# Medicine in the Information Age

Clinical Information Systems

Internal Medicine Grand Rounds
University of Texas Southwestern Medical School

July 11, 1996

W. C. Yarbrough, MD

It is simply unrealistic to think that individuals can synthesize in their head scores of pieces of evidence, accurately estimate the outcomes of different options, and accurately judge the desirability of those outcomes for patients. . . All confirmed what would be expected from common sense: The complexity of modern medicine exceeds the inherent limitations of the unaided human mind.

D. M. Eddy (1)

...a good man knows his limitations.

Clint Eastwood, Magnum Force

*"Capable of solving scientific problems so complex that all previous methods of solution were considered impractical, an electronic robot known as ENIAC - Electronic Numerical Integrator And Computer - has been announced by the War Department. It is able to compute 1000 times faster than the most advanced general purpose calculating machine, and solves, in hours, problems which would take years on a mechanical machine. Containing nearly 18,000 vacuum tubes, the 30 ton ENIAC occupies a room 30 x 50 feet."(2)*

This announcement occurred 50 years ago last month. Next month, we will witness Olympic competitions which measure human physical capabilities. This Grand Rounds will concern our mental capabilities and how to extend them for the benefit of our patients. I quickly developed information overload during the preparation of this Grand Rounds, so I have limited the scope of this paper to clinical information systems, mostly as they relate to the Intensive Care Unit (ICU), and in particular those aspects that aid the physician most.

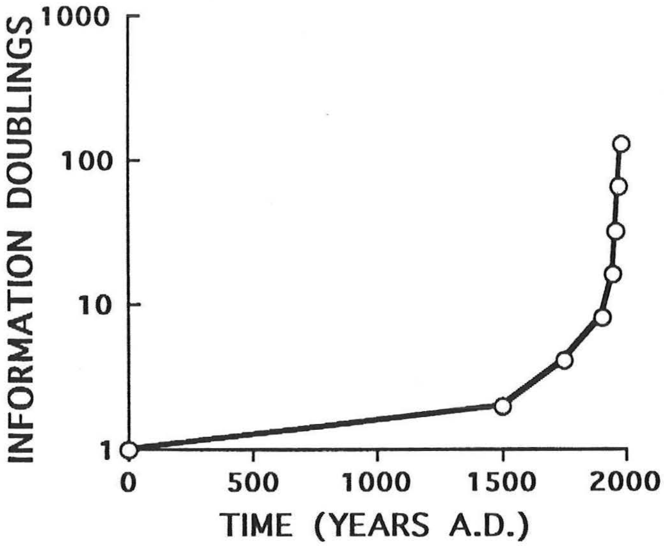
Since the time of the ENIAC, computational speed has increased several hundred fold. The size and energy consumption of computers have decreased by 10,000 times. If the aircraft industry had evolved as rapidly as the computer industry, a Boeing 767 would cost $500.00 today, would circle the globe in 20 minutes, and would do it on 5 gallons of fuel. It seems apparent that, as computer cost drops and critical care complexity increases, computers will occupy a more central role in intensive care units.

## Information Growth

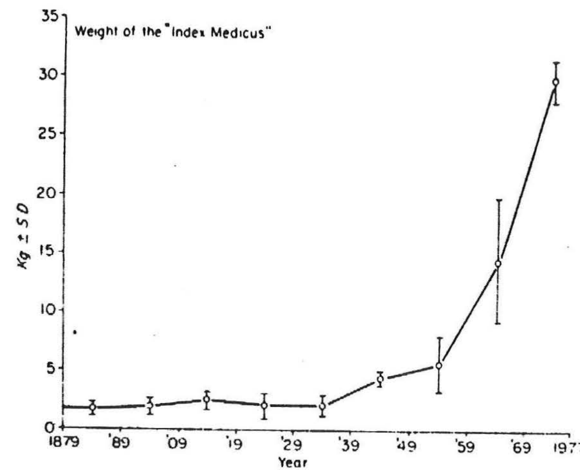
Method advances in the biomedical sciences have led to a rate of growth in medical knowledge that far outstrips any previous historical period. The first volume of Index Medicus contained approximately 17,000 citations from 700 periodicals. Today a single month's volume has an average of 21,000 citations from over 3000 journals and other publications (3). In 1930, Cecil's Textbook of Medicine consisted of 1584 pages. In 1980, the same text required 2357 pages to cover the same discipline.

The two graphs shown below represent the accumulation of scientific information on the left and medical information on the right. The graph on the left is from a study performed by the French economist, Georges Anderla, for the organization of economic cooperation and development (OECD) in 1973 (4). Anderla took as his unit of measurement the known scientific facts of the year 1 AD According to Anderla's estimates it required 1500 years to double this amount of knowledge. The next doubling required 250 years which brought us close to the time of the formation of this country. After that, doubling occurred 150 years later and then 50 years bringing us to 1950 after two world wars. The next doubling took only 10 years, then 7 years, then 6 years to bring us to 1973, the year in which Anderla completed his study.

Scientific Information

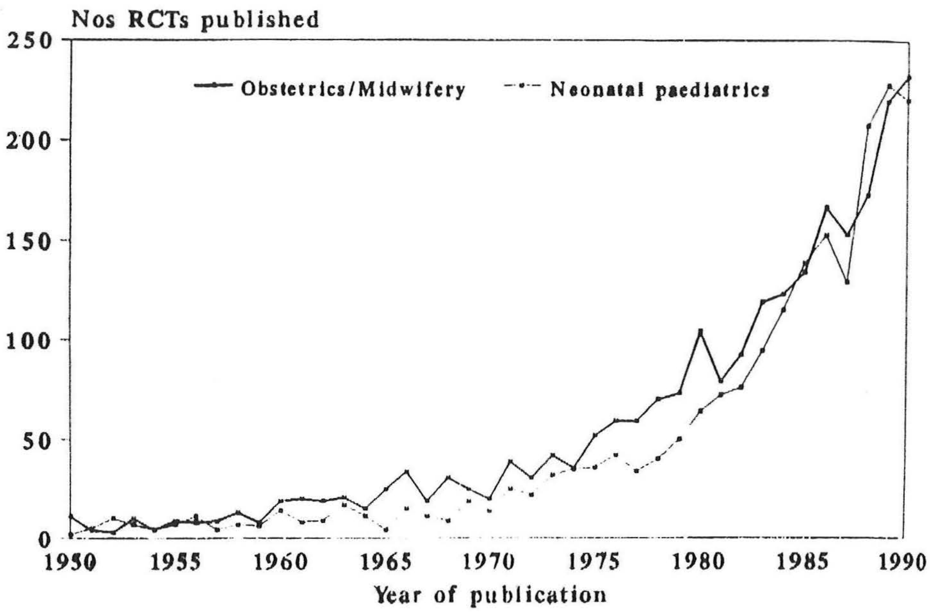



Medical Information




A similar curve is shown on the right from a small paper in the NEJM called "The Weight of Medical Knowledge" in which David T. Durack used the weight of the Index Medicus as a measure of medical knowledge accumulation (5). This study was published in 1978, some 5 years after Anderla's study. Durak extrapolated the acceleration to predict that the cumulated Index Medicus for 1985 would weight in at approximately 1000 kg. However, a follow-up to this study published in 1989 from a letter to the editor by Dr. Madlon-Kay reported that the actual weight of the 1985 cumulated Index Medicus weighed 37.32 kg, that the exponential growth had leveled off and that if the present rate of growth was extrapolated, the Index Medicus would not reach 1000 kg until approximately 2027 (6).

Archie Cochrane published an influential book, Effectiveness and Efficiency: Random Reflections on Health Services in 1972 (7). He suggested that limited resources should be used to provide health care which had been shown in properly designed studies to be effective. He stressed the importance of using evidence from randomized controlled trials (RCTs). A Cochrane collaboration was developed to provide up-to-date reviews of all relevant RCTs of health care. The pregnancy and childbirth review group, for example, comprises approximately 30 reviewers who maintain between 500-600 systematic reviews of RCTs and deal with approximately 200-300 new reports of trials each year (8). The graph below illustrates this daunting task.



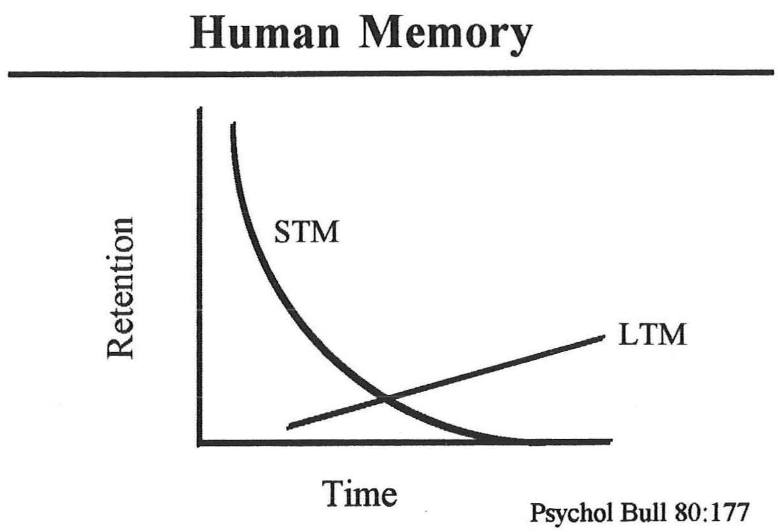

These exponential curves represent an accumulation of knowledge impossible for any single human to accumulate in their lifetime (as shown below). This information explosion points out the impossibility of staying "current" with the medical literature. If you read two articles every day for one year, at the end of that year you would be 800 years behind in your reading! Staying current on drug therapy alone would require a daily reading log of 273 pages (9).

In a recent survey, fewer than 25% of physician leaders regularly did personal on-line searching of the literature and only 5% of community physicians did so (10). Robert Brook of the Rand Corporation (11) has suggested that the purpose of journals is not to disseminate information, but to promote faculty and that this is the sole reason and justification for the journal's existence. The Rand Corporation undertook a detailed, random survey of doctors in this country concerning reading habits as part of an evaluation of the NIH consensus development process. Their conclusions were that doctors read virtually nothing in journals, not even the abstracts. This was also shown in a study of physicians in private practice, 20 to 50% of whom were unaware of advances such as hemoglobin-A1C for diabetic control (10). The RAND surveys revealed that doctors basically want guidelines or short statements available to them at the point in time when they need them. Brook suggests that what the physician needs is a "talking wall" that a doctor can ask a question of while examining a patient. The wall would then respond with an answer. Of course, the question that must be asked is how the information published in journals is going to produce the answer that the wall gives (11). Clinical Information Systems (CIS) are one step toward the realization of that "talking wall."

## Human Memory

According to Thomas K. Landauer (12), who was the Head of Bellcor Laboratories (the old Bell Laboratories), the maximum input to long term memory is approximately 2 bits per second. This input rate was relatively constant regardless of how the information was presented (reading, pictures, listening, music, or nonsense syllables). If someone could sustain this level of input for 12 hours per day, for 100 years, the total amount of information would be approximately 3 billion bits of information. This amount of information could fit on a 5 1/2 inch compact disc (CD). This limit to long term memory is compatible with Dr. Luria's monograph of a mnemonist, an individual with an unbounded memory. Even this individual required approximately 1 second between input of words or digits (13).

Short term memory(STM) has been shown to consist of approximately 7±2 bits of information (14). There is an exponential decay to short term memory as shown in the following graph.



