentheses are the same thing as category codes. If a category for obstetrical conditions is then added, and the author thinks it belongs in the middle of the list, it may become number (4), which displaces injuries into being number (5). Neonatal conditions probably should follow the obstetric, and become number (5), with injuries now number (6). We have taken these sorting (sequencing) numbers to use as our codes – and so their meanings change, too.

"Leaving Room" Doesn't Work

The fact that there are inevitable changes in classifications is recognized and anticipated, so designers of classifications postpone the periodic major change by "leaving room" for new things. But this can't handle the kind of unpredicted explosion seen in genetics, for example, let alone changes in thinking as to how the various kinds of entities should be grouped. AIDS was first thought to be (and classified as) a disease of the immune system. Later it was moved into the infections disease section of the classification. Never has "leave room" provided a permanent solution to the "space" problem, nor has it solved the problem of the need for outright changes in where, in the schema of a classification, a given entity belongs.

"Decennial" Revisions of ICD

ICD is intended to be revised every ten years, but this time frame has not been generally followed. For example, it was 17 years between *ICD-9* and *ICD-10*.

However, modifications are introduced each year in *ICD-9-CM* in the United States. Although they are relatively few, their effect is not trivial.

Annual modifications introduced in October, 1998, for example, added 81 new codes to the roughly 13,000 codes in *ICD-9-CM*. Of these 81, 30 are "V" codes, adding to the Chapter entitled "Supplementary Classification of Factors Influencing Health Status and Contact with Health Services." Fifteen codes were deleted. The 1998 *ICD-9-CM* is truly a different classification than the 1997 *ICD-9-CM*, but the impact of such "minor" changes has never been recognized in our information system.

In describing these changes, a coding advisory publication warned the users to immediately make changes in their coding reference materials, but that carriers might vary in their dates of implementation. But of course, not all who should do so will even read the modifications list.

Trails Are Lost

It is impossible to maintain an information trail from one classification to another. Data once aggregated cannot be disaggregated. Statistics developed under one classification cannot be fairly compared with another, unless the grouping is so broad as to be of little value for most purposes – as, for example, "cancer" or "heart disease" or "injuries."

The Guillain-Barré Story, Part II: How Big Was the Epidemic?

Guillain-Barré Syndrome is not a reportable disease, so another source of information was sought. Its victims are invariably hospitalized, and a poll of hospitals was suggested as a method to to quantify the actual situation. Had there been an actual increase, and, if so, how large was it?

At the time, the code for classification of Guillain-Barré was *ICD-8* code 354.0, "Polyneuritis and polyradiculitis." A reasonable assumption might have been that changes in the size of the group of cases in this category would be a strong clue that the incidence of GBS had increased, although it would still have been necessary to "look inside the group" to be sure.

But prior to 1968, the classification in use was *ICD-7*, in which code 354 meant "migraine." After 1978, code 354 meant "Mononeuritis of upper limb and mononeuritis multiplex" (*ICD-9-CM*).

It was impossible to determine with any precision the incidence of the syndrome prior to the perceived increase, because there was no inherent way to tell which generation of a given classification had been used for coding – the codes all look alike. A coder could, without detection, have simply continued to use the 7th Revision even after the 9th Revision appeared.

In fact, the senior author learned that one developed nation had simply skipped the 8th Revision of *ICD* in its national statistics, going directly from the 7th Revision to the 9th (in violation of the international treaty to which it was a party). One can only speculate on the undetected effects of this violation on any international statistical comparisons in which it was included during that decade. For example, was its neonatal death rate really that low? Or were we comparing apples with oranges?

Flaw 4: Our "Single Classification" System of Statistics Can't Serve All Needs

At present, because we are using only category codes, we are restricted to using only one diagnosis classification for all purposes. But a single classification can't serve all masters. It is much better to use classifications specifically tailored to their purposes. Separate classifications should be used, for example, for billing, evidence-based medicine, policy making, planning, epidemiology. Plus, a researcher should be able to custom-tailor the statistics as optimum for the study.

Purposes of *ICD*

ICD originated as simply a classification to serve a single purpose – to present statistical comparisons of mortalities of nations for public health purposes. The first expansion, still for public health, was to include morbidity, which meant adding pigeonholes for non-fatal conditions and events.

Even these early uses of *ICD* reflected competing demands on the classification: was the purpose to classify the causes or the effects of processes? Infections and trauma, especially, were of concern because they often could be prevented – an argument for the causes. Infections could be classified by organism or by disease. For trauma, an "External Cause" chapter was provided, so that an injury could be classified under both its physical effects and its external cause.

Then came hospital indexing, followed by efforts to make *ICD* useful for a wider array of activities, such as quality measurement, billing for care, establishing health policy, studying office and ambulatory care, facility planning, and evidence-based medicine. Not only did these changes in the organization and content of the classification complicate the construction of and learning to use each successive generation, but they also made the product – the classification itself – less useful for any of the different purposes. Common sense would suggest that a classification for billing would have different attributes than one for quality assessment, which in turn would differ from one designed to compare health and healthcare among nations.

