Xoetronics LLC
advanced instrumentation, diagnostic and digital technologies

Company Headquarters 113 Wertsville Road, Ringoes, NJ 08551
e-mail info@xoetronics.com
voice 908-788-8705

R&D Headquarters 121 Yancy Road, Madison, AL 35758
e-mail R&D@xoetronics.com
voice 256-895-9865

web site: http://xoetronics.com

Smart Medical Database

FAQ and Responses

(Updated February 9, 2002)

Table of Contents

1. Why is there a need for a Smart Medical Database?

With the enormous amount of medical information available, making a decision on a factual basis is difficult. Until now, a readily accessible compilation of collective medical wisdom simply did not exist: virtually all information, which found its way to efficient dissemination was of the research-results kind. While absolutely essential to support medical progress, this kind of data is of little value in the day-to-day practice of optimum medicine. Clinical experience forms the proper basis to informed decisions in health care, but it must be collected and carefully monitored. However, faced with a presenting patient, a physician has no practical way to know what the statistical experience of the last, say, 1000 patients so presenting has been. The same physician, having made a well-founded diagnosis, has no immediate, practical way to compare treatment choices and their relative outcomes for the last 1000 or so similarly diagnosed patients. But it is precisely this kind of data on which the best medical decisions must be made. The Smart Medical Database (SMDB) resolves this dilemma.

2. What is the Smart Medical Database?

The SMDB is an integrated software tool based on Artificial Intelligence and Bayesian probability statistics, with a real time, "dynamic" data structure, which provides a tool for the physician. The SMDB is not an expert system, which seeks to replace or at least replicate the judgment of the physician, rather the SMDB gives the physician the ability to observe the effects of various diagnosis and treatment regimens on a statistically significant portion of a relevant patient population.

3. Why is there a need for a physician's tool like the Smart Medical Database?

Currently, there is no readily accessible, timely body of clinical data that is useful in the practice of medicine. The SMDB provides a desperately needed, comprehensive inventory of clinical experiences involving patient presentation, test results, diagnoses, treatment protocols and eventual outcomes. Additionally, the SMDB integrates clinical experience with cost, payment data and strategic planning for truly informed and effective health care policy. Relevant data, useful in the making of day-to-day diagnostic and treatment decisions, will lower medical costs and provide better quality care. Studies by Dr. Wennberg at Dartmouth[1], and others, incontrovertibly show that better health care is not linearly linked to increasing health care expenditure. Based on these studies, use of the SMDB can result in a 25% savings in health care expenses[2].

4. There are many prior art computerized systems, which take information about symptoms and claim to assist in providing a diagnosis. How is the Smart Medical Database different?

In most of the prior art systems, the data does not come from "clinical experience" as it does with the SMDB. Rather for each disease, a specific list of symptoms is prepared. Each symptom is assigned a weight, which represents likelihood that the patient has the disease, given the symptom. This describes a method of diagnosing a disease. In the prior art systems even a minor difference of

---

[1] *Dartmouth Atlas Of Health Care in the United States*, 1998, American Hospital Publishing

[2] 1995 payments by HMO and non-HMO insurers was estimated at $210 + Billion in the United States alone.

a single data field would change the applicability of the statistical model used. This is essentially classical list processing based on a traditional Boolean approach to determine a diagnosis based upon the responses to the list of symptoms for each disease. The static list processing approach does not vary the diagnosis as a function of expanded clinical data and experience.

In contrast, the SMDB avoids the necessity of preparing a particular list of symptoms corresponding to a particular diagnosis by collecting the symptoms for actual clinical experience and correlating that information to subsequent observations. The SMDB, by accessing a collection of clinical medical data, makes the diagnosis and its corresponding probability of correctness functions based on an enormous body of accumulated clinical data and experience. The diagnosis is refined and updated based upon subsequent medical data obtained for a particular patient, and subsequent additions and updates to the collection of clinical medical data. Likewise, once a confident diagnosis (or set of probable diagnoses) is obtained, treatment decisions can then be made on the basis of optimizing the outcome for that patient, once again using present patient specific criteria and preferences in the face of treatment choice characteristics provided by the SMDB. In fact, by not having to specify a particular diagnosis model for a given disease, the symptoms and clinical treatments options can be used to essentially speak for themselves, dynamically presenting the physician with options and probabilities that are most relevant for the particular patient presented to the physician. This ability to adapt, in "real time" is a particularly unique feature of the SMDB.