Psychol Bull 80:177

Long term memory(LTM) accumulation is, however, linear (15). There is a saying that individuals use only 10% of their brain capacity or that they take in only a certain percentage of what they hear or see. To confirm the "10%" notion we would have to find someone that actually used 100% of their brain. The trouble is that the best estimates of the difference between the mental capacities of a genius and the dumbest normal individual is about a factor of two (16). While I could find nothing about the "10%" question, it is obvious that no one person could remember it all in his or her lifetime, even assuming total retention. There are many reports of eidetic, or so called photographic, memory. However, there are no reports of anyone memorizing even 100 books or an entire encyclopedia (12). Our strength lies not in what any one individual can remember, but in what we as a society can remember, and more importantly, use.

## Human Capabilities of Data Processing

The working memory for humans consist of approximately 5 "chunks" of knowledge (17). Experts differ in the size of their chunks, not their processor capabilities (18,16). This was demonstrated by comparing the ability of various chess experts to reproduce actual positions of play on a chess board. When individuals were exposed to a board with the pieces in actual play positions consisting of approximately 20-25 pieces for 10 seconds, the grand masters were able to reproduce the board with 100% accuracy. The novice players were only able to reproduce approximately 4-6 pieces on the board. Intermediate players were able to reproduce the board somewhere between the experts and novices. When the board consisting of 20-25 pieces, was presented in random order rather than in actual play positions the chess masters were no better than the novices at reproducing the board. With further analysis, it was found that the chess masters remembered the board in "chunks" of 4-6 pieces. It was also estimated that the masters kept some 50,000 to 100,000 of these "chunks" in their memory (18,16). It is interesting that this is the same range for the number of words in the vocabulary of college educated individuals. This idea of chunks of knowledge is easily demonstrated. For instance, if I give the following sequence of letters:

**T, E, A, I, T, E, A, H, C, T, N, H, H, T**

it is doubtful, with an exposure of only 10 seconds, that you would be able to remember this sequence of letters, however, if I rearrange the letters into the following phrase:

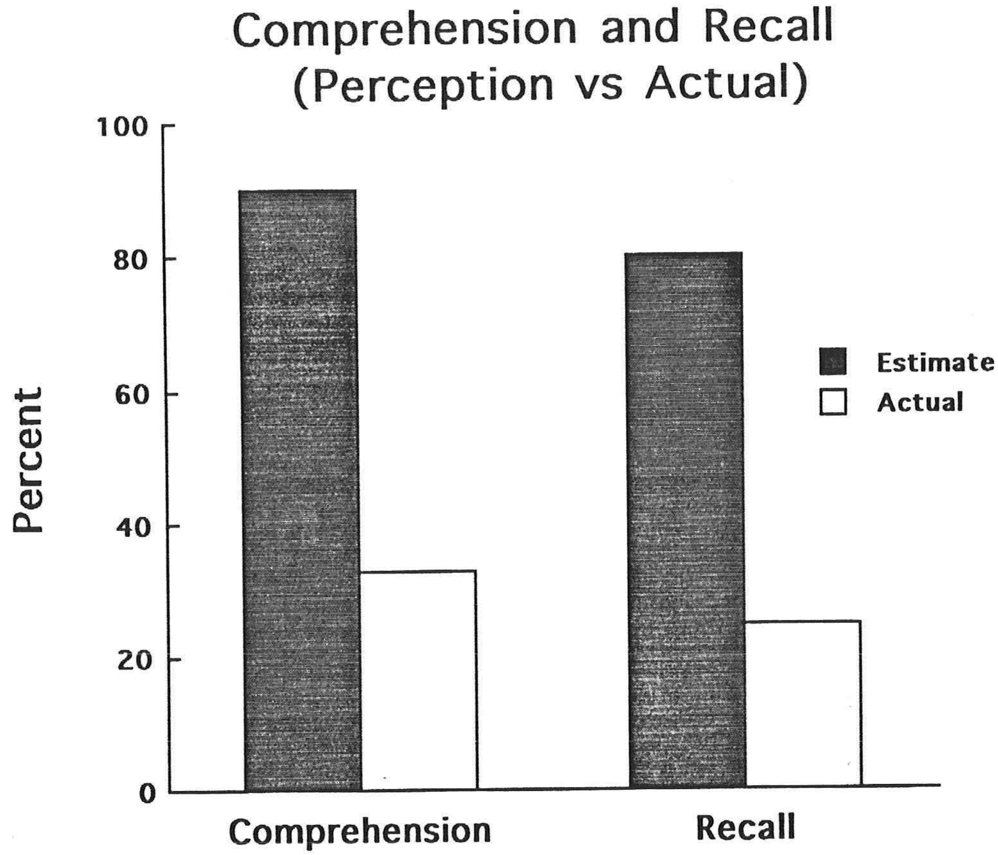
**THE CAT IN THE HAT.**

you would probably have no difficulty in remembering it short term or long term. The reason is that the letters, when given in a random sequence, consist of 14 bits of information. However, when rearranged into the sentence they consist of 5 "chunks" of information or, if you are familiar with Dr. Seuss' book, 1 chunk of information, the title of his book. Therefore, you only have to remember 5 bits or chunks of information or, in fact, only 1 chunk of information, the title.

When physicians are confronted with a clinical problem, they divide the problem into pieces that they can handle, using only a few attributes (19). They process one piece of the problem, then proceed to the next piece until the problem is solved. There is no evidence that any mathematical operations are performed, and any probabilities used are not precise numbers, but rather represent landmark values used in a "representativeness heuristic" or an "anchoring heuristic." This method is necessary because the processing ability of humans is relatively slow with a limited working memory and subject to frequent interruptions, but at the same time allows sophisticated reasoning with these limited resources. The piecemeal method minimizes lost work from interruptions, however introduces biases which we are unaware we possess (19,20) (see "Decision Support" below).

## Comprehension and Judgment

In a study conducted at the University of Laval in Quebec, (21,22) medical students and physicians with various levels of clinical experience were given a portion of text from Harrison's *Principles of Internal Medicine*. They were asked to read the passage and then give two percentages; one would represent the percentage of their comprehension of the text, the second percentage would represent their recall at some later time. They were then formally tested for comprehension and recall. All the subjects were overconfident regardless of their level of training (see graph below).




J Am Cell Cardiol 9:1385, 1987

Measured comprehension averaged 33% which was significantly lower than the subject's estimated comprehension of 90%. Their measured recall averaged 25%, again much lower than their estimated recall of 80%. This represents a bias that our memory is much better than in actuality. This is much the same as our lack of awareness of our own visual blind spot, or the inability of someone with cortical blindness to admit that they are blind.

These same Quebec investigators evaluated the origin of diagnostic errors (23). They presented 20 case histories of patient's who had been previously misdiagnosed to a new group consisting of 59 physicians at various levels of training. The subject's were asked to make a diagnosis on the basis of the available clinical information. They were then given a true-false test to determine if they possessed the factual knowledge that was needed to correctly diagnose the cases. Almost half the cases were misdiagnosed by the new physicians. However, most of the errors (57%) were made by individuals who possessed adequate factual knowledge as evidenced by the true-false test. The errors were attributable more to faulty reasoning than to inadequate knowledge. Certainly we have trouble with both data input as well as interpretation of that data that we do get into our processor (19).

Physicians tend to hold a bias or preconception toward an assumption of illness. An example of this bias can be seen for physician referrals for tonsillectomy from a 1934 survey performed by the American Child Health Association (24). Of 1000 children, 611 had already had their tonsils removed. When the remaining children (389) were examined by a new group of physicians, 174 (45%) were thought to be in need of tonsillectomy. This left 215 children with apparently normal tonsils, however, when they were again examined by another group of physicians, 99 (46%) were thought to be in need of tonsillectomy. This left 116 children with apparently normal tonsils, however these children were again examined by a third group of physicians who thought that 51 (44%) of these children were in need of tonsillectomy. These physicians were inclined to remove a certain proportion of tonsils regardless of the signs and symptoms observed.

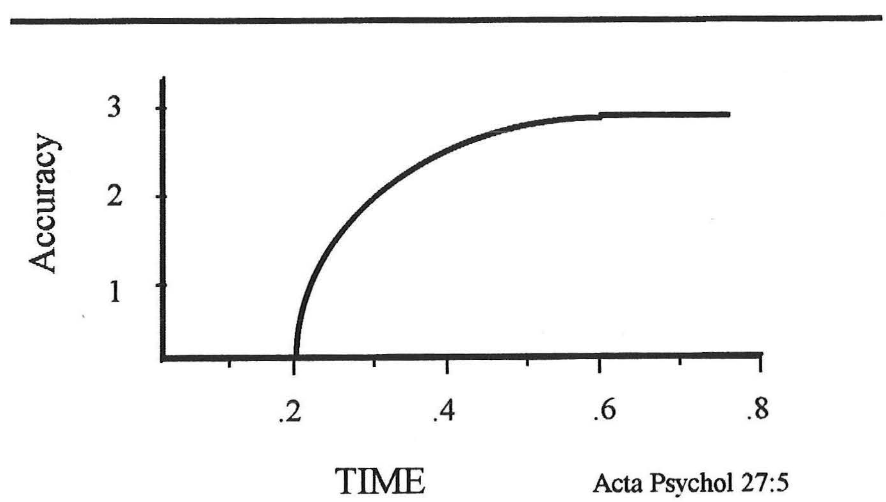
Hingston, et al, (25) asked physicians attending medicine grand rounds if they felt they knew how to interpret blood gas data. Sixty-one percent answered in the affirmative and 71% felt that an expert system to interpret blood gas data was unnecessary. However, these same physicians were only able to answer 39% of posed questions correctly.

## Error Rate

In studies with pilots, it is obvious that if you give the pilot too little information then the pilot will crash. However, something that is less obvious is that too much information also results in a crash in simulated flying (26). Studies have also been carried out with generals in battle simulations (27), computer programmers (16), and clinicians, yielding similar results. For critical care specialists, approximately 4 variable determinants overwhelm the human processor and result in an inability to handle inverse ratio mechanical ventilation in a patient with 4 variable determinants (28,29).

It is also obvious from the graph below that there is a trade-off between speed and accuracy for human processing capabilities. As more and more time is given to accomplish a given processing task, accuracy increases to an asymptotic maximal level. Unfortunately, even this maximal level has an inherent error rate (30).