Using Only One Classification – and Coding Process – Has Enormous Consequences

If we stick to only one kind of coding – category coding – we must change the entire system periodically, with serious results. The effects of making the change are far greater than most realize:

- Information quality slumps
- Continuity is lost
- Large amounts of money are spent

Information Quality Slumps

Decisions based on faulty information are one serious result of periodically changing classifications. With our present, single-code coding system, a change from one classification to another causes a serious deterioration of information quality during the changeover period, deterioration which typically lasts two or three years.

The introductory period of any new classification creates less accurate information, because any new process has to be learned and corrections made in light of actual experience. Less apparent are the problems of timing of the adoption of the new system. In the hospital, for example, are the new codes to be used for patients on the basis of their admission or their discharge dates? Individual institutions will be ready for the change at different times.

Intermediaries and carriers have similar readiness problems. Their solutions and responses are difficult or impossible to police or trace. In the 1960s some carriers, for example, had not allowed for enough digits in the diagnosis fields in their databases for the longer codes used by *ICDA-8* and *H-ICDA* as compared with their predecessors. Without detection, some of the carriers and intermediaries simply truncated the incoming data (a situation similar to the Y2K problem, and to the problem noted above of the country which simply skipped one revision of *ICD*). Many of these problems have widespread impact.

Continuity Is Lost

Another substantial effect of a category coding change is on the research which can be conducted and the validity of inferences which may be drawn from the data. Vitally important longitudinal studies – those which span two or more classifications – often cannot even be attempted. It is essential that we know what is going on in health and in healthcare over periods longer than the nominal decade between revisions of *ICD*. Are genetic defects really increasing or declining? We might be able to judge a question that broad if we could aggregate all the kinds of genetic defects. But we may be completely thwarted if we wish to pinpoint the specific kinds of defect which are increasing or declining. And the explosion in knowledge in this field compounds the issue.

What Happened With AIDS

AIDS presents a real-life example of the longitudinal study problem, even within a single generation of the classification (*ICD-9-CM*). AIDS was first described in 1981. There was no agency to which coders could turn in order to find where to classify it so, in order to keep the paperwork flowing and reimbursement coming in, they simply made their own decisions – and we don't know where they put it.

In October 1982, an official coding decision was published in the Journal of the American Medical Record Association (of course, we don't know how many coders read this issue). The Journal gave, as preferred, code 279.19, "Other deficiencies of cell-mediated immunity," and thus it was put into a "waste-basket category" – not a classification specific to AIDS. Alternative codes were also given: 279.10, "Immunodeficiency with predominant T-cell defect, unspecified," and 279.3, "Unspecified immunity deficiency" (more waste-baskets).

Because of these classification decisions, prior to October 1982 retrieval of AIDS cases would have required searching countless classification pigeonholes, and even after that date, one would still have had to go back to all of the original medical records having any of the three codes suggested.

Unique coding for AIDS did not begin until late in 1986, when an addendum to ICD-9-CM was issued assigning the unused codes 042, 043, and 044 to AIDS – a jump from the " – Immunity" chapter to the "Infectious – " chapter in the classification. This addendum contained nearly five pages of instructions as to how to classify AIDS cases, AIDS-like syndromes, and HTLV-III/ LAV infections, taking into consideration various symptoms and "with" and "due to" relationships with other conditions. Further changes occurred two years later when the nomenclature of HIV infection and the diagnostic criteria were revised.

The delay in uniquely identifying AIDS in the data system was due in part to debate over just how to classify it, but also to administrative and bureaucratic factors – in the words of one nosologist, a "political/resource consumption quagmire."

Large Amounts of Money Are Spent

Making the change is very expensive.

Outside the immediate healthcare setting – hospitals, clinics, physician's offices – a sizeable industry has grown up for teaching and supporting healthcare coding. A new classification requires this industry to develop training materials and training facilities, and provide consultation and reference services. New coding software for use in clinical sites must be written, installed, taught, and supported. A tremendous amount of resources go into this (raising the costs of coding and, therefore, of healthcare).

Then there are far-reaching effects for the coders themselves. Replacing one entire classification with another requires many steps. Preparation of the coders usually takes time away from productive work, during the educational process. Meanwhile, the daily coding with the old system must go on. Something must be compromised. There is an inevitable decline in productivity as familiarity is being gained in actual coding.

And perhaps greatest of all, the change disrupts the reimbursement/payment system; this is discussed further below.

The Solutions – How to Live Happily With "Code-Dependency"

Flaw 1: "The System" Encourages Distortion of Information

Solution: Support the Truth

It's time for our society and the healthcare industry to support and encourage the truth.

