



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Case Detection Algorithms

Fu-Chiang Tsui, Michael M. Wagner, and Jeremy Espino

RODS Laboratory, Center for Biomedical Informatics, University of Pittsburgh, Pittsburgh, Pennsylvania

Richard Shephard

Australian Biosecurity Cooperative Research Centre for Emerging Infectious Disease, Brisbane, Australia

1. INTRODUCTION

This chapter focuses on algorithmic methods for case detection. As we discussed in Chapter 3, the objective of case detection is to notice the existence of a single case of a disease. Case detection is a core activity of biosurveillance. Detection of an outbreak usually depends on detection of individual cases, although it is also possible to detect an outbreak from data other than case data (e.g., retail sales of thermometers or satellite imagery).

For most of the 20th century, governmental public health relied almost exclusively on the astute clinician and the clinical laboratory to detect and report cases via the notifiable disease system. Algorithms for case detection did not exist, with the possible exception of case definitions discussed in Chapter 3, which are algorithms in the sense that they are formal specifications that a clinician or epidemiologist can follow to classify an individual as a suspected or confirmed case.

Around the beginning of the 21st century, emerging diseases and the threat of bioterrorism began to stress the capabilities of the notifiable disease system. Health departments endeavored to improve it by creating web-based forms for disease reporting and electronic laboratory reporting. Their goal is to increase completeness of reporting and decrease time latencies inherent in paper-based reporting.

After the anthrax postal attacks of 2001, health departments redoubled their efforts to improve disease reporting by increasing the “astuteness” of clinicians through education and training (Gerberding et al., 2002, Hughes and Gerberding, 2002). A specialist in infectious diseases had detected and reported the first case of inhalational anthrax (Kolata, 2001); thus, the conventional wisdom was that the notifiable disease

system worked. For the foreseeable future, increasing the ability of front-line physicians to diagnose rare diseases through better training was our best defense against bioterrorism. The conventional wisdom was correct insofar as case detection by clinicians is important in outbreak detection. However, it was overly sanguine about the ability of training to improve the existing capability.¹

There is a limit to which additional training can improve a clinician’s ability to detect and report rare diseases to a biosurveillance organization. Humans are not perfectible in this manner, as noted first by Dr. Clem McDonald, who entitled his seminal paper on physician performance *Protocol-Based Computer Reminders, the Quality of Care and the Non-Perfectibility of Man* (McDonald, 1976b).²

His research, conducted in the early 1970s at the Regenstrief Institute in Indiana, demonstrated that even for common conditions, such as diabetes and hypertension, and for common preventive measures, such as immunizations, physicians often failed to deliver required services. However, when reminded by a computer system that monitored electronic patient data about the need to vaccinate a *specific patient* or order a needed test, physicians complied with standards of care at twice the rate as when not reminded. When the system stopped reminding the physicians, however, their compliance rates quickly returned to baseline; thus, any “education” or “training” that the system provided had no lasting effect on compliance.

McDonald’s research spawned a new line of system-level thinking in medicine that continues to this day (Kohn et al., 2000, Leavitt, 2001, Yasnoff et al., 2004). McDonald concluded this influential paper with the observation that, although man is not perfectible, *systems* of care are.

1 Studies of continued medical education programs also suggest that the yield is low and the cost is high (Haynes et al., 1984, Leist and Kristofco, 1990, Williamson et al., 1989, McDonald, 1983). The one study of efficacy of training (of a web-based educational program) on physicians’ knowledge about diagnosis and management diseases caused by known weaponized biological agent showed no retention of information (chung et al., 2004).

2 We do not know whether the always playful Dr. McDonald and the sophisticated editors of the *New England Journal of Medicine* misspelled the word *perfectibility* intentionally. Also reinforcing the main conclusion of his research, it took a computer’s spell checker to bring this error to our attention.