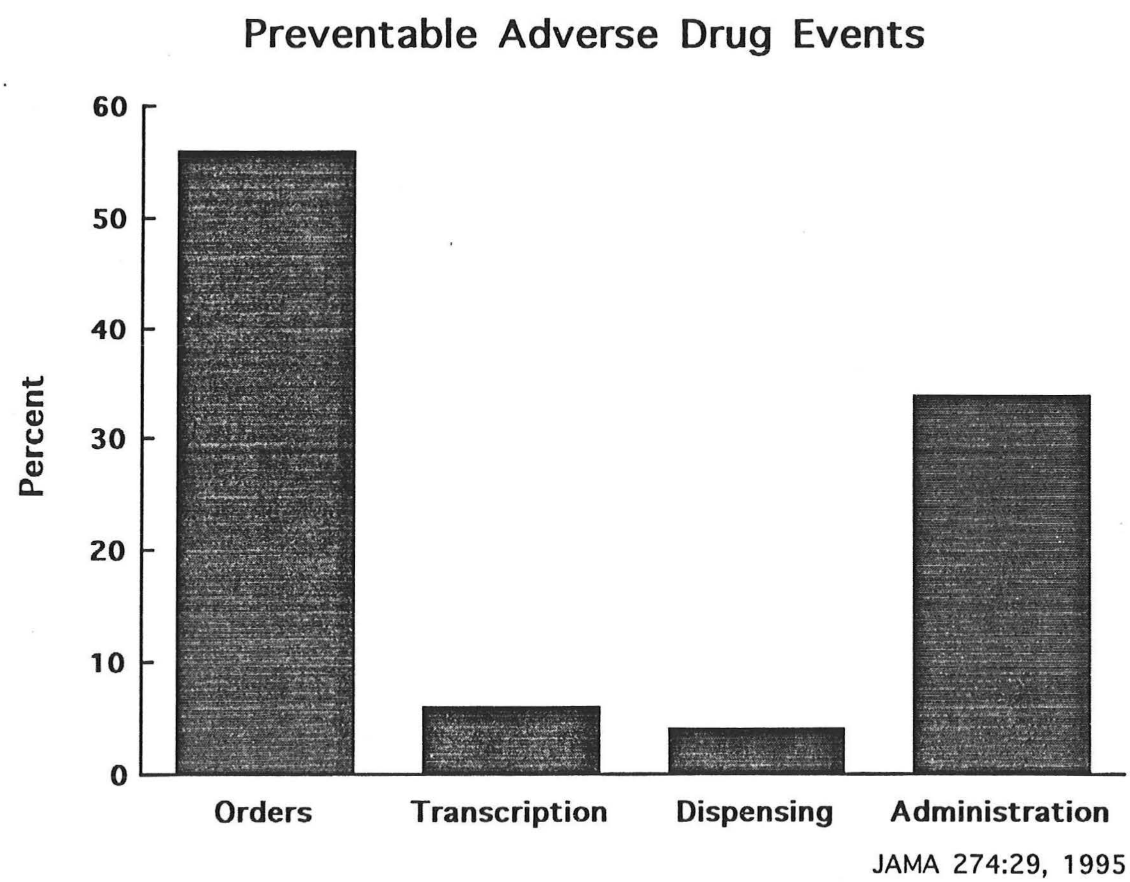
**Speed vs Accuracy for Humans**




The poor function of humans as data processors should be reflected in an error rate in decisions that are based on data. While it is impossible to estimate the true frequency of medical mistakes, they are undoubtedly much more common than most of us would like to admit (31,32). There are two published studies that could help us at least guess at our error rate. One study, conducted in California in 1974, concluded that 3 million hospital admissions led to 140,000 injuries, giving an incidence of 4.7% (33). The Harvard medical practice study conducted in New York state in 1984 concluded that the 2.7 million hospital admissions that year led to 98,600 adverse events, giving an incidence of 3.7% (34). If this were extrapolated to the rest of the country we could estimate that over 1 million patients are injured each year due to their inpatient treatment. Approximately 180,000 of these patients would be expected to die each year as a result of these injuries. To put this in perspective, approximately 45,000 people die each year in automobile accidents in this country. In 1990, nearly 92,000 people died in any kind of accidental death in this country (35).

In 1963 there were 650 FDA approved drugs available for use by physicians. Today there are 9500 FDA approved medications for physicians to prescribe, representing a 1500% increase over these 30 years or so (9). The number of possible combinations and interactions with this volume of drugs is astronomical. (much less than the $2^{9500}$ combinations and 9500! (9500 factorial) permutations, but still astronomical)

A recent study (36) found an incidence of 6.5 adverse drug events (ADE) and 5.5 potential ADEs per 100 non-obstetrical admissions. Of all the adverse drug events, 1% were fatal, 12% were life-threatening, 30% were serious, and 57% were deemed to be significant. The researchers found that 28% of the adverse drug events were preventable. However, the fraction of ADEs that were preventable varied with the severity of the adverse drug event. Of the life-threatening and serious ADEs the researchers found 42% were preventable. Whereas, only 18% of the ADEs that were deemed to be significant were actually preventable.



The rate of ADEs was highest in medical intensive care units. The rates for the medicine ICUs were 19.4 per 1000 patient days compared to surgical ICUs and general medical or surgical wards all of which had rates of 8.9 to 10.6 per 1000 patient days. As can be seen in the graph above, most of the errors occurred when the drugs were ordered or when the drugs were administered, both of which could be affected by clinical information systems.

It has been estimated that each ADE increases patient costs by $2000 per ADE (37). This estimate did not include the costs of the injuries to the patients nor the malpractice costs that are associated with ADEs. Extrapolation of these numbers to Parkland or the VA would be an annual cost of between $3,000,000 and $4,000,000, of which approximately $1,000,000 would be due to events that were preventable (36).

## The Relativity of being Wrong

In the early days of civilization the general feeling was that the earth was flat. This was not based on superstition, but rather, on relatively sound evidence. Another way to express the curvature of the earth is with the deviation from perfect flatness. A perfectly flat earth would have a curvature of 0 to the mile. We are taught that the flat earth theory is wrong, that it is absolutely wrong. But it isn't. The curvature of the earth is actually very close to 0 per mile, which accounts for why the theory lasted so long.

As early as 350 BC the Greek philosopher, Aristotle, summarized why the flat earth theory was unsatisfactory and suggested that the earth was a sphere. Approximately 1 century after Aristotle the Greek philosopher, Eratosthenes calculated the size of the earth's sphere to be approximately 25,000 miles in circumference. The curvature of such a sphere is about 0.000126 per mile(or 8 inches to the mile), a quantity very close to 0 per mile which accounts for the fact that it took so long to pass from the flat earth to the spherical earth theory.

However, the earth is not a sphere, at least not in the strict mathematical sense. Isaac Newton, in the 17th century, demonstrated that a massive body would form a sphere under the pull of gravitational forces but only if it were not rotating. Actual measurements of the curvature of the earth were carried out in the 18th century and Newton was proved correct. The earth was then labeled an "oblate spheroid" rather than a sphere. The "oblateness" of the earth or its departure from true sphericity is 1/3 of 1 percent. This means that the earth curves under the straight line everywhere approximately 8 inches per mile. However, as an oblate spherical surface the curvature varies from 7.937 inches to the mile to 8.027 inches to the mile. The correction in going from spherical to oblate spheroid is much smaller than going from the flat earth to the spherical. This means that although the notion of the earth as spherical is wrong, strictly speaking it is not as wrong as the notion of the earth as flat.

In 1958 when the satellite, Vanguard 1, was put into orbit around the earth it was able to measure the local gravitational pull of the earth, and therefore its mass and shape, with unprecedented precision. It turned out that the equatorial bulge south of the equator was slightly larger than the bulge north of the equator. This pear-like deviation from oblate spheroid perfect was a matter of yards rather than miles and the adjustment of the curvature was in the millions of inch per mile.

Scientific concepts, then, are not so much wrong as simply incomplete. A good concept is gradually refined and extended with greater and greater subtlety as instruments of measurement improve (the above discussion is paraphrased from reference # 38).

Certainly, medicine follows a similar path. In Durack's study the percentage of single author articles from the NEJM dropped from 98.5% in 1886 to 4% in 1976 (5). This may reflect the complexity of experiments such that no single person could carry out all of the experiment (or it reflects another purpose of journals, see above). This trend of multiple authors culminates in publication of articles by multi-national collaborative groups whose authors have to be listed in an appendix at the end of the article. Another trend that I believe has occurred is the number of subjects in clinical trials that are necessary to show statistical differences between experimental groups. The number of patients can reach to the tens of thousands. Surely, the results of these large trials add subtle, but important, corrections to our practice of medicine, much the same as measurements of the earth added subtle, but important, corrections.

## The Limits of Being Right

We may be reaching the limits of our ability to do the right thing. Much of this limit has to do with knowing what is the right thing to do in every situation.

In 1986, John Simes conducted two separate meta-analyses. One meta-analysis used data only from studies that had been published. The other meta-analysis used data from studies listed in a cancer trials registry, some of which had never been reported. The meta-analysis using published studies found that a combined chemotherapeutic regimen for ovarian cancer was statistically significantly superior to the use of a single alkalating agent. However, when the results of all registered trials were considered, no statistically significant advantage for combination chemotherapy was seen (39).

So many groups publish "Practice Guidelines" that to date there are over 1500 guidelines. Unfortunately, these are not uniform in their development, are not seen by practicing physicians, and are not used by those physicians who do see them. In a physician survey conducted by the American College of Physicians 1/3 of the respondents reported significant concerns such as a potential loss of autonomy from increased guideline use. The survey results also showed that fewer than 15% of internists reported that guidelines greatly affected their clinical practice (40). Guidelines are produced by governing bodies of various medical societies, government affiliated agencies, and health maintenance organizations. Some of these may have overriding priorities, such as cost, that weigh heavier than they should.

In 1993 both the British Thoracic Society and the American Thoracic Society published recommendations for the empirical treatment of community acquired pneumonia. Even though both these groups reviewed similar data in the literature, their recommendations are quite different. A recent "current concepts" article in the NEJM gives a third set of recommendations, again, utilizing the same literature (41-43).

Tobian et al examined the opinions of 7 "experts" in hypertension concerning the impact of "official" guidelines on clinical practice. The individual therapeutic recommendations of these 7 experts reflected the diversity of first-rate treatment plans that aim to reduce the cardiovascular sequelae in individual patients with essential hypertension. However, not one of these 7 treatment strategies followed the "preferred" treatment of the US guidelines developed by the Joint National Committee (JNC) which recommended diuretics and beta-blockers as first line therapy (44).

I could find no studies that demonstrated improved outcome from any "practice guidelines" from any governing group. Despite this, practice guidelines have been used as a measure that generalists do not provide as "good" care as specialists (45). One reason for the lack of literature to support nationally distributed "practice guidelines" could be that they are not specific enough for any one institution or practice. They are forced to include fuzzy terms that do not help the physician enough. A recent study on practice guidelines for antibiotics used locally generated guidelines, which use more local clinician time for development, but produce practice-specific recommendations that are agreeable to all the physicians and give demonstrably improved outcomes (46) (see below under "Critiques").

It is difficult to know that our good intentions are not doing harm. One result of the Cochrane initiative is demonstrated in Effective Care in Pregnancy and Childbirth, where only 35% of the 283 interventions examined were demonstrably beneficial. Of the remaining 248 interventions 22% were found to do more harm than good (8). A recent study found that application of intensive primary care actually resulted in more frequent readmissions, rather than the more intuitive opposite (47).

The Rand corporation has applied ratings to many procedures in the United States and abroad and have demonstrated that as many as 1/3 of procedures may be performed for less than appropriate reasons. This rate varies by country, hospital, and procedure (48). It has been suggested that health care interventions be evaluated differently (49).

## The Need for Clinical Information Systems

The quantity of data available on any sick patient in the intensive care unit is truly phenomenal. The output from the patient monitor for heart rate alone can reach 1.5 megabytes of data per day. The total output from the data port of a Puritan Bennett 7200 ventilator reaches over 3 megabytes of information per 24 hours. Allan Morris and his colleagues (50) counted 236 classes of information gathered on a sick patient in their ICU. An example of one class of information would be blood pressure where you might have many readings of blood pressure throughout the day, however, that's only one class of information out of the 236 (50).

The institute of medicine has suggested that an electronic medical record (EMR) will help lower health care costs, maintain quality of care, and provide physicians with better information (51). Tierney et al, (52) undertook to test whether this suggestion was valid. They used micro-computer work stations linked to a comprehensive EMR system for writing all inpatient orders on an inpatient medicine ward. A control group wrote orders in the chart in a conventional manner. The intervention teams generated charges that were significantly lower (12.7%) per admission than did the control teams. Significant reductions were also shown for bed charges, diagnostic test charges, drug charges, and total hospital costs. The mean length of stay for the patients of the intervention group was nearly a day shorter, however this did not reach statistical significance. They estimated a hospital wide savings of $3,000,000 per year, $300,000 on drug supplies alone.

The poor function of the human processor and the huge amounts of information gathered in sick patients clearly demonstrate that physicians need help in processing all this information. Clinical information systems (CIS) can be designed to do just that. And unlike human processors, computers perform calculations just as well at 3:00 a.m. as at 9:00 a.m. The computer does exactly what it is told to do (programmed), does it with great speed, with virtually no mistakes, and with constant vigilance.