In short, prior art computerized systems attempt to give algorithms or prescriptions for interpreting evolving real time patient data, virtually always based on correlations with data that have proved 25% capricious, and help direct a more meaningful search for more. In contrast, the SMDB directs diagnostic evaluation and treatment selection on the basis of previous relevant outcomes. The difference lies in the fact that the data needed to make solid decisions have not been available in a clinically useful form previous to the SMDB.

5. Isn't the Smart Medical Database an expert diagnosis system?

No, the SMDB is actually a data base tool, facilitating access to relevant data within a vast collection of information. Essentially, the SMDB can find the relevant data, which would otherwise be lost in the enormous amount of information that is not relevant, recovering the relevant data signal from the noise presented by the non-relevant data. The difference between an expert system and a data base tool such as the SMDB lies not so much in the execution of the idea, but in its ultimate goals. The essence of an expert system is found in its ultimate goal of imitating, mimicking, and, at least theoretically, replacing, an expert. A tool's weight in a complex decision making process is constantly under review by a human expert. The SMDB may be seen as a very sophisticated portal that allows access to data, and to allow its meaningful input into expert decisions, where these data either totally unavailable previously, or available in a manner, or only with such effort and time investment, that its use was not a practically meaningful option. Expert systems code specialized knowledge to recommend decisions. Expert systems crisply dichotomize rules at artificial boundaries. Expert systems can solve only general classes of problems and cannot handle uncertainty and imprecise knowledge very well, which is typical of a medical problem. While clinical experience forms the proper basis to informed decisions in health care, it must be collected and carefully monitored. It is precisely this kind of data, which the best medical decisions are made on.

6. How is the Smart Medical Database implemented?

In order to determine a probability of successful diagnosis as a function of clinical data and experience in which a diagnosis is refined and updated based upon subsequent medical data obtained for a particular patient or subsequent additions and updates to the collection of clinical medical data, Bayesian methods and fuzzy logic based methods are utilized. Bayesian systems incorporate information used by expert systems but not vice versa: both types of systems use data common to both, and data specific to either. A Bayesian based system is more general than an expert system. The Bayesian approach gives the posterior probability density as a function of the prior density functions, while a utility function is used to order preferences. These parameters are defined carefully in terms of the SMDB structure and possible queries. Decisions relying on the Bayesian approach use the following:

- a space of possible actions by the decider – the clinician;
- a space of possible patient states, diagnoses for example;
- a collection of possible experiments, or tests, or findings;
- a space of possible outcomes for these tests, experiments; and,
- a utility function defined over all the possible consequences, which defines outcome preferences.

The above are then used to "condition" the prior probability density function as more results (information) become available, yielding the posterior probability.

The SMDB is designed as a "dynamic" structure which continually updates itself with the addition of new clinical experience as well as a host of other possible sources of information such as newly found research results, and specific other information about the particular patient. The dynamic aspect of the structure is built into it via Conditioned Probabilities. The key word is "conditioned" since with each entry of new information the Conditioned Probability changes. And, the patient specific information is entered interactively in real time. So in a very real sense, the SMDB is a Dynamic, Interactive, Real Time Structure.

On the other hand, expert systems are static structures. They essentially operate, or return probabilities, based on the frequencies with which certain symptoms arise. There is nothing "conditioned" about the probabilities these structures return. They are not interactive and cannot automatically (in real time) adjust for specific information about the patient other than look in a systematic way to match up symptoms.

Simply put, the SMDB uses inductive logic and Boolean Medical Databases use deductive logic. The SMDB's inductive approach almost demands that the database be dynamic and capable of interaction in real time.