2. PERFECTING CASE DETECTION

The best case detection system imaginable would be one in which every individual in a community is examined every morning by the best diagnostician in the world. This diagnostician would have all the time in the world to interview and examine each person. Since she would be examining everyone every day, she would notice patterns (clusters) of early illness in a community. Her awareness of patterns would appropriately bias her diagnostic thinking (and treatment) of individual patients. Physicians are taught (and reminded incessantly) “when you hear hoof beats, don’t think of zebras.” This adage is an informal statement that when the evidence available about a particular patient supports equally a diagnosis of either influenza or SARS (e.g., the patient has constitutional symptoms and no history of exposure to SARS), they should conclude that the diagnosis of influenza is far more likely than SARS. This diagnostician would also never fail to report immediately each fever, early syndromic presentation, or reportable disease to governmental public health.

This ideal scenario recognizes the importance of the knowledge, judgment, and skills that clinicians bring to bear on the diagnosis of disease. The importance of expert knowledge and judgment underlies the opinion, expressed by experts in public health after the anthrax letters in 2001, that there is no substitute for an astute clinician. This ideal scenario also recognizes that knowledge of the prevalence of disease in a population influences the diagnostic work up and management of individual patients.

Of course, such a system is not feasible due to the impossibility of having every person seen by the same diagnostician every day or the alternative of cloning the best diagnostician and ensuring that the clones could instantly share information about individuals they were seeing. Nevertheless, it represents a benchmark against which we can compare other schemes that are perhaps superior to current approaches.

In the next section, we discuss diagnostic expert systems, which are computer programs (algorithms) that embody the diagnostic knowledge and diagnostic skills of expert clinicians. Diagnostic expert systems are far less expensive than clinicians, never tire, and we can clone them at will. They make it reasonable to imagine a biosurveillance system in which a highly competent diagnostician examines thousands of individuals in a community every day with consistent diagnostic quality—reporting fevers, syndromes, and reportable diseases to a health department without fail and without delay. They make it possible to imagine a biosurveillance system in which the health department analyzes highly improved case data and communicates up-to-the-minute information about patterns of illness in the community back to the diagnosticians.

3. DIAGNOSTIC EXPERT SYSTEMS

Diagnostic expert systems are computer algorithms that automate the cognitive process of medical (or veterinary) diagnosis.

Many readers will be surprised to learn that researchers in the fields of artificial intelligence and medical informatics have been developing and fielding such systems since the 1960s. The first fielded system, developed by Dr. Homer Warner in Salt Lake City, provided diagnostic assistance for children with congenital heart disease (Warner et al., 1961, Warner et al., 1964). Congenital heart diseases are severe birth defects of the valves and structure of the heart. In the 1960s, the exact nature of heart malformations was very difficult to diagnose without invasive and risky angiographic procedures. Dr. F. Timothy de Dombal developed a similar system for the differential diagnosis of the acute abdomen, another high-stakes diagnostic problem. A surgeon must differentiate between conditions that require emergency surgery, such as appendicitis, and conditions, such as pancreatitis, for which surgery is relatively contraindicated (de Dombal et al., 1972, de Dombal et al., 1974, Wilson et al., 1975, de Dombal, 1975, Wilson et al., 1977, de Dombal, 1984, Adams et al., 1986, McAdam et al., 1990, de Dombal, 1990, de Dombal et al., 1993, American College of Emergency Physicians, 1994).

3.1. How Diagnostic Expert Systems Work: Data Collection

We can perhaps best explain how a diagnostic expert system works by analogy to the process that a physician uses to diagnose a patient. Like a physician, a diagnostic expert system begins by collecting information about the patient’s illness—symptoms, observations from physical examination, results from laboratory tests, risk factors for disease (e.g., travel to a foreign country) and pre-existing medical conditions (e.g., diabetes). We refer to this diagnostic information collectively as the findings. Of course, the computer usually does not interview the patient and (at present) never examines the patient. Rather, a physician interviews and examines the patient after which she or an assistant enters the findings into the program (e.g., Warner and Bouhaddou, 1994, London, 1998, Buchanan and Shortliffe, 1984, Miller et al., 1986, Shwe et al., 1991, Heckerman et al., 1992). Increasingly, diagnostic expert systems acquire findings automatically from clinical information systems (Aronsky et al., 2001, Burnside et al., 2004, McDonald et al., 1991). There are also examples of diagnostic expert systems that interview patients directly to obtain their medical histories (Pynsent and Fairbank, 1989, Wald et al., 1995).