We have to set up the reimbursement system in such a way that frankness is not "punished." It is essential that payment accurately covers necessary care, so that no one is tempted to "fudge" the information. People should be rewarded for doing the right thing.

Measures taken to "ensure" honesty – such as the federal regulations requiring specific information supporting payment – must be carefully considered. First, are they really necessary? If so, how can the situation best be handled so as not to perpetuate the very "evil" we're trying to avoid?

These are difficult problem areas, but they must be addressed.

Luckily, the other constraint on information – technology – is much easier to tackle. We can design the electronic medical record so that it does not force expedient codes at the cost of information distortion. We can find ways so that, in fact, we get much more, and much more accurate, information than ever before. The other solutions proposed below will directly contribute to our meeting this challenge.

Flaw 2: The System Can't Capture Exact Diagnoses

Solution: Introduce "Entity Codes"

Flaw No. 2 – lack of clinical detail – is due to the fact that we presently have no way to capture individual diagnoses in our code-dependent system. Since retrieval of clinical information for both patient care and secondary uses is almost exclusively via codes, it should be clear that, if clinical entities such as exact diagnoses are ever to be retrieved, the clinical entities themselves must be coded.

We need to add to our medical records a feature called "entity coding." Entities are defined as the "most detailed information available" – for example, diagnostic entities are the exact diagnoses as recorded by the physician. If the diagnostic entity is as crude as "heart disease," so be it. If the record contains anatomic detail as to the branch of a coronary artery involved in an occlusion, that level of detail is also an entity.

Entity coding tells us what the original diagnoses are, and does it with permanent, unalterable codes which, when decoded, give us the physician's statements verbatim.

Implementing entity coding requires adding a new "front end" to our medical record system to incorporate the necessary detail. A master database of clinical diagnostic entities and codes is needed, and the medical record information system must be modified to properly accommodate them.

The "Entity Coding" system must have the following essential features:

- The database of entities and their codes must contain all diagnoses, no matter how long the list.
- There must be only one master database (list) of entity terms and entity codes for each type of

information (diagnoses, for example). Competing lists would defeat the entire purpose. We might entitle the master list for diagnoses, "United States Standard Diagnosis Entity Codes"

- Each diagnostic entity must have its own permanent, unalterable, never-to-be-reused code.
- Each code must be tagged, just as the category codes will be tagged, to identify it as an entity code (this tag might be as simple as "USSDEC").
- The system, once begun, must remain permanently available.
- Entity codes must be instantly available to any coder.

We don't have an entity coding system yet – the tool has not been built. A master database of entity codes must be compiled. Software must be written, and technical details solved, so that entity coding can work. "Someone" must create and service the entity coding system. For lack of a better word, let's call this institution the "secretariat."

The first job of the secretariat would be to establish the master diagnosis entity database. Populating it would initially be done by simply taking the diagnostic terms found in the various classifications in use (*ICD-9, ICD-9- CM, ICD-10, ICD-10-CM, SNOMED, MESH*, Read Codes, and others), and from textbooks and standard publications as has been done with its *Unified Medical Language System (UMLS)* by the National Library of Medicine. More than a million terms would result from such an initial search. This starting list would be relatively easy to compile, and it would be expected to handle a very high percentage of the demands.

Once initially populated, the master database would simply grow by acquisition of new terms as they appear in the literature and as they are presented to the secretariat by coders requesting entity codes. Each diagnostic term encountered would automatically be an "entity," with no argument. It would be entered in the master entity database, and it would be given its own unique, permanent "entity code." This would basically be an accession number (the next number).

It is essential to capture the language used by the clinician, which of course varies with the physician's medical school, graduate training, cultural background, and geographic area of practice. Biomedical progress is often enhanced by the fact that various investigators and research centers use their own terminology, reflecting subtle differences in concepts. And history tells us that, except in highly controlled and motivated situations (such as in academic medicine and specific research projects), the clinician cannot be constrained to use a "prescribed" language.

No one should even try to tell the physician (or patient, or anyone on the team) what to say. The entity code must be essentially just an accession number. No effort should be made to decide whether or not the term is the "preferred" term, a controlled term, a synonym, an alternate spelling, a new entity, or even whether it is a "legitimate" term at all. The secretariat should simply capture the terms actually used in medical records and assign codes to them. It should not make judgments about them.

To compromise on this issue and expect entity coding to change the language actually used in healthcare would, we contend, prove fatal to its implementation.

The Language of Healthcare Can't Be "Prescribed"

Attempts to control clinical language – for example, those of the American Medical Association with *SNDO*, and later with its *Current Medical Information and Terminology* – have been abandoned. Other "languages," such as *SNOMED,* have never been widely adopted except within such settings as the pathology laboratory – and even here the changes in codes with successive generations of *SNOMED* have been frustrating.