7. How is the Bayesian approach applied to the Smart Medical Database to make a diagnosis?

This can be best understood by looking at a simple example. Mr. John Doe visits his doctor with a health problem. John explains his health problem to the doctor. Based on the information received, the doctor did not come to a firm conclusion as to the nature of the illness or the treatment thereof. He decides to consult with the SMDB system. He supplies the identification of the patient and other routine details. He then selects the query option, which presents a standard list of symptoms.

The doctor highlights all the relevant symptoms as well as enters additional symptoms—prior information, **I**, about the specific patient—that are not part of the standard list. The SMDB then presents a list of possible diseases consistent with the symptoms in the form of a table containing the sample of patients used for the response and the probability of successful diagnosis. Also presented are the patient's symptoms that matched with the data used in arriving at the correlations indicated.

The probability for the disease is derived using standard statistical tools and Bayes theorem. Let **H** be the hypothesis whose validity is being examined, (all boldfaced symbols imply vector quantities,) in other words,

$\mathbf{H}(\mathbf{x})$: The symptoms and test results imply disease(s) $\mathbf{x}$.

Since the set of symptoms and test results are never exhaustive, the validity of **H** in general has to be statistical based on evidence. The confidence in the validity of $\mathbf{H}(\mathbf{x})$ is summarized by the quantity $P(\mathbf{H}(\mathbf{x})|\mathbf{DI})$ defined by

$P(\mathbf{H}(\mathbf{x})|\mathbf{DI})$ = Probability Density(confidence) that $\mathbf{H}(\mathbf{x})$ is true.

Conditioned on the evidence **D** and prior information **I**,

D: Set of all symptoms and test results, and $0< P(\mathbf{H}(\mathbf{x})|\mathbf{DI}) < 1$.
$P(\mathbf{H}(\mathbf{x})|\mathbf{DI}) =1$ implies 100% confidence in $\mathbf{H}(\mathbf{x})$
$P(\mathbf{H}(\mathbf{x})|\mathbf{DI}) = 0$ implies $\mathbf{H}(\mathbf{x})$ is ruled out

The estimation of $P(\mathbf{H}(\mathbf{x})|\mathbf{DI})$ is done using Bayes theorem:

$$P(\mathbf{H}(\mathbf{x})|\mathbf{DI}) = P(\mathbf{D}|\mathbf{H}(\mathbf{x})\mathbf{I})\ P(\mathbf{H}(\mathbf{x})|\mathbf{I})\ /\ P(\mathbf{D}|\mathbf{I})$$

where $P(\mathbf{H}(\mathbf{x})|\mathbf{I})$ = Prior probability density that $\mathbf{H}(x)$ is true based on previous symptoms and tests on patients other than the one currently under examination, $P(\mathbf{D}|\mathbf{H}(\mathbf{x})\mathbf{I})$ = Likelihood of $\mathbf{H}(\mathbf{x})$ based purely on current evidence. Confidence based on current symptoms and tests, and $P(\mathbf{D} \mid \mathbf{I})$ = Probability Density for **D**. (See Appendices A and B for a general discussion and example.)

The system offers an estimation of $P(\mathbf{H}(\mathbf{x})|\mathbf{I})$ using all the information available in its database. It will also estimate the likelihood based on current evidence. These algorithms are new and unique to the medical community. Although theoretical Bayesian techniques have been known for some time, their actual application to any practical area is still in its infancy, and their application to medicine is virtually non-extant. A large portion of the reason for this lies with technology itself. Only very recently hardware and software necessary to handle very large databases became available. The system can also recommend what symptoms to look for and what tests to order based on previous case histories.

The query responses are a function of the number of symptoms that matched with the list of known symptoms for a particular disease state, the history of the patient based on this match of the symptoms, and commonality characteristic with symptoms in other illnesses. Unless requested this response uses the entire sample of patient data on the database. The form also has options to choose specifics that would restrict the sample by cross matching to selected demographic and medical variables. The software to choose a sub-sample of the entire data set and come up with a new real time response would then use this later option.

The system would also query the doctor if he wants to see an additional list of symptoms that might increase the probability of a particular diagnosis. The doctor in consultation with the patient might look for additional symptoms over a stipulated time period, or order tests that might reveal additional symptoms. He would then repeat, if necessary, the above procedure to come to a firm decision as to the nature of the illness. At this stage we assume that the doctor in consultation with the patient and through his use of the software and additional tests, has come to a decision regarding the illness.