There are several limitations to the traditional medical record:

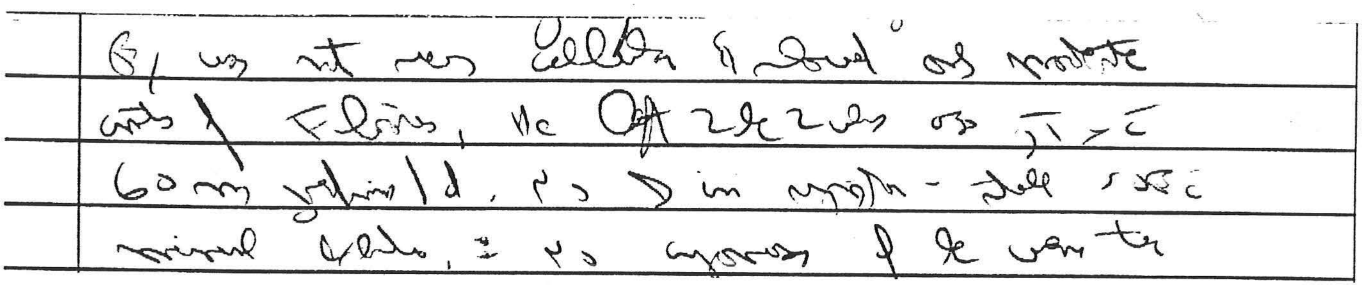
- It is physically inaccessible.

I am sure you have seen patients in clinic without an old chart. In the emergency room, it is usual that the patient's record cannot be recovered and delivered in time to be of benefit.

- Information is available at only one location.

  If the patient has two clinic visits, there is a good chance that the chart will not make it to the second clinic.
- The chart is usually poorly organized.

  Anyone trying to locate a subspecialty consult in a chart has run into this poor organization.
- Illegible handwriting may make information nearly impossible to retrieve.



excerpt from H & P by Parkland Attending

- A lack of standardization for record-keeping makes it difficult to compare charts.
- Manual chart review must be done by trained personnel, making research costly and cumbersome.
- Data recorded from electronic instruments must be hand-written or manually attached to the patient's chart, opening a source of error.

In one study of physician's information needs, 454 clinical questions were recorded during 17 hours of observed clinical activity consisting of approximately 90 patients with 24 physicians and medical students. Over fifty percent of these questions could be found in the patient's chart (53). The following table indicates how computers can be used to assist physicians in solving the problems with record-keeping, data management, and decision making in ICU's. Below, are specific aspects that have been shown to be of benefit.

**Uses of the Computer in Intensive Care.**

1. Assist in data collection.
2. Provide computational capability.
3. Assist in data communication and integration of data.
4. Record-keeping.
5. Report generation.
   a. Variable report format
   b. Available at multiple sites
   c. Data communications
   d. Eliminate redundancy
   e. More structured reports
   f. More accurate reports
   g. Current information
6. Assist in decision making.

## Tabulation of Data - The Flow Sheet

In the 1960's Lawrence L. Weed advocated an organized approach to the medical record. He suggested problem oriented progress notes, time dependent flow sheets, acceptance and use of paramedical personnel, and a more positive attitude about the computer in medicine (54).

Physicians gather data in order to make diagnostic and therapeutic decisions. Computers can accumulate data on patients rapidly and accurately and transmit that data quickly to the physician. For instance, data from the clinical laboratory can be viewed in the intensive care unit virtually as soon as the test is performed (see "What Data do we Need" below).

Tierney and co-workers showed a 13% decrease in lab tests ordered when physicians were given the results of previous test results at the time of order entry (55). A subsequent study showed an almost 9% decrease in lab tests ordered when the computer displayed the probability (0-100%) that the test would be positive for the main abnormality being tested for (56) (see What Data do we Need, below).

## Reminders

*Man more often needs to be reminded than instructed.*
Boswell

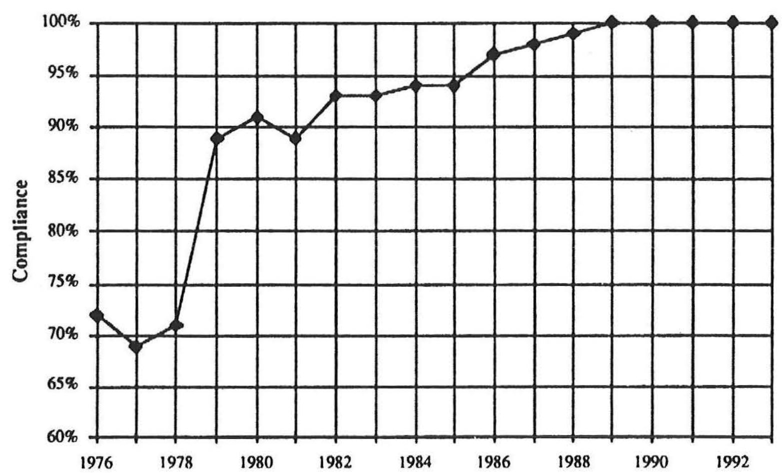
We would all like to think that the process of medical care is a sophisticated, intellectual activity, forming hypotheses based on extensive knowledge of underlying pathophysiologic mechanisms. However, McDonald has pointed out that much of the day to day work in medicine consists of simple recognition-response arcs, such as a fall in hematocrit triggering a test of the stool for occult blood, or a newly positive PPD triggering a chest x-ray. These "reflexes" consist of a stimulus (a simple clinical event) and a response consisting of a corrective or clarifying action. The logic between the stimulus and the response is defined by simple rules or protocols. Much of medicine, both routine and important, requires a compulsive attention to the rules or protocols rather than intellectual brilliance (57).

As we have seen from the underlying error rate in medicine, physicians' "clinical reflexes" fail. McDonald (57) has shown that this failure rate can be improved by the use of a protocol driven computer reminder system. He demonstrated that clinicians responded to 36% of alerts when the computer reminders were given and 11% without the computer reminders. In a subsequent controlled, cross-over design study (58) physicians responded to 51% of events when given computer reminders but only 22% when the computer suggestions were withheld. The level of postgraduate training was unimportant. Another interesting finding was that the order in which the physician's served as study and control subjects had no significant effect on the results. McDonald's conclusion was that the prospective reminders do reduce errors and that many of these errors were due to man's limitations as a data processor rather than to correctable human deficiencies (58). The physicians given the reminders first would have done better if their problem was learning. Instead, they went back to the same "error" rate as the physicians in the "control first" group. We often try to push students further into a system, thinking that if we can just teach them more, they will perform better. McDonald's study, and others (59,60) suggests that this is not the case. A change in the system may be needed in order to improve care.

## Critiques

Pestotnik et al (46) recently published the results of a seven year study in which they implemented antibiotic practice guidelines through computer assisted decision support. This study involves both critiques and reminders, but illustrates the basic premise. They found that nearly 40% of their patients over the 7 year period received antibiotics. The proportion of hospitalized patients who received antibiotics increased each year from 32% in 1988 to 53% in 1994. Likewise, the use of broad spectrum antibiotics increased from 24% of all antibiotic use in 1988 to 47% in 1994. Despite these increases and an additional increase in annual Medicare case-mix index, their total costs of antibiotics, adjusted for inflation, decreased from 25% of pharmacy budget in 1988 to 13% in 1994. They were able to decrease their costs per treated patient by more than half, from $122 per patient in 1988 to $52 per patient in 1994. The percentage of patients receiving appropriately timed preoperative antibiotics increased from 40% in 1988 to 99.1% in 1994. Antibiotic associated adverse drug events decreased by 30% over this time period. Their antimicrobial resistance patterns and length of stay remained stable during the time period. They were able to show a statistically significant decrease in mortality rates, from 3.65% in 1988 to 2.65% in 1994. The authors concluded that much of this improvement in outcomes was secondary to their computer assisted decision support program and the use of local clinician derived practice guidelines.

The practice guidelines utilized in this study were developed locally and were subject to a continuous review and update over time. The figure below demonstrates the increasing compliance with the system's suggestions over time.

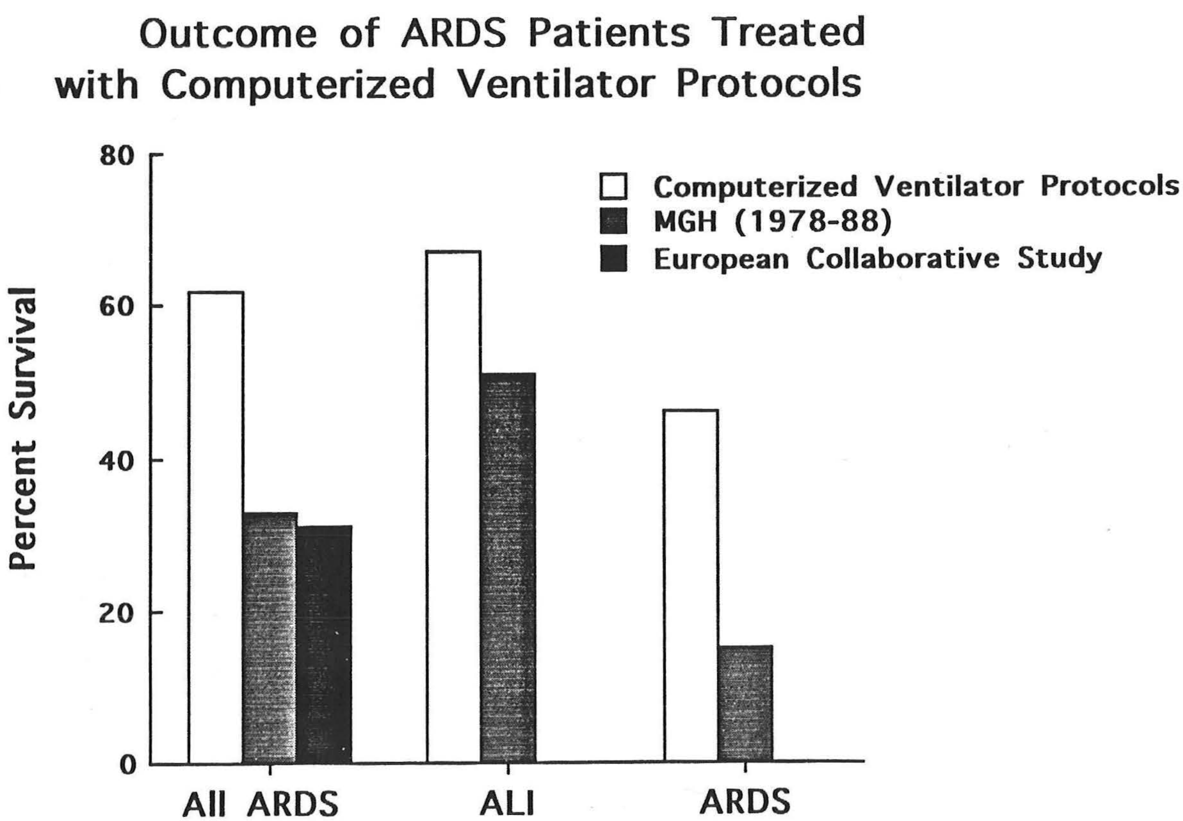



## Suggestions or Instructions

A good example of the use of an expert system to generate suggestions or instructions comes from the LDS Hospital in Salt Lake City, Utah (50,61,62). A European study published in 1986 showed an unusually high survival of patients with severe ARDS with the use of pressure controlled inverse ratio ventilation accompanied by extracoporeal $CO_2$ removal (ECCO2R) (61). The physicians, headed by Alan Morris, wanted to repeat this study, but were aware that the management of acute respiratory distress syndrome (ARDS) differed from patient to patient depending on the course the disease followed and the previous experience of the physicians and staff caring for the patient. They sought to standardize therapy to increase the interpretability and credibility of the clinical trial results.