However, *SNOMED*'s "compromise," the provision of a module entitled "*Diseases/Diagnoses*," is being offered by some vendors of electronic medical records as an answer to the specificity problem. But that module offers (1999) only about 41,000 terms (and 4,000 of these are for veterinary use only). This isn't even close to the millions of terms used in real clinical life.

At the same time the master database is being created, we can prepare the medical record – both paper and electronic formats – to incorporate tagged entity codes, as well as tagged category codes.

Today's medical records, electronic as well as paper-based, are only prepared to handle, as codes, the category codes (dictated by reimbursement). In the case of paper-based records, these codes are recorded for transfer to the bills. In the EMR, there may be only the codes, without any supporting language. (Except as that language is generated "in reverse," because the computer knows what the label is for the code recorded, and it can print that, or a truncated version of it to fit the space allowed. Truncation usually, of course, further distorts the meaning.)

System designers for both paper-based and electronic records will have to accommodate the needed flexibility of input. This requires physical space, in both the paper-based and electronic records – but this is a technical problem which can be solved. In the individual medical record, there is likely to be essentially a one-to-one correspondence between a given entity and the category into which it is to be classified. The fact that a category may contain ten or a hundred entities does not mean that each patient has that many entities; the likelihood is that one patient has only one entity belonging in one category. Thus the additional space demand in the medical record to accommodate entity codes would be minor.

The other major task is "deployment" of the codes (and coding system). This will initially involve publicity to announce the concept, question-and-answer type sessions, and then specific education and training for those using the system.

Finally, getting the codes to those who need them will be almost a "snap." Access to the master list can be in hard copy and CD-ROM. Terms not available in those formats could be sought via an "800 number"or the Internet (where, on a web site, the list is updated in real-time). Instant response is possible because either the entity would have been encountered earlier, in which case it would already have its code, or it would be accepted and given its code without any delay, since the code will be just the next available number on a list of code numbers (it has no inherent meaning).

Why Category Coding Can Never Be Updated Immediately

In today's category coding system, there is nothing resembling an instant response service for answering coding questions. When an entity new to the system (or a configuration of entities for which there is no precedent) is encountered, an "official ruling" often takes months or years (see the examples of AIDS and Gulf War Syndrome).

Even though, of course, with category coding, the decision is only where to put the new entity along with others. Rarely is an entire new category offered for a newly encountered entity – this would require even a slower decision.

Entity coding does not need any decision on the classification of an entity before it can be coded and placed in the information system (both in the original medical records, and in the secondary records derived from them). Classification decisions require a separate step.

Of course, coders will have to learn the new "front end." But entity coding has a distinct advantage over category coding. To do category coding, one must learn to classify – a process requiring a good deal of knowledge, training, experience, and judgment. With entity coding, the coding process – substituting a code for a term – is essentially as simple as looking up a telephone number or calling for directory assistance.

Over time, we may expect the healthcare information system to evolve to the point where the category coding (the classifying process) is delegated to the computer. The computer would take the entity codes for a given case, pick up the other information needed from the rest of the medical record, and classify the case. Classifying would become more uniform, and costs would be lowered. Most importantly, health information professionals – not having to deal with the "quagmire" of coding anymore – would have the time to practice their profession, concentrating on the uses of information.

Flaw 3: We Can't Tell For Sure What Any of the Codes Mean

Solution: "Tag" All Codes

It is time to add to every code in all our code series, both those series coming from classifications and from the entity code master lists, a tag which tells:

- The name of the Classification ("code series") to which it belongs, e.g., *ICD-9, ICD-10*, "United States Standard Diagnosis Entity Codes" (USSDEC), etc.,
- The specific Version of the classification, e.g., WHO or United States, and
- The Series, e.g., year of modification or correction

Once a category code is combined with its tag, that "code + tag" combination will become unique, and we will always be able to accurately decode it.

With the *ICD* series, the tag would tell at least the *ICD* and Revision number + "Version" (e.g., Original from Geneva or U.S. Clinical Modification) + Series (e.g., "1998").

Thus tags could have solved the swine flu vaccine incidence problem in the earlier example, in which the investigator was trying to find at least which category would contain the condition. The category in which to find Guillain Barré Syndrome might have been given "code + tag" combinations as follows (note: No such tagging system exists at present):

Example of Possible Code-Tagging

Tag Key Components

Classification – requires both Basic name and Revision number
- Basic name of classification = *ICD*
- Revision = e.g. "7"

Version
- WHO original (Geneva) = G
- US Clinical Modification = CM

Series (when applicable, as in CM)
- Year = 1978

Decoding The Tagged Codes

[Tag] **ICD-7-G** + [Code] **357** =
"Other diseases of spinal cord" (this code contained Guillain Barré Syndrome (GBS))

[Tag] **ICD-8-G** + [Code] **345** =
"Polyneuritis and polyradiculitis" (which then contained GBS)

[Tag] **ICD-9-CM-1978** + [Code] **357.0** =
"Acute infective polyneuritis" (which represented only GBS, but which retrieved the preferred term rather than the term entered)

Permanently identifiable, unique codes is not a new idea. The number 355-34-3222 means nothing, until we prefix it with SSN to denote a social security number. If we order a car part, we usually have to give both the car model number (MN) and the part number (PN), and sometimes even the SN – serial number. And the book world depends on the ISBN number.