8. How is the Smart Medical Database used to determine a choice of treatments?

The system presents details about the samples of patients and the response to a particular treatment as a function of time. A number of statistical processing results are shown depending on the queries posed up by the doctor (chosen out of a list or entered by the doctor) such as probability of cure versus probability of failure of treatment. This is done interactively as a function of time. The system will give a set of alternatives as a guide to proceed. It will also provide a list of contacts for consultation should it prove necessary. It will also search additional databases should further processing be necessary. If the doctor incorporates the patient treatment and response into the software, it can in real time compare and match with similar patient scenarios and give useful pieces of information such as recommendations for midcourse change of medicines and treatment.

9. How is the Smart Medical Database updated?

The system is updated continually as additional tools are available. The underlying database is updated on a regularly as additional information is available. It will have a hierarchy of network connections to other databases available all over the world. It will provide patient histories to other agencies with the approval of the doctor and the patient. For example, a person may go to a different state or country and an unfortunate illness might necessitate fast access to this information. The system also has access to libraries of on line information to all the latest developments in the relevant fields. The various functions of the SMDB happen in real time, allowing immediate analysis and decision making at the time and place of physician contact. Such a massive integration of clinical data will allow relatively effortless retrospective studies of diagnostic accuracy, treatment efficacy, and safety. It will permit, within confidentiality legislation and guidelines, the convenient access of individual patient histories and data by those people and agencies entitled to that data.

10. How does the Smart Medical Database distinguish which data is relevant?

From a Bayesian perspective ALL possible causal data, determined by prior information stored in the database concerning patients with overlapping symptoms, prior information about the patient's physical condition, lab testing and the physician's subjective input, is relevant and this is precisely the power of the method. Unlike query systems, the SMDB does not look to correlation alone to determine the probability of disease, although statistical correlation can be, and is, incorporated into the method as prior information. Obviously, the more the database is employed the more data there will be to look for correlation. Nevertheless, the SMDB can be initiated and operated effectively without huge volumes of data. And, in a very real sense, the more the SMDB is used it becomes "stronger" and much "smarter". It does indeed learn from its experiences.

11. Can the Smart Medical Database be used for veterinary medicine?

The SMDB is equally well suited for use in veterinary medicine, as the issues confronting a veterinarian are analogous to those confronting a physician. Although the patient typically cannot communicate well with the veterinarian, clinical experience is the basis to inform all decisions in veterinary medicine. It is precisely this kind of data on which the best veterinary medical decisions must be made.

12. Can the fundamental principles of the Smart Medical Database be demonstrated prior to full implementation of the system?

A practical demonstration of principles involves starting with simple well-defined clinical cases. A simplified user interface will initially confine the user to these established test selected cases. The user-interface either consists of:

- a web like entry of few keywords; or,
- a pre-established set of data elements from which the user have to select a subset.

The user-selected response would then be a prioritized list of case studies with corresponding Bayesian thresholds. The user-selected response can also be an inconclusive, with insufficient data to make a recommendation and thus prompts the user to either enter more data or answer new interactive questions on the screen.

13. How can a full prototype Smart Medical Database system be implemented?

One recommended model for development is to select initially one or two medical domains (such as the Mayo Clinic) and engage via incentives leading experts in such domains to participate as users, providers of case studies and advisors. These experts would be essential for the validity, usability, and marketing credibility of such a database. Also, these experts are very needed to identify any relevant open source data that can be used as well as advise on the inconsistencies, gaps and noise that is likely to exist among the various data collected.

The supervised learning and inclusion of new case studies is to be initiated securely by only the medical experts. A specialized user interface needs to be developed for such purpose where new relevant attributes are added to the schema of such a database. Because of the numerical integration needed for Bayesian computations, the Hidden Markovian algorithms required for supervised learning and the stringent performance needs of relational database, the platform is recommended to be a Linux or Unix server. It is anticipated that a minimum of two computer scientists, one designer, and one "field" medical person are needed for this development. A prototype server that can be incrementally scaled would be available for demonstration within one year of the availability of resources.