A group of physicians, nurses, respiratory therapists, and medical informaticists were gathered to create protocol logic that would guide the ventilation of patients in both the control and experimental group. The initial work of this group took 18 months with the major problem being physician agreement on a standard protocol. Physicians had to agree to give approaches to therapy that had to do with individual style and agree on a detailed, standard approach to patient care. The protocols were initially developed on paper, however, as they became more complex it became difficult to follow the protocols manually. Therefore, the protocols were computerized.

The resulting system was successfully used during the ECCO2R study (61). The study was terminated after 40 patients were treated with 42% survival in the conventional therapy group and 33% survival in the ECC02R group, demonstrating no significant difference between the groups. The 42% survival in the control group was unexpected since reported survival in these patients was between 0-15%. The researchers suspected that the quality and uniformity of care provided through the use of the computerized protocols were the reason for the improved patient outcomes. Comparison of the computerized ventilator protocols with historical controls is shown below. (ALI = acute lung injury, or mild ARDS)



The success of these mechanical ventilation protocols clearly indicates the feasibility of using expert systems to direct management of care for critically ill patients. Specific instructions were generated and 95% of these instructions were followed by the clinical staff. The largest fraction of instructions that were not followed by the clinical staff were due to inaccuracies in the data, particularly respiratory care and blood gas data. Less than 1% of the instructions were declined secondary to physician disagreement with the protocol logic.

A randomized, controlled, prospective, clinical trial has been developed to compare the LDS computerized protocol control of mechanical ventilation with conventional physician-directed care. One part of this study is whether the protocols can be exported to other hospitals and be accepted by physicians who are not involved in the development of the protocols. Preliminary data shows that patients at other institutions were kept on the protocol nearly 93% of the time with around 96% of total instructions followed (63).

## Retrospective Quality Evaluations

One of the fastest growing segments of any modern hospital has been quality assurance. This function goes under many names, such as quality assurance, quality improvement, continuous quality assurance, continuous quality improvement, quality management, medicine performance improvement, and performance improvement committees. Documentation of performance, outcome measurements, and other aspects of quality assurance are requirements of many accreditation bodies including the Joint Commission for Accreditation of Health Care organizations (JCAHO).

The Parkland Respiratory Therapy department recently acquired an information system. They record approximately 6,000 data items each 8 hour shift, occupying 6-7 megabytes of storage every 24 hours. This represents an enormous amount of data that must be reviewed for quality assurance. They routinely track the productivity of each employee and the workload per therapist, but also monitor the treatment frequency within each ward or unit of the hospital. This allows a more or less real-time performance improvement by shifting workload to a therapist who may not be as busy on another ward. This allows intra-shift reallocation of resources to improve efficiency.

Typical quality assurance surveys are done manually and retrospectively. They are usually done on only a fraction of the patients, are time consuming, and often do not produce results until months after the project was initiated. The usual performance of quality assurance is to define a benchmark of quality and then perform a manual chart review to determine the pre-existing compliance with this benchmark. Education or, less frequently, changes in procedure are used in an attempt to improve the quality. A second manual, random chart review is performed to gauge the success of the intervention. In contrast, a computerized data base allows a real-time improvement of care by comparing practice with detailed "standards" previously identified as associated with quality care. When a transgression is made against that "standard" a report can be generated, an alert given, and the problem corrected immediately(again, point-of-care). This type of performance improvement can be done on every patient for every clinical situation. This real-time quality improvement is far preferable to the manual, retrospective, delayed quality assurance usually performed.

The Parkland medicine intensive care unit has utilized a computer data base of severity of illness on all patients admitted over the last 6 years utilizing the APACHE II (acute physiology and chronic health evaluation) score (64). This has allowed us to compare our performance in the MICU as reflected in patient mortality with other medical centers (65).

## Education

Feedback about performance and comparison of clinical performance with that of one's peers are significant motivations for physician learning and have been shown to actually change physician behavior (66,67). The feedback is best if individualized to the physician's actual performance and the sooner the feedback occurs, the better (68). A Clinical information system could give an analysis of performance compared to pre-set benchmarks, as well as provide comparison to some aggregate performance of the physician's peers. This could be accomplished in real-time or on some pre-set scheduled time.

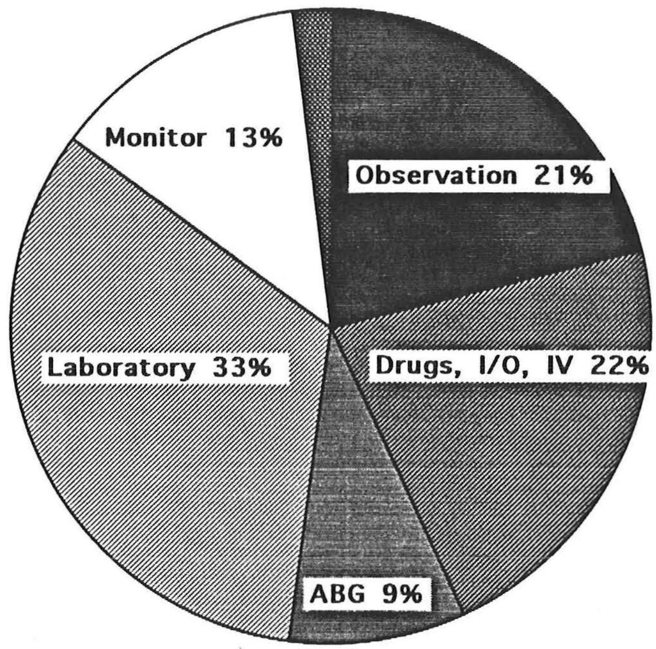
The National Library of Medicine (NLM) initiated a study of their MEDLINE data base because of a shift in patterns of searches. More searches were being done to answer clinical questions rather than academic or research questions as was previously the case. They concluded that MEDLINE was a valuable asset to clinicians as well as the academic and research associated users (69). However, even when a librarian was used to perform the search it took an average of 47 minutes per question and the physician was left only with a pointer of where to get the information (70). And the use of a librarian increased the search cost to around $45 per question (70). In one study of physician's information needs 454 clinical questions were recorded during 17 hours of observed clinical activity consisting of approximately 90 patients with 24 physicians and medical students. Twenty five percent of these questions required a library, textbook, journal, or MEDLINE search (53). Leape and colleagues (71) found that drug knowledge dissemination was the most frequent system error, accounting for 29% of all errors associated with adverse drug events. Certainly all the physicians in their study knew where and how to get information on the drugs that they prescribed, however it was probably too much trouble and their "memory bias" allowed them to proceed with insufficient knowledge. It is not hard to imagine how point-of-care dissemination of knowledge would aid in preventing these adverse drug events. Engines that provide just such knowledge have already been developed and are in use.

Actual journal articles, scanned into computer memory and linked to reminders, critiques, or instructions could provide education at the point of care. Both the evidence and the logic behind the reminder could be reviewed. Policies and protocols for various procedures could also be provided in similar fashion to provide help and education for physicians, nurses, and therapists.

## What Data do we Need?

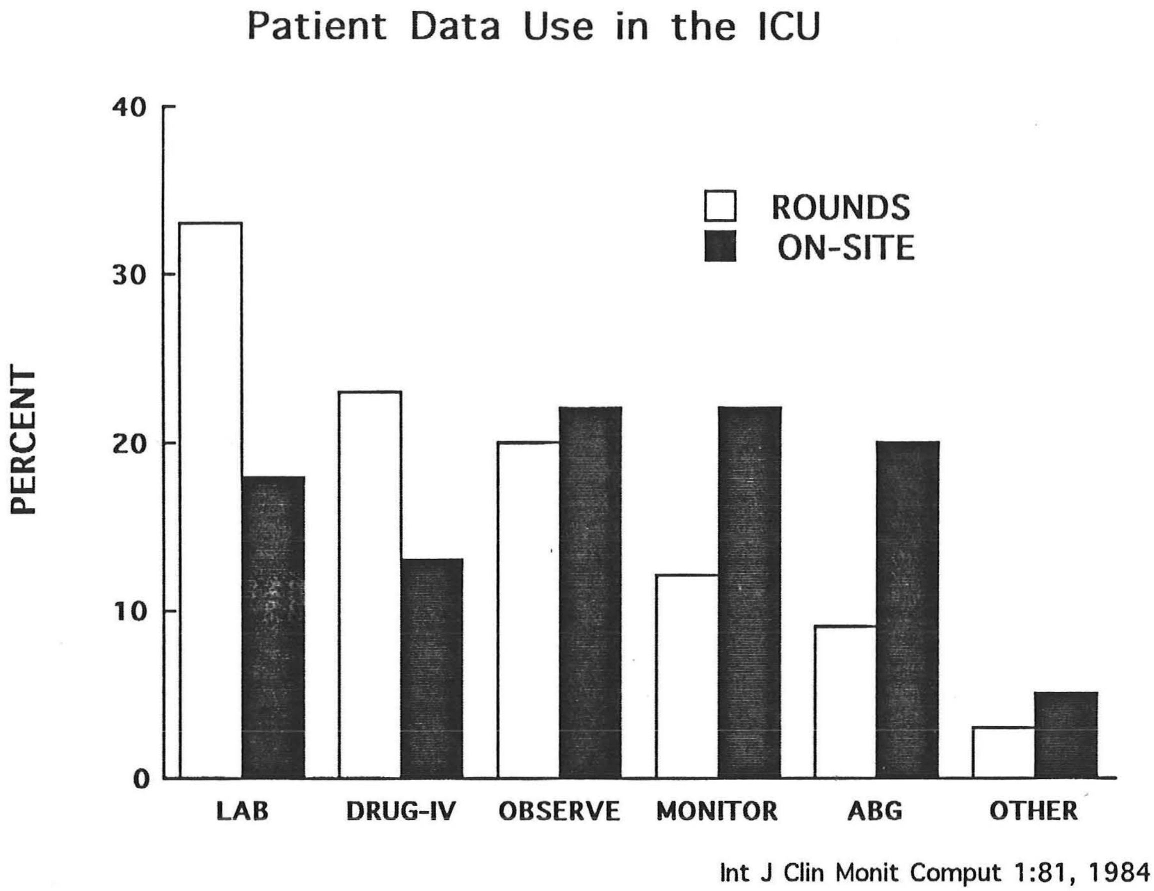
Bradshaw et al (72) examined what data was actually used by the physicians for decision making during teaching rounds in an intensive care unit. The data was divided into six categories shown in the following figure.

## Information Used in Clinical Decisions in the ICU




It is clear from this pie chart that data from multiple sources must be combined for the physician to make effective diagnostic and therapeutic decisions. The data come from multiple sites and the need for this data to be current and located in an easily accessible medium is obvious.

The graph shown below is from a study done in the Shock-Trauma ICU at LDS Hospital in Salt Lake City, UT (73). This is a combined medical surgical intensive care unit. One portion of the study was observation of morning physician teaching rounds where pertinent data on each patient being cared for in the unit is reviewed and plans are formulated for the patient's care during that day. The other portion of the study evaluated the use of patient data outside of rounds, termed "on-site". Physician usage of data, shown in the graph, was compared to data storage in the computer to pinpoint the areas that could be improved for each category. While laboratory data usage was over 30% on teaching rounds and nearly 20% on-site the total laboratory data occupied only 16% of the computerized data base. Because of the high usage to storage ratio for laboratory data it was felt that optimization of speed and ease of use for that data was the most indicated correction to make.