ISBN numbers

The International Standard Book Number (ISBN) is used internationally to identify each edition (and within that edition, each binding) for every book published. It is a unique, unambiguous code.

Each number identifies just one book, forever. The number includes "ISBN" as a prefix to denote the kind of code it is.

The number itself contains some information about the book. For example, the number 0-915516-21-7 tells us:

- 0 - the book originated in an English speaking country
- 915516 - identity of the publisher
- 21 - the particular title and edition (hard or soft cover) of book
- 7 - a check digit, used to verify the accuracy of the other digits.

If an ISBN were – like a category code – not unique, it would be useless. We would not tolerate a book identification system, for example, in which an ISBN number pointed to a shelf (or shelves!) full of books rather than to a specific volume. Nor could we be expected to buy all the books on the shelf, just to get the one we really wanted.

Flaw 4: Our "Single Classification" System of Statistics Can't Serve All Needs

Solution: Use "Tailored" Classifications

We must work toward changing the "system" so that cases can be allocated to more than one classification – each tailored to a specific use.

Especially important would be classifications truly appropriate for billing under a variety of circumstances, e.g., inpatient vs. home care. Our acknowledged problems with *DRG*s are primarily due to the compromises which had to be made in *DRG* design because of the limitations of *ICD-9-CM*, the only available source of diagnosis input at the time the *DRG*s were created.

We can't have multiple classifications in the "system" today because, with the data already aggregated (into the *ICD-9-CM* categories), it can only be placed in broader, rather than different, groups. It can never be disaggregated so that one can start over and classify it another way.

With entity coded information available, however, each case can be coded to any classification, past, present, or future.

Wouldn't it be great if –

– we had classifications which were:

(1) broad, to be used for public policy,

(2) clinically-based, for evidence-based medicine,

(3) etiologically oriented, for epidemiologic surveillance,

(4) "product line," for facility planning,

(5) custom-made for your needs ?

The Coding Situation Today

The National Center for Health Statistics (NCHS), responsible for disease classification in the United States, has been working under the assumption that, as for the past forty-five years, our healthcare system would continue to use each successive *ICD* Revision as the basis for coding diagnoses – that category coding would be the way we would continue to handle the diagnostic information in and from medical records.

In 1989 WHO completed, and in 1992 published, *ICD-10, International Statistical Classification of Diseases and Related Health Problems, 10th Revision. ICD-10* is a much larger set of volumes than *ICD-9*; growing to about 2,100 pages from 1,400. Many conditions were added, others were moved from one Chapter to another, and some categories were modified. One-to-one correspondence between the categories of *ICD- 9* and those of *ICD-10* does not exist. Only a concordance (which also does not exist) would permit tracking a diagnosis from one version to the other. All category codes in ICD-10 start with an alphabetic character, including "I" and "O," which are ordinarily avoided in codes because of their confusion with the numerals "1" (one) and "0" (zero).

NCHS examined *ICD-10* and determined that, like *ICD-7, ICD-8,* and *ICD-9* before it, *ICD-10* would have to be modified for clinical use in the United States. As a result, a new clinical modification was created by the United States National Center for Health Statistics: *International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM).*

A Look at the Clinical Modification of *ICD-10* for Diagnoses

ICD-10-CM, the volume the United States government has produced for diseases, is a huge expansion over *ICD-9-CM* . In October 1998, *ICD-9-CM* contained 12,628 categories. *ICD-10-CM* in its Internet draft has about 60,000 categories.

The bulk of the growth seems to have been in the widespread adoption of "combination categories." Combination categories use early computer theory in direct opposition to the more modern relational database theory. For example, today's theory would have one database giving the site of a disease (e.g., stomach) and a second database giving the condition (e.g., ulcer). Conditions and sites would only be linked dynamically when appropriate. A search for "stomach" problems could be carried out with a single look at that code.

In combination coding, each condition for which site is important has the site put right into the code for the condition. Thus a search for stomach problems requires a search of every code where the element "stomach" might have been included in the code itself. It's hard to overstate the complexity this introduces into the coding process as well as the retrieval process. For example, look at the following from *ICD-10-CM:*

"S02.974" is the code for

"Open fracture of skull and facial bones, part unspecified with subarachnoid, subdural, and extradural hemorrhage with prolonged [greater than 24 hours] loss of consciousness, without return to pre-existing conscious level, or when an unconscious patient dies before regaining consciousness, regardless of the duration."