14. Doesn't the Smart Medical Database system require a very large amount of clinical data in order to be useful?

While it is true that a large amount of data will provide optimal functionality, by initially selecting a limited number of medical domains, through the use of leading experts and suitable incentives for their participation, the experts will identify any relevant open source data that can be used as well as advise on the inconsistencies, gaps and noise that is likely to exist among the various data collected. Also, a classification scheme (taxonomy) can be introduced to limit query searches to only relevant subsets of the data thus minimizing the search space and easing the need for collecting all data at once. This approach will allow the SMDB system to be phased in, starting with simple well-defined clinical cases, building and expanding initially in a limited number of medical domains and being able to evolve and expand, utilizing existing data collections and new resources.

15. There are differences amongst the users and the differences amongst requirements of the same user of the Smart Medical Database, how will these differences be addressed?

A goal of the SMDB is providing useful integration of clinical, administrative, and payment data. This data is used, through controlled access, by a variety of individuals for different purposes. The differences amongst the users and the differences amongst requirements of the same user can be illustrated best by looking at the clinical use of and three incremental levels of the SMDB.

Level I - One use to which a clinician may put the SMDB is a retrospective analysis of associated findings (signs, symptoms, complaints, test results) with diagnoses. The clinician can then use this information in determining how to proceed. The clinician may be made aware of a correlation under appreciated before, or debunked of a false but believed correlation that would have received more weight (unwarranted) before reviewing the data. Hence the clinician's choices become more efficient and more accurate in terms of diagnoses. Then, after a working diagnosis has been established, a similar search can improve the efficiency of treatment selection to optimize variables for each patient case (longevity, comfort, intellectual functioning, mobility, etc.) given their respective importance in each patient case. So far, the SMDB is just another 'fact tool.'

Level II – The SMDB tool itself begins to perform active processing (initially just filtering) of the data, and presenting selective displays depending upon the inquiry. This level of sophistication is analogous to selective representation of flight data on visual displays for pilots. Compared to use described in Level I, this use assumes and demands a higher level of knowledge about the dynamics of medicine and will require some learning, supervised, unsupervised, or both.
Level III - This involves informing prospective research, in that the SMDB may point to avenues of work unappreciated before, and perhaps help foresee dead-end lines of work long before they might have been so determined before. Helping in designing drugs (purpose built molecules) and nanotechnology are potentially two of this kind of use.

It is important to note that Level I is about as far from an expert system function as one can get. The Level III SMDB system is very advanced using fuzzy system and neural network techniques. Level II can be viewed as pushing the Level I envelope or simplified use of the Level III SMDB system.

The different levels build upon each other with increasing complexity to implement as a functional application. Essentially, Level I will yield the most immediate results and consume the least resources. Levels II and III will take progressively longer and consume progressively more resources. The use of IDT's proprietary use of Bayesian systems finds critical application at all three of Levels. Level I can be implemented by using an Oracle like database that fuses deterministic diagnostic data with administrative, cost, and billing data. However, Levels II and III require something like a neural net approach, which cannot be supported by an Oracle like database and thus requires innovative code development.

16. What are some ways that the Smart Medical Database can initially be marketed?

One business model uses co-branding with a recognized medical organization. The organization is selected initially for one or two medical domains. Leading experts in such domains participate as users, providers of case studies and advisors through incentives. The co-branding essentially offers the organization the ability to publicize and/or market their proven expertise in a selected medical domain.

A powerful marketing tool for on-line databases, that has been proven in the legal profession in the United States by Lexis® and WestLaw® in the legal profession is to provide students with essentially free, although limited, access to the database. Upon graduation, the students are both quite familiar with the system and have become somewhat dependent upon it. The SMDB would be offered to medical students, initially at the co-branding medical organization, and subsequently expanded to other medical schools.

Complementary to co-branding with a medical organization, would be to establish relationships with key health insurance organizations.