3.2. How Diagnostic Expert Systems Work: Knowledge Representation

Like the physician, the diagnostic expert system is a storehouse of medical knowledge. A diagnostic expert system stores its medical knowledge in tables of diseases and their findings. There are typically a table of prevalences for each disease (e.g., Table 13.1) and tables with every finding of every disease that the system knows about (e.g., Table 13.2). The latter tables usually represent the strength of association between diseases and findings as conditional probabilities.

TABLE 13.1 Prior Probabilities and Prior Odds of FMD and MCD

Disease	P(Disease)	Odds(Disease)
FMD	0.001	0.001001
MCD	0.001	0.001001

Notes: The notation $p(\text{FMD})$ actually is shorthand for a probability of an event variable called FMD that can take the values true or false. So, $p(\text{FMD})$ actually stands for $p(\text{foot and mouth disease is present})$ or $p(\text{foot and mouth disease is not present})$. If we to write Table 13.1 out in its full form, it would have four rows corresponding to $p(\text{FMD is present}) = 0.001$, $p(\text{FMD is absent}) = 0.999$, $p(\text{MCD is present}) = 0.001$, $p(\text{MCD is absent}) = 0.999$. The laws of probability require that the sum of the probabilities for all the event variable outcomes will be equal to 1 (e.g., $p[\text{FMD is present}] + p[\text{FMD is absent}] = 1$, because it is a certainty that something is either present or it is absent), so knowledge engineers usually save space by only writing one row per disease. FMD, foot and mouth disease; MCD, mad cow disease.

3.3. How Diagnostic Expert Systems Work: Differential Diagnosis

Like the physician, the computer program generates a differential diagnosis for the sick individual. A differential diagnosis is a list of diseases that are most likely to account for the findings in a patient. The diagnostic expert system typically creates its differential diagnosis by computing the posterior probability for every disease given the findings. A disease’s posterior probability is the chance that the patient has the disease, given the findings.

Diagnostic expert systems use Bayes rules to compute the differential diagnosis. The relevance of Bayes rules to medical diagnosis was first introduced theoretically by Ledley and Lusted (1959) and first used in a diagnostic expert system by Homer Warner in 1962. Developers of diagnostic expert systems continue to use the same methods as did Homer Warner, as well as more complex Bayesian methods (but the original technique generally works well).

3.4. How Diagnostic Expert Systems Work: Question Generation

A differential diagnosis must be “resolved”; that is, the diseases in the list must be ruled in or ruled out. Like an expert

TABLE 13.2 Conditional Probabilities for FMD and MCD

Finding	Disease	$p(\text{Finding} \text{Disease})^a$
Drooling of saliva present	FMD present	0.95 (sensitivity)
Drooling of saliva present	FMD absent	0.05 (1 – specificity)
Drooling of saliva present	MCD present	0.001 (sensitivity)
Drooling of saliva present	MCD absent	0.05 (1 – specificity)
More than one animal affected	FMD present	0.95 (sensitivity)
More than one animal affected	FMD absent	0.2 (1 – specificity)
More than one animal affected	MCD present	0.001 (sensitivity)
More than one animal affected	MCD absent	0.2 (1 – specificity)

^aSimilar to the previous table, the knowledge engineers have left out half of the combinations because they can be derived from those listed listed by subtraction from 1. For example, $p(\text{drooling of saliva is not present}|\text{FMD is present}) = 1 - p(\text{drooling of saliva is present}|\text{FMD is present}) = 1 - 0.95 = 0.05$. FMD, foot and mouth disease; MCD, mad cow disease.

Probabilities from BOVID, a cattle diagnostic program. Courtesy of Animal Information Management Pty. Ltd, Victoria, Australia, BOVID, a cattle diagnostic program.

physician, the diagnostic expert system engages in a cyclic process often referred to as “hypothesize and test” to resolve the differential diagnosis. A diagnostic expert system uses value-of-information calculations (Weinstein, 1980