Simultaneously, the Health Care Financing Administration (HCFA), responsible for procedure classification, determined that the U.S. needed a new procedure **coding** system. Neither *ICD-10* nor *ICD-10-CM* contain codes for procedures. So HCFA created a brand new product, and named it *ICD-10 Procedure Coding System (ICD-10-PCS)*, with "ICD-10" in the title, although its only relationship to *ICD-10-CM* is temporal.

The New Codes For Procedures

The procedure coding system commissioned by HCFA – *ICD-10 Procedure Coding System (ICD-10-PCS)* is exactly that. It is not a classification, as was found in Volume 3 of *ICD-9*, and generally in *CPT*[4] and *HCPCS*[5], nor is it related to the classification *ICD-10.*

ICD-10-PCS is an extremely complete but complex system which, like *SNDO* and *SNOMED*, requires the coupling of elements from several modules to form the code (with its nomenclature) to describe a procedure. There are 7 lists (modules), each with a potential of 34 alphanumeric entries (unlike *ICD-10* , the letters "I" and "O" have been omitted to avoid confusion).

Not all of the 52 billion possible combinations have been used, but some 200,640 codes are available for the male reproductive system, 482,000 for the female.

"Appendectomy" in the Alphabetic Index leads to the Tabular List, where one finds it possible to code 9 kinds of diagnostic partial appendectomies, 9 kinds of therapeutic partial appendectomies, and 9 kinds of complete appendectomies: 27 in all. Code "0DBJ6ZX," for example, means:

[medical and surgical = 0]

percutaneous

intraluminal

endoscopic

diagnostic

partial

appendectomy

Presumably this is a valid code, because the *ICD-10-PCS "Working Paper"* states that "combinations of characters that do not constitute a valid procedure are not contained in the Tabular List."

Despite the huge number of specific procedures which may be coded, *ICD-10-PCS* still must rely on wastebaskets such as "other device" to accommodate future inventions; e.g., a new procedure which uses tiny robots does not have a code.

The original plan reportedly was to exchange *ICD-10-CM* for *ICD-9-CM* by 2001. The Department of Health and Human Services (DHHS) has now published its *Final Rules for Health Insurance Reform: Standards for Electronic Transactions.*[6] With regard to the coding of diagnoses and procedures, the rules call for the continued use *ICD-9-CM* for diagnoses and *CPT-4* for procedures. So implementation of *ICD-10-CM* has been delayed indefinitely as have any changes in procedure coding.

[4] *CPT – Current Procedural Terminology*, the American Medical Association.

[5] *HCPCS – Healthcare Financing Administration Common Procedure Coding System*

[6] The Rules, required by HIPAA, the Health Insurance Portability and Accountability Act, were published in the *Federal Register* on 17 August 2000. The Rules are effective 1 October 2001.

Problems with Switching from *ICD-9-CM* to *ICD-10-CM*

The delay in changing diagnosis and procedure coding systems is a good thing, because every such change – where one category coding system replaces another at the point of origin of information, the medical record – is very costly. As many as perhaps 500,000 coders must be retrained. Software must be rewritten. Information quality suffers. Longitudinal studies in particular become difficult or impossible. And in today's healthcare scene, the most serious impact is on the reimbursement system.

How a Category Coding Change Affects Our Reimbursement System

Billing in the United States is "prospective," in that prices are established for groups of patients ahead of time. The grouping for hospital inpatients is called *Diagnosis Related Groups* (*DRG*s), about 475 groups into which patients are fitted, primarily on the basis of their *ICD-9-CM* codes. All patients within a given *DRG* have, by definition, roughly the same resource consumption, so each patient episode falling within the *DRG* should be worth about the same payment. To develop the proper values for *DRG*s, the resource consumption is approximated by the hospital cost.

*DRG*s were originally defined from cases coded with *ICD-8* (modified), at a time before hospital reimbursement was dependent upon the prospective payment system (PPS). When *ICD-9-CM* was introduced, series of cases coded under *ICD-8* were also coded under *ICD-9-CM*. This "duplicate" information, along with the accompanying cost information for each case, was provided to the researchers and the effects on the dollar values of each *DRG* were studied. Significant differences were found, and the DRGs got new definitions based on *ICD-9-CM* rather than *ICD-8 – DRG*s were recalibrated.

Only after the comparative cost study under both *ICD-8* and *ICD-9-CM* was completed, and the necessary adjustments made, were *DRG*s given the green light to govern payment for hospital care.

Now there has been a dramatic change in the situation.

Hospitals are critically dependent on *DRG*s as major determinants of their incomes. Medicare, Medicaid, and most other payers are dependent on *DRG*s in planning their expenditures.