Digital medical record systems are evolving and being implemented. In Australia, a nation wide government lead effort is underway regarding Telemedicine, which among other things, creates a virtual medical record, freeing patients and health care providers from dependence on paper. Currently, the majority of health care records in Australia, and virtually every other country, exist as discrete paper-based entities held at a variety of different locations, resulting in a fragmented picture of individuals' health needs and health histories. Increasingly, the potential of electronic health records in improving efficiency, safety and quality of care over paper-based systems is being recognized across the health sector. In Australia, the National Electronic Health Records Taskforce has proposed the concept of a national health information network (HealthConnect) that would allow personal health information to be collected, safely stored and exchanged online. Under Health*Connect*, health-related information about an individual would be collected in a standard, electronic format at the point of care (such as at a hospital or a general practitioner's clinic). This information would take the form of event summaries, not all the notes that a health care provider may choose to keep about a consultation. While implementation of digital medical record systems in the US is slower, they are currently are being implemented in stages in a number of hospitals. Therein exists a significant opportunity to co-brand with the digital medical record software/system providers.

17. What are some ways that the Smart Medical Database would be accessed?

The SMDB can be accessed through any of a variety of remote access means, including virtually any form of internet/web access device. One embodiment uses a personal digital assistant for accessing the SMDB. This enables the physician to have immediate access to a collection of clinical medical data and a means for querying the collection of clinical medical data to determine a diagnosis and probability of successful diagnosis for a patient based upon assessment of the patient to obtain medical data. Additionally, the physician can enter prescription information, which is then transmitted to the SMDB, analyzed, and immediately reported back to the physician, indicating any anomalies. This system also produces a readable hardcopy and can even directly transmission the prescription to a provider pharmacy for dispensing.

18. How can the Intellectual Property rights for the Smart Medical Database computer implemented method and underlying algorithms be protected?

Traditionally, computer software, a process based on a mathematical formula, and a method of doing business were three types of intellectual property that were excluded from patentability. The U.S. courts have referred to these exceptions as the "mathematical algorithm" exception and the "business method" exception. The U.S. Supreme Court has upheld Federal Court decisions essentially indicating that computer software, a process based on a mathematical formula, and a business method are patentable subject matter. On January 11, 1999, the U.S. Supreme Court left intact the decision of the Federal Circuit Court in *State Street Bank*, which held that a computer system designed to implement an investment structure is patentable. This decision has opened the door to new and powerful ways to protect these kinds of intellectual property. This trend leads to the interesting situation in which software can be protected simultaneously by patent, copyright, and trade secret laws. A patent can protect a software idea, while copyright and trade secret laws can protect the details of the software as an unpublished work. This multifaceted legal protection can provide significant financial and competitive advantages for the software developer, as well as establish barriers. Essentially, computer software that utilizes a mathematical formula and methods for doing business are now patentable subject matter, which means that we have new and powerful ways to protect these kinds of intellectual property. Outside of the U.S. computer software that utilizes a mathematical formula is increasingly being recognized as patentable subject matter.

19. Please provide an Intellectual Property summary covering the Smart Medical Database.

Pending patents include:

- US Application Serial No. 09/495,185 filed February 1, 2000, entitled Method And System For Accessing Medical Data, which claims priority from US Provisional Application Serial No. 60/141,191 filed on June 25, 1999;

- US Application Serial No. 09/553,162 filed April 19, 2000, entitled Method And System For Accessing Medical Data, which is a Continuation-In-Part of Application Serial No. 09/495,185, which claims the benefit of US Provisional Application Serial No. 60/141,191;

- PCT Application No. PCT/US00/10727 filed April 20, 2000, entitled Method And System For Accessing Medical Data, electing all member countries except the US, which claims priority from Application Serial No. 09/553,162 filed April 19, 2000, Application No. 09/495,185 filed February 1, 2000, and Provisional Application Serial No. 60/141,191 filed on June 25, 1999; and,

- US Application Serial No. 09/836,067 filed on April 19, 2001, entitled Smart Medical Database Interfaces, which claims the benefit of US Provisional Application Serial No. 60/198,590 filed on April 19, 2001 and is a Continuation-In-Part of Application Serial No. 09/553,162 filed on April 19, 2000.