However, data obtained from bedside monitors during rounds made up only about 13% of total data used in the decision making while occupying almost 33% of the computer record. The use of this monitor data increased to 22% for on-site decisions, however this still left a gap of over 10 percentage points between the amount of data stored and the amount of data actually used. This would indicate an area in which data storage could be re-evaluated and optimized so that medical and legal requirements for the patient data base can be efficiently met. Analysis such as this could help match data stored to data used and decrease the noise level of data in the intensive care unit.

## Data Accuracy

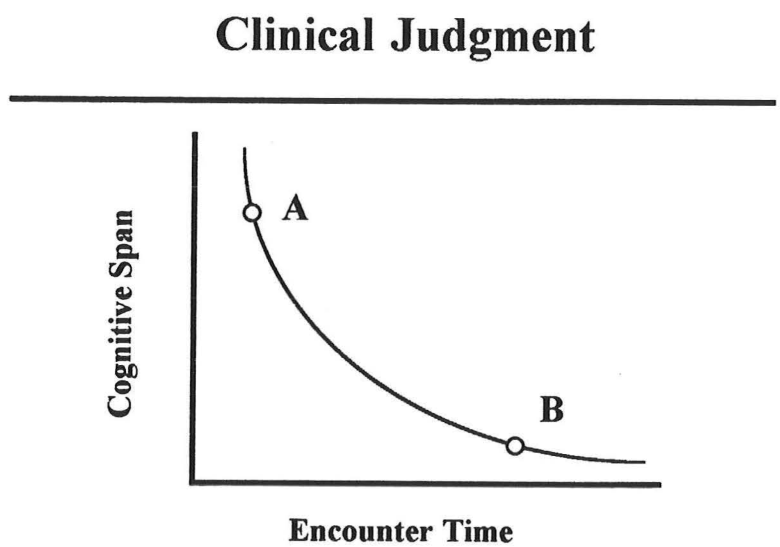
The best control over the quality of ICU data is to use it constantly (50). The more the computerized data is used by the health care team, the greater the likelihood that any errors present will be found. Those who know the patient best are most likely to recognize errors when they are using the data to make the decisions for the patient. The health care teams are legally and ethically liable for the decisions that they make based on the data that is collected. They are therefore very demanding of the data and will insist on its accuracy. Through the use of the data, errors can be found and corrected when found.

Just as you sign your initials when striking through a error in the paper chart an audit trail should be present to document changed or corrected data including the time and identification of the person making the change for legal purposes.

There is often redundancy in the data received on a patient in the intensive care unit. For example, heart rate can be input from the patient monitor with an EKG, the pulse rate detected from the arterial blood pressure monitor, and the pulse rate detected by the pulse oximeter. Which ones of these should be recorded in preference to the others or should they be compared or averaged is problematic.

Hand entry of data into the clinical information system presents problems of entry errors, however automatic data entry is not without problems either. A transient rise in heart rate, blood pressure and drop in $PaO_2$ in a patient being suctioned would probably not be recorded by a nurse or respiratory therapist until the patient had settled back down. Nor should those data values be used in clinical decisions for the patient. However, automatic data entry could certainly detect and enter values during this non-representative time. However, nurses perceptions of representative data have frequently been shown to be in error as well (74). Some combination of automated data perhaps with verification or editing by the nurse or therapist would be most appropriate.

## Decision Support




J Am Cell Cardiol 9, 1987

Blois illustrates the process of a clinician encountering a patient by the use of a diagram shown above. The horizontal axis represents the time of the physician-patient encounter, and the vertical axis is the cognitive span or amount of knowledge required to solve the patient's problem. When the clinician first meets the patient he must be prepared for a large number of possibilities. However, when the problem is better defined he is required to focus his attention on only a few alternatives. Point A in the figure would be early in the clinical encounter when a broad, although probably superficial knowledge base is needed, and Point B is a later stage when a small, but detailed, knowledge base is needed. Blois contends that humans are designed to function best at Point A, whereas computers best function at Point B (75).

Diagnosis is often a matter of opinion (76). Computer programs are, as a rule, not very good at diagnosis (or opinion) (75). Dr. Frenkel participated in a recent comparison of four computer-based diagnostic systems (77,78). The results were disappointing. The programs provided the correct diagnosis 52 to 71% of the time, and provided relevant diagnoses from 19 to 37% of the time. These rates were unfortunately not compared to human diagnosticians, however, all the cases were actual cases that had been diagnosed by a physician. One other disappointing aspect of the programs is that each had a specific lexicon needed for input of signs and symptoms. Diagnosis is one thing that we do better than the computer. At least for the time being, diagnosis is probably best performed by humans. However, we would benefit from some support for our "opinions" and this section includes sections on decision analysis and clinical prediction rules which are promising as aids in diagnosis if, for no other reason than that it makes us define the problem (see below).

As a pulmonary fellow, I was impressed by Dr. K. Joy Robertson's ability to tell when a patient was ready to be weaned from the ventilator. However, when questioned as to how she knew a patient was ready for weaning or extubation, she was able to point to a few criteria but then remarked that "you can just tell". I took this to be part of her "art" of medicine. The difficulty in analyzing this "art" is revealed in the WEANPRO system developed by Tong and colleagues for weaning patients from mechanical ventilation. The knowledge for the system was obtained from four physicians who were experts at weaning patients from mechanical ventilation. This knowledge base consists of 406 rules and 133 "metafacts". A review of the suggestions made by the WEANPRO system compared with expert physicians showed a 96% acceptance rate. This program is presently in use clinically at the Baptist Memorial Hospital in Memphis (79,80).

Formal analysis of decisions will probably expose many of the inconsistencies and lack of focus that make the decision difficult in the first place. This formal analysis, however, is beneficial because it exposes important flaws in either the decision maker's understanding of his decision or his logic (81,19,82,83).

Decision making has three fundamental components:

Alternatives - what you can do.

Preferences - what you want.

Information - what you know.

Decisions are often difficult to make because there is no dominating alternative. The lack of a dominating alternative requires that doctors know more about the problem. Perhaps about the patient's preferences before they can identify the best alternative. Preferences directly represent the desires of the decision maker, whether it is the patient, or the physician who acts on behalf of the patient. In medicine, preferences almost always concern the patient's length of life, quality of life, comfort and financial or other economic considerations. Achieving clarity of action requires explicitly quantifying the trade-offs between these fundamental quantities. Knowledge or information about possible consequences of action is essential to decision making. Decision analysis treats uncertainty explicitly and can yield optimal recommendations that would be discarded if uncertainty were ignored. Uncertainty is treated using probability functions, and probability is dependent on the amount of information available.

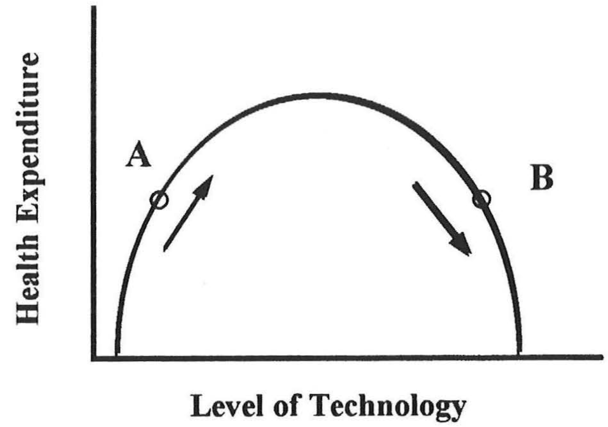
Decision analysis uses probability to design a course of action in the face of uncertain information. However, probability is dependent on the available information rather than describing a natural phenomenon. For example, a 35 year old man comes to the Emergency room complaining of mild chest pain. As the physician seeing the patient, you order an EKG and lab work. Your initial estimate of the probability of this patient having a myocardial infarction is presumably quite low, a young patient with mild chest pain. However, if you were to obtain the family history that both his father and grandfather died at his age of heart disease then your assessment of the probability of myocardial infarction may increase. Then, when the EKG returns and there is only non-specific ST-T wave abnormalities you may be less suspicious and decrease your probability of MI. Finally, if the laboratory work returns with an elevated creatinine phosphokinase (CPK) with a significantly elevated myocardial band (MB), your probability would increase dramatically and acute myocardial infarction would become your diagnosis. Throughout this diagnostic maneuver, the patient himself has not changed only your information has changed. Therefore, there is no correct probability, it depends on what you know.

Many important medical decisions involve uncertainty so there must be a distinction between the quality of decisions and the quality of outcomes. For example, a 20 year old man presents with nausea, epigastric pain, point tenderness at McBurney's point, fever and leukocytosis. The patient is diagnosed with appendicitis and undergoes an appendectomy. Most surgeons would consider this a good decision. However, suppose the appendix is found to be normal during the operation which is then confirmed by pathology. According to decision analysis, the appendectomy would be considered a good decision because it was consistent with the decision maker's alternatives, preferences, and knowledge at the time the decision was made. Thus, the quality of decisions and the quality of outcomes should be measured separately.

Unfortunately, most performance is measured in terms of outcomes rather than decisions. Rewarding outcomes rather than decisions results, ultimately, in bad decisions. Defensive medicine, which has been estimated to account for $160,000,000,000 spent each year, or 22% of the total health care expenditure reflects this effect (84). For example, a young man admitted for an outpatient operation has a massive myocardial infarction and dies. The surgeon is sued for not obtaining a preoperative EKG even though this was not indicated. Therefore, the surgeon then obtains a preoperative EKG for all his patients regardless of indications. The surgeon has thus made bad decision making a routine part of his practice (81).

Lewis Thomas (85) described medical technology on three qualitatively different levels. He called the first level "non-technology", an example of which would be hospice care directed at an untreatable condition for which medicine can only offer supportive care. He termed the second level "half-way technology" and described it as directed toward the treatable consequences of an incurable disease (as with renal dialysis). The third level which he calls "high technology" is directed exclusively at prevention and cure (immunization and antibiotics). Thomas saw these levels of technology as quanta, however Weisbrod (86) described these as lying along a continuum. His model reveals that half-way technology is more expensive than either non-technology or high technology. An advance in technology can either increase or decrease health expenditures depending on the current position along the Weisbrod curve, shown below (85,86).

# Medical Technology and Health Expenditure




J Am Cell Cardiol 9, 1987

According to this model a computer decision aid directed at a problem located at Point A will not be cost effective unless it represents a major leap. In contrast, a computer decision aid directed at Point B will be cost effective even if it accomplishes only a small advance.

Clinical prediction rules can reduce the uncertainty in medical decisions by defining how to use clinical findings to make predictions. Wasson et al (87) set forth the following methodologic standard:

1) explicit definition of outcome,
2) explicit definition of findings used to predict the outcome,
3) patient age and sex stated,
4) study site described,
5) mathematical modeling technique described,
6) test of misclassification rate,
7) effects of clinical use prospectively tested.

Clinical prediction rules will probably not become an integral part of patient care until computer based medical information systems provide them as point-of-decision assistance. Physicians presently do not use quantitative estimates of probability extensively and may not know how to integrate them into test ordering decisions. Patients often desire or even demand certain tests and may gain some psychological benefit from a negative result of that test. Even the way the information is presented can alter the diagnostic approach of physicians (87,88).

## Some Down Sides

ICU clinical information systems (ICU-CIS) have not gained widespread acceptance as a method of recording information despite the huge amount of clinical data collected in the ICU. The primary reason for this has been cost justification. The idea that ICU-CISs are cost justified based on reduced nursing or therapist (FTEE) requirements is controversial and probably untrue. Hammond et al (89) showed that an ICU-CIS significantly reduced the number of errors found in paper flow sheet charts. An ICU-CIS could also be expected to improve the quality, accuracy, timely capture, and easy retrieval of data (90). There is some evidence that less nursing time is spent on direct patient care (from 49%-43%) and that more time is spent on clinical data entry (from 18%-24%) with the institution of an ICU-CIS (91).