Prior to the introduction of any change, another study must be made, just as with the change from *ICD-8* to *ICD-9-CM*. There should be a study of statistically significant numbers of the same cases coded to both classifications, the financial impact measured, and the payment groups redefined as necessary before the nation is asked to shift. Needless to say, such a study would be both time consuming and very costly. But without such a study, financial chaos is almost a certainty.

Entity coding and code tagging offer an alternative to continuing with category coding

If it proves possible to implement entity coding and code tagging prior to a move to replace *ICD-9-CM* with *ICD-10-CM,* all of these bad effects can be avoided. The reason is that the classification process – the placing of cases into the categories of the classification – can be made a computer process, "entity-code classifying,"exactly as the determination of the *DRG* in which each case belongs is today done by computer. In other words, the category coding to *ICD-9-CM* (or *ICD-10-CM*) can then be done by computer just as the *DRG* coding is today done by computer.

Placing cases in *DRG*s is done by a computer program called GROUPER which, based on the *ICD-9-CM* codes and other information about the case, determines the proper *DRG*. This is done uniformly for all cases in the United States.

In exactly the same fashion, given entity codes as the building blocks, a computer program can be written which places the cases uniformly in the proper categories of *ICD-9-CM*. A similar program can place the same cases in *ICD-10-CM*. And programs can be written which place the same cases in any classification desired, including classifications specifically "tailored" for such disparate needs as development of public policy, facility planning, and evidence-based medicine.

In fact, the healthcare system might avoid the costs of implementing *ICD-10-CM* at the coder level entirely, because once entity coding has "incubated," and standard healthcare code identification is a reality, we may decide to take a different route. For example, it might then be feasible to develop a classification specifically designed for reimbursement, as well as the other "tailored" classifications mentioned above.

Thus the transition from one generation of *ICD* to another would cause no burden or expense on the handling of medical records. And the adverse effects on information and reimbursement of inflicting a category coding change at the point of origin in the system would be entirely avoided.

What needs to be done.

We must remedy flaws 2 and 3 – the "coding" flaws – as soon as possible. This will go a long way in helping to cure the other flaws. The remedies will require us to:

1. Tag all codes.
2. Develop and implement entity coding.

Obviously, creating and implementing these solutions, especially adding entity coding to the medical record and its information system, cannot be accomplished overnight. It will take several years and a good deal of developmental work. And, until entity-code classifying is available, providers will be put to extra expense to put the entity coded "front end" into their record systems. The steps required are:

- Mount a careful project to create the basic entity coding tool – the master database of entity terms and their codes. This would require close collaboration and "beta testing" in real life clinical settings.
- Simultaneously prepare the medical record systems to carry the entity coded data. Computer-assisted encoding programs should find adding entity codes relatively easy – they must already "recognize" many of the entities as they decide their category assignments. Vendors of electronic medical record systems should welcome the added information their products, with entity coded-diagnoses in the records, would provide their customers.
- Implement entity coding as rapidly as the healthcare system finds it desirable.
- Begin developing the algorithms for entity-code classifying.
- Make sure all codes used in healthcare are "tagged" for identification; establish standards for the tags to be applied to all codes in the system.

To be ready for the future, we can also undertake a study of the whole health information system in the light of current information technology and the opportunities which will be made available with the corrections we propose. We have time to study the needs and how best to meet them.

A Great Opportunity

The coming of *ICD-10* is serving to focus attention on our health information system, at a time such attention is sorely needed. It is also providing a wonderful opportunity to enhance the system. A system with entity codes as well as category codes can better meet our needs, and save money. The addition of entity coding will allow us to comply with one of the basic principles of information management:

Keep data in its most discrete form.

From the discrete "building blocks" it can then be aggregated in any desired way. Entity coding keeps discrete information discrete. Category coding, although essential for its purpose, can never adhere to this rule, because, by its very nature, it aggregates data.

Once aggregated, data can never be disaggregated.

Entity coding will permit the electronic medical record to have the flexibility needed to accommodate a vast range of information, allowing it to live up to its potential.

"Tagging" both entity and category codes will permit us to comply with two other basic principles:

Never change the meanings of the codes

Be sure you can tell what the codes mean

Even if a classification changes the meanings of its codes, the tags will ensure that the "code + tag" combination will never change meaning, because it has become unique. We will always be able to find out what the tagged code means.

If we take advantage of the opportunity offered at this time, there is much to look forward to:

- Classifications custom-tailored for reimbursement under various circumstances would become a welcome possibility.
- Changes in international demands for statistics and in clinical knowledge reflected in succeeding generations of *ICD* would never distort the information in medical records.
- In a reasonably short time, a substantial saving in information management cost could be achieved.
- We can rely on a much sounder base of health information.

Our medical records can get back to their fundamental business of helping in medical care, and our national knowledge of health and healthcare can become trustworthy.

Vergil Slee, MD, FACP, FACHE (Hon) is an internist who developed the Professional Activity Study (PAS), the first hospital discharge abstract computer system. He founded the Commission on Professional and Hospital Activities and served as its Director and President from 1956 through 1980. He was President of the Council on Clinical Classifications which wrote, in collaboration with the National Center for Health Statistics, the *International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM),* the disease and hospital procedure classification now used in the United States. Dr. Slee is a co-founder and Chair of the Board of the Health Commons Institute, Portland, Maine. He is senior author of *Slee's Health Care Terms* (4th Edition, Tringa Press, 2000) and of *The Endangered Medical Record: Ensuring Its Integrity in the Age of Informatics* (Tringa Press, 2000).

16 Udoque Court, Brevard, NC 28712 828-884-6508 – vslee@juno.com

Addendum (1 Nov 2001)

Florida Discovers False Increase in HIV/AIDS Deaths Due to Coding Changes

The foregoing White Paper describes problems with AIDS statistics, showing how changes in definition and classification rules have distorted the picture over time, resulting in invalid longitudinal studies (page 15). Cases have been added to and subtracted from the AIDS population, and to its mortality rates, solely in response to "rules" as to who was to be counted. A recently published Florida study demonstrates this problem.

The HIV/AIDS mortality rate reported for Florida unexpectedly increased in 1999, reversing the national and state trends for previous years. Florida's rate had declined each year, beginning in 1966. The Florida Bureau of HIV/AIDS suspected a coding effect rather than a failure of prevention or care.

Effective January 1, 1999, the United States changed from coding (classifying) death certificates using the International Classification of Diseases, 9th Revision (ICD-9), published by the World Health Organization in 1975, to the International Statistical Classification of Diseases and Health Related Problems, 10th Revision (ICD-10), published in 1992.

The Florida Department of Health went back and coded again all 1999 deaths which had any mention of HIV/AIDS, this time using the ICD-9 rules. As a result, 81.1% (1445 cases) were allocated to HIV/AIDS, compared with the 92.6% (1651 cases) which had been so allocated with the ICD-10 rules. ICD-9 rules would have resulted in a 6.6% decline in the death rate, continuing the previous trend, rather than the 6.7% increase in death rate which had been reported under the new, ICD-10, rules.

[Grigg, et al, "Research Letter: "Coding Changes and Apparent HIV/AIDS Mortality Trends in Florida, 1999," JAMA, Vol. 286, No. 15, page 1839, October 17, 2001]

The cost of this recoding effort was not reported, but it was substantial. More important would be effects of the misinformation on HIV/AIDS programs. And the statistics for other states and for the nation were, of course, equally suspect.

In fairness, the United States National Center for Health Statistics is aware of the nature of the problem of comparing statistics compiled under changed rules, and has attempted to address it by studying a sample of cases from 1966 death certificates "double-coded" to both ICD-9 and ICD-10, and from the two sets estimating a "comparability ratio." This yielded a ratio of 1.06 for HIV/AIDS, whereas in the Florida data on 1999 deaths the actual ratio was 1.14. But any comparison of data which has been compiled under different classification rules can only be an approximation; a true comparison requires going back to the "records of original entry," here the detail in the death certificates, as was done in Florida, rather than depending on a "mass correction" using an estimate.

Requesting your health records demystified. How to Request Your Medical Records? Easily request your health records with our online form.Â Records that may get denied: Psychotherapy notes. Information that could endanger the physical safety of the patient or another person. Information compiled for use in a lawsuit. Information that is part of a research project while it is still in progress. What are medical records? Although some NHS patients may have all their healthcare information at their fingertips, others may be grateful to know that their healthcare information is centrally recorded and can be accessed by different healthcare professionals looking after them. This is especially important for anyone on lots of different medication, with long-term or chronic conditions and those for whom English is not their first language. If you're registered with a GP in England details of your health history is kept on an electronic file. If you are registered with a GP practice in Medical records contain valuable patient information. Medical professionals have obligations to keep these records in a confidential manner. Additionally, they are also obliged to ensure that records are legible, accurate, and that the documentation is presented in an orderly fashion. Doctors may also find that there are times when they are legally obliged to release confidential information contained in medical records. This chapter also looks at the care.data medical record IT system in the National Health Service. Organizations that are responsible for medical records are supposed to have a The Electronic Medical Record (EMR) is essentially a cash register. It was developed by technocrats as part of a mandate of the Obama administration in 2008, to help make medical records more efficient. It was a good idea: to make all clinical data from a patient's medical history readily available electronically to doctors and other health care workers. It would have worked, if it were used only for that. But somehow the for-profit insurance industry got into the EMR, and linked the medical data part tightly to the money partâ€”through billing.Â Even more important are the human costs. These costs endanger the health of both you and your doctor. The most widely-sold EMR system is called Epic. It is so unpopular that if you mention the name, some doctors will literally start to scream.