There are obvious costs to the hardware and software. In the study by Tierney, et al, (52) the costs were approximately $20,000 per ward. However, there were also costs in terms of time spent writing orders for patients. This amounted to 33 more minutes per day writing orders for the intervention group versus the control group. This averaged out to 5 1/2 minutes per patient. However, it increased the order writing time even more for admissions and discharges, both of which took approximately 10 minutes longer in the intervention group than in the control group. These increased times may seem insignificant, but they hold a huge impact on acceptance of the system and usability.

It will be difficult to show that improvements in the quality of patient charting accomplished through an ICU-CIS will result in improvement of patient outcome. A reduction in mortality from 16 to 14.4% would require a study of at least 6000 patients (92). Indirect measures of quality of patient care such as length of stay, incidence of mistakes, and adherence to guidelines may show significance in smaller studies. Nevertheless, one review of the effects of computer based clinical decision support systems on clinician performance and patient outcomes showed that 3 of 4 studies of computer assisted dosing, 1 of 5 studies of computer aided diagnosis, 4 of 6 studies of preventive care reminders, and 7 of 9 studies of computer aided quality assurance showed improvements in clinician performance under the computer based clinical decision support system. Only 3 of 10 studies that assessed patient outcomes reported significant improvements (93). I suspect that some of the lack of improvement relates to a lack of user-centered design (see below).

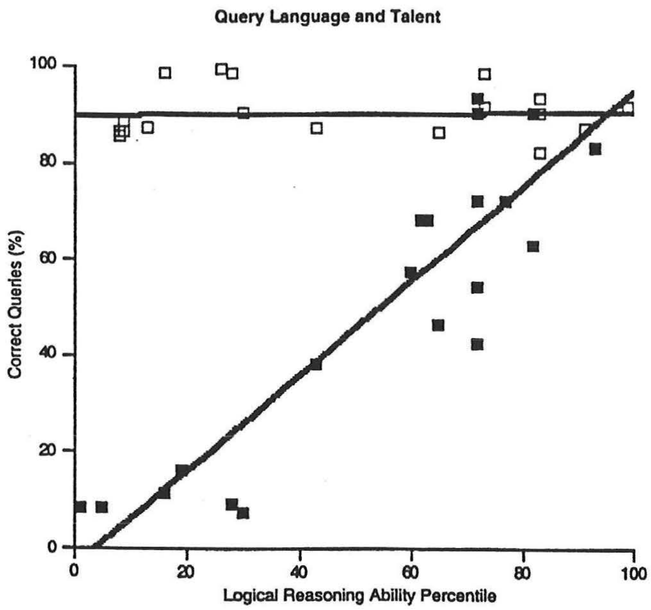
People read anywhere from 5% to 30% slower from computer screens than from printed pages. In order to have the same resolution as offset printing, a computer screen would have to have at least 100,000 dots (or pixels) per square inch of surface. Computer screens usually have less than 20,000 pixels per square inch. In order for a clinical information system to be effective, the down time of the computer has to be minimized. The HELP system at the LDS Hospital in Salt Lake City is available 99.6% of the time (50). This averages out to a down time of 5.77 minutes per day, about half of which is planned down time. One problem of various computer systems is their incompatibility with each other and with medical devices. There are a number of standards which help with this problem including the Help Level 7 (HL-7) and the DICOM 3 standards (94,95).

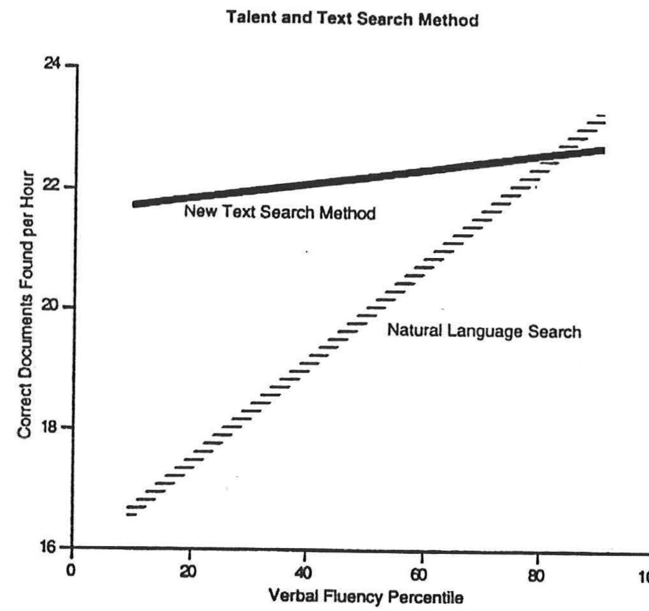
## One Solution? - User Centered Design

Abilities that have been found relevant for a wide range of computer based tasks are age, logical reasoning, memory for the location of things in space, high scores on a logical reasoning test, and verbal fluency. The reason that these characteristics determine computer based proficiency probably has to do with the fact that most computer programmers score very high in these characteristics and too often use themselves as model users. They therefore program for a young, logical, high end user familiar with computers (96).

By altering the way that a data base was queried, Greene and co-workers (97) were able to remove logical reasoning ability as a prerequisite for querying a data base. Dumais and

Landauer (98) altered search techniques for finding documents that removed the requirement for verbal fluency from the search method. These are illustrated in the graphs below. These alterations allowed most individuals to perform at the same high level.




Thomas K. Landauer (96) uses the following disclaimer:

*"It won't work - not until it is based on thorough task analysis in a realistic setting, not until it is mocked up and tried and revised, then again; then prototyped and tried and revised, then again; then experimentally implemented and evaluated and revised, then again".*

The Dallas VA Medical Center is upgrading the clinical information system in our intensive care units. The Central Office of the VA has mandated that the decentralized hospital computing program (DHCP) be redesigned with a graphical user interface (GUI, pronounced "GOOEY") design. This new integrated information system is presently under alpha testing and should be available within the next couple of years. The VA Central Office is committed to the development of an entirely electronic patient medical record.

For the Dallas County system a large sum of money has been designated for clinical information systems, the first of which will be placed in ambulatory care locations including Parkland and the community oriented primary care clinics (COPCs). It is imperative that each of us aid in the design of these information systems to ensure that their use helps us and helps our patients instead of adding work and decreasing productivity.

> *The Internal Revenue Service invested over $50,000,000 in PCs for its agents. The systems were supposed to help agents enter and look up data and make calculations more quickly and accurately. But the number of cases processed by each agent in a week went down by 40%* (96).

The key to our success involves user-centered development. Unless we, as physicians, are involved in the design, development, and deployment of clinical information systems these systems are likely to be significantly less than they should be and we will be doomed to making the same mistakes with our poor processors.

# References


1. Eddy DM: Clinical Decision Making. JAMA 263: 1265-1275, 1990.
2. Scientific American, June, 1946.
3. Schoolman HM: The Role of the Medical Library in Support of Clinical Decision Making. Presentation at the Symposium on Medical Information in the Information Age. Sponsored by the Association of the American Medical Colleges. Washington, DC, March 7, 1985.
4. Wilson RA: Right where you are sitting now: Further tales of the illuminati. Ronin Publishing, Berkley, CA, 1992.
5. Durack DT: The weight of medical knowledge. N Engl J Med 298: 773-775, 1978.
6. Madlon-Kay DJ: Weight of medical knowledge: Still gaining. N Engl J Med vol321: 908, 1989.
7. Cochrane AL: Effectiveness and efficiency. Random reflections on health services. Nuffield Provential Hospital's Trust. London (reprinted in 1989), 1972.
8. Grant A: Randomized trials in perinatology: Major achievements and future potential. Doing More Good Than Harm: The evaluation of health care interventions. Editors Warren KS, Mosteller F, The New York Academy of Sciences, New York, 1993.
9. Shrier R: Multum Expert Medical Information, Multum Information Services, Inc., 1996.
10. Williamson JW, German PS, Weiss R, Skinner EA, Bowes F: Health science information management in continuing education of physicians: A survey of US primary care practitioners and their opinion leaders. Ann Intern Med 110: 151-160, 1989.
11. Chalmers, Iain. 1993. The Cochrane Collaboration: Preparing, Maintaining, and Disseminating Systematic Reviews of the Effects of Health Care in Doing More Good Than Harm: The Evaluation of Health Care Interventions, editors Kenneth S. Warren and Frederick Mosteller, The New York Academy of Sciences, New York, 1993.
12. Minsky M: Will robots inherit the earth? Scientific American 271(4): 109-113, 1994.
13. Luria AR: 1987. The mind of a mnenomist: A little book about a vast memory. Harvard University Press, Cambridge, MA, 1987.
14. Miller G: The magical number 7, $\pm$ 2: Some limits on our capacity for processing information. Psychol Rev 63(2): 81-97, 1956.
15. Welford AT: Single-channel operation in the brain. Acta Psychologica 27: 5-22, 1967.
16. Kahney H: Problem Solving: Current Issues. Second edition. Open University Press, Philadelphia, PA, 1993.
17. Simon, HA. 1974. How big is a chunk? Science 183:482-488.
18. Simon, HA and WG Chase. The mind's eye in chess. In Models of Thought, HA Simon (Ed.) Yale Univ. Press, New Haven, Conn. 1979.
19. Moskowitz, AJ, BJ Kuipers, and AP Kassirer. 1988. Dealing with Uncertainty, Risks, and Tradeoffs in clinical decisions: a cognitive science approach. Ann Intern Med 108:435-449.
20. Nisbett, RE and TD Wilson. 1977. Telling more than we know: verbal reports on mental processes. Psychol Rev 84:231.

21. Bordage G, Zacks R: The structure of medical knowledge in the memories of medical students and general practitioners: Categories and prototypes. Med Educ 18: 406-16, 1984.
22. Diamond GA, Pollock BH, Work JW: Clinician decisions and computers. In Decision Support Systems in Critical Care, eds. Shabot MM, Gardner RM. Springer-Verlag, New York, 1994.
23. Bordage G, Allen T: The etiology of diagnostic errors: Process or content? An explatory study. Proceedings of the 21st Annual Conference on Research in Medical Education, pp. 171-176, 1982.
24. Bakwin H: Pseudodoxia pediatrica. N Engl J Med 232: 691-697, 1945.
25. Hingston DM, Irwin RS, Pratter MR, et al: A computerized interpretation of arterial pH and blood gas data: Do physicians need it? Respir Care 27: 809-815, 1982.
26. Drinkwater BL: Performance of civil aviation pilots under conditions of sensory input overload. Aerosp Med 38: 164-168, 1967.
27. TRW uses AI tools to control system problems. Aviation Week and Space Technology. February 7: 79-81, 1986.
28. Morris A: ARDS and new modes of mechanical ventilation: Reducing the complications of high volume and high pressure. New Horizons 2(1): 19-33, 1994.
29. Miller G: The magical number $7 \pm 2$: Some limits on our capacity for processing information. Psychol Rev 63(2): 81-97, 1956.
30. Wickelgren WA: ACTA Psychol 41: 67-85, 1977.
31. Kirch, W. and Schafii, C. Reflections on misdiagnosis. *Journal of Internal Medicine* 235:399-404, 1994.
32. Kingsford, D.P. A review of diagnostic inaccuracy. *Medicine, Science & the Law* 35:347-351, 1995.
33. Smith R: When Things Go Wrong. Br Med J 293: 461-2, 1986.
34. Harvard Medical Practice Study: Patients, Doctors and Lawyers: Medical Injury, Malpractice Litigation, and Patient Compensation in New York. Harvard Medical Practice Study, Boston, 1990.
35. National Center for Health Statistics: Vital Statistics of the United States, 1990. Washington, DC, Public Health Service, 1993.
36. Bates DW, et al: ADE Prevention Study Group. Incidence of Adverse Drug Events and Potential Adverse Drug Events, Implications for Prevention. JAMA 274: 29-34, 1995.
37. Evans RS, Classen DC, Stevens LE, et al: Using a hospital information system to assess the effects of adverse drug events. Proc Annu Symp Comput Appl Med Care 17: 161-165, 1993.
38. Asimov I: "The relativity of being wrong." The Magazine of Fantasy and Science Fiction, October 1986.
39. Simes RJ: Publication bias: The case for an international registry of clinical trials. J Clin Oncol 4: 1529-1541, 1986.
40. Feussner JR, White LJ: The clinical efficacy assessment program of the American College of Physicians. Doing More Good Than Harm: The evaluation of health care interventions. Editors Warren KS, Mostellar F, The New York Academy of Sciences, New York, 1993.

41. British Thoracic Society: Guidelines for the management of community-acquired pneumonia in adults admitted to hospital. Br J Hosp Med 49: 346-50, 1993.
42. Niederman MS, Bass Jr, JB, Campbell GD, et al: Guidelines for the initial management of adults with community-acquired pneumonia: Diagnosis, assessment of severity, and initial antimicrobial therapy. Am Rev Respir Dis 148: 1418-26, 1993.
43. Bartlett JG, Mundy LM: Community-acquired pneumonia. NEJM 333: 1618-1624, 1995.
44. Tobian L, Brunner HR, Cohn JN, Gavras H, Laragh JH, Materson BJ, Weber MA: Modern strategies to prevent coronary sequelae and stroke in hypertensive patients differ from the JNC V consensus guidelines. Am J Hypertens 7: 859-72, 1994.
45. Ayanian JZ, Hauptman PJ, Gaudagnoli E, Antman EM, Pashos CL, McNeil BJ: Knowledge and practices of generalist and specialist physicians regarding drug therapy for acute myocardial infarction. N Engl J Med 331: 1136-1142, 1994.
46. Pestotnik, SL, DC Classen, RS Evans, and JP Burke. 1996. Implementing antibiotic practice guidelines through computer-assisted decision support: clinical and financial outcomes. Ann Intern Med 124:884-890.
47. Weinberger M, Oddone EZ, Henderson WG, et al: 1996. Does increased access to primary care reduce hospital readmissions? N Engl J Med 334: 1441-1447, 1996.
48. Brook RH: Using scientific information to improve quality of health care. Doing more good than harm: The evaluation of health care interventions. Edited by Warren KS, Mostellar F, The New York Academy of Sciences, New York, 1993.
49. Sharpe VA, Faden AI: Appropriateness in patient care: A new conceptual framework. The Milbank Quarterly 74: 115-138, 1996.
50. Morris, AH, RM Gardner, and TD East. 1995. Clinical applications of Computers in monitoring in Critical Care Monitoring editors: RL Levine and RE Fromm, Mosby, St. Louis, MO.
51. Committee on Improving the Medical Record, Institute of Medicine: The computer-based patient record: An essential technology for health care. Washington, DC, National Academy Press; 1991.
52. Tierney WM, Miller ME, Overhage JM, McDonald CJ: Physician inpatient order writing on Micro-computer work stations: Effects of resource utilization. JAMA 269(3): 379-383, 1993.
53. Osheroff JA, Forsythe DE, Buchanan BG, Bankowitz RA, Blumenfeld BH, Miller RA: Physician's information needs: Analysis of questions posed during clinical teaching. Ann Intern Med 114: 576-581, 1991.
54. Weed LL: Medical records that guide and teach. N Engl J Med 278: 593- 600, 1968.
55. Tierney WM, McDonald CJ, Martin DK, Hui SL, Rogers MP: Computerized display of past test results; effect on outpatient testing. Ann Intern Med 107: 569-574, 1987.
56. Tierney WM, McDonald CJ, Hui SL, Martin DK: Computer predictions of abnormal test results; effects on outpatient testing. JAMA 259(8): 1194-1198, 1988.
57. McDonald CJ: The Use of a Computer to Detect and Respond to Clinical Events: Its Effect on Clinician Behavior. Ann Intern Med 84: 162-167, 1976.
58. McDonald, CJ: Protocol Based Computer Reminders, Quality of Care and the Non Perfectability of Man. NEJM 295: 1351-1355, 1976.

59. Schroeder SA, Myers LP, McPhee SJ, Showstack JA, Simborg DW, Chapman SA, Leong JK. The failure of physician education as a cost containment strategy: Report of a prospective control trial at a university hospital. JAMA 252: 225-230, 1984.
60. Williams SV, Eisenberg JM, Kitz DS, Carroll JG, Beck LH, Rubin SI, Ruff GE: Teaching cost effective diagnostic test use to medical students. Medical Care 22:535-542, 1984.
61. Morris, A.H., Wallace, C.J., Menlove, R.L., Clemmer, T.P., Orme, J.F.,Jr., Weaver, L.K., Thomas, F., East, T.D., Pace, N.L., and et al., Randomized clinical trial of pressure-controlled inverse ratio ventilation and extracorporeal CO2 removal for adult respiratory distress syndrome [see comments] [published erratum appears in Am J Respir Crit Care Med 1994 Mar;149(3 Pt 1):838]. *American Journal of Respiratory & Critical Care Medicine* 149:295-305, 1994.
62. East, T.D., Bohm, S.H., Wallace, C.J., Clemmer, T.P., Weaver, L.K., and Orme, J.F.,Jr. A successful computerized protocol for clinical management of pressure control inverse ratio ventilation in ARDS patients. *Chest* 101:697-710, 1992.
63. Wallace CJ, Franklin MA, Kinder T, Sailors RM, Carlson D, Bradshaw R, Morris AH: Medical Informatics Academia and Industry: A symbiotic relationship that may assure survival of both through health care reform. SCAMC Proceedings, New Orleans, LA, November 1995.
64. Knaus WA, Draper EA, Wagner DP, Zimmerman JE: APACHE II: A severity of disease classification. Crit Care Med 13(10): 818-829, 1985.
65. Knaus WA, Draper EA, Wagner DP, Zimmerman JE: An evaluation of outcome from intensive care in major medical centers. Ann Intern Med 104: 410-418, 1986.
66. Sivertson SE, Meyer TC, Hansen R, Schoenenberger A: Individual physician profile: Continuing education related to medical practice. J Medical Education 48: 1006-1012, 1973.
67. Manning PR, Lee PV, Clintworth WA, Denson TA, Oppenheimer PR, Gilman NJ. 1986. Changing prescribing practices through individual continuing education. JAMA 256: 230-232, 1986.
68. Berwick DM, Coltin KL: Feedback reduces test use in a health maintenance organization. JAMA 255: 1450-1454, 1986.
69. Lindberg DAB, Siegal ER, Rapp BA, Wallingford KT, Wilson SR: 1993. The use of MEDLINE by physicians for clinical problem solving. JAMA 269: 3124-3129, 1993.
70. Veenstra RJ, Gluck EH: A clinical librarian program in the intensive care unit. Critical Care Medicine 20(7): 1038-1042, 1992.
71. Leape LL, Bates DW, Cullen DJ, et al: Systems analysis of adverse drug effects. JAMA 274(1): 35-43, 1995.
72. Bradshaw KE, Gardner RM, Clemmer TP, et al: Physical decision making: Evaluation of data used in a computerized ICU. Int J Clin Monit 1: 81, 1984.
73. Bradshaw KE, Gardner RM, Klemmer TP, Orme Jr., JF, Thomas F, West BJ: Physician Decision Making: Evaluation of data used in a computerized ICU. Decision Support Systems in Critical Care, eds Shabot MM, and Gardner RM. Springer-Verlag, 1994.
74. Shabot MM, LoBue M, Leyerle BJ, and Dubin SB: Decision support alerts for clinical laboratory and blood gas data. Int J Clin Monit Comput 7: 27-31, 1990.
75. Blois MS: Clinical judgement and computers. NEJM 303: 192-7, 1980.

76. Wulff HR: Rationale diagnosis and treatment. Blackwell Scientific Publications, Oxford, 1976.
77. Berner ES, Webster GD, Shugerman AA, et al: Performance of 4 computer-based diagnostic systems. N Engl J Med 330: 1792-1796, 1994.
78. Kassirer JP: 1994. A report card on computer-assisted diagnosis - the grade: C. N Engl J Med 330: 1824-1825, 1994.
79. Tong DA: Weaning Patients from Mechanical Ventilation: A Knowledge Based Approach. In Miller RA, editor: Symposium on Computer Applications in Medical Care, Washington, DC, IEEE Computer Society Press, pp. 79-85, 1990.
80. Tong DA: Weaning Patients from Mechanical Ventilation: A Knowledge Based System Approach. Comput Meth Progr Biomed 35 (4): 267-278, 1991.
81. Seiver A and Holtzman S: Decision Analysis: A Framework for Critical Care Decision Assistance. Int J Clin Monit Comput 6: 137-156, 1989.
82. Pauker, AG and JP Kassirer. 1987. Decision Analysis. N Engl J Med 316:250-258.
83. Berwick DM: Harvesting knowledge from improvement. JAMA 275(11): 877-878, 1996.
84. Health Care Leadership Council. US Health Care Reform: A Primer.
85. Thomas L: Notes of a Biology Watcher. The Technology of Medicine. NEJM 285: 1366-8, 1971.
86. Weisbrod BA: Economics and Medical Research. Washington, DC: American Enterprise Institure, 1983.
87. Wasson JH, Sox HC: Clinical prediction rules: Have they come of age? JAMA 275 (8): 641-642, 1996.
88. Wasson JH, Sox HC, Neff RK, Goldman L: Clinical prediction rules: Application and methodologic standards. NEJM 313: 793-799, 1985.
89. Hammond J, Johnson MH, Ward CG, et al: Clinical evaluation of a computer based patient monitoring and data management system. Heart-Lung 20 (2): 119-124, 1991.
90. Hammond J, Johnson HM, Vares R, et al: A qualitative comparison of paper flow sheets versus a computer based clinical information system. Chest 99 (1): 155-157, 1991.
91. Bradshaw KE, Sittig DF, Gardner RM, et al: Computer based data entry for nurses in the ICU. MD Comput 6(5): 274-280, 1989.
92. Hilberman M, Kamm B, Tarter M, et al: An evaluation of computer based patient monitoring at Pacific Medical Center. Comput Biomed Res 8: 447-460, 1975.
93. Johnston ME, Langton KB, Haynes RB, Mathieu A: Effects of computer-based clinical decision support systems on clinician performance and patient outcome: A critical appraisal of research. Ann Intern Med 120: 135-142, 1994.
94. Alsafadi Y, Liu Sheng OR, Martinez R: Comparison of communication protocols in medical information exchange standards. Seventh Annual IEEE Symposium on Computer-Based Medical Systems, pp. 258-263, 1994.
95. Kuzmak PM, Norton GS, Dayhoff, RE: Using experience with bidirectional HL7- ACR-NEMA interfaces between the Federal Government HIS/RIS and commercial PACS to plan for DICOM. SPIE 2435: 132-143, 1995
96. Landauer TK: The trouble with computers: Usefulness, usability, and productivity. The MIT Press, Cambridge, 1995.

97. Greene SL, Gomez LM, Devlin SJ: A cognitive analysis of database query production. In Proceedings of the Human Factors Society, Santa Monica, CA: Human Factors Society, pp: 9-13, 1986.
98. Dumais ST, Wright AL: Reference by name vs location in a computer filing system. In Proceedings of Human Factors Society, Santa Monica, CA: Human Factors Society, pp: 396-402, 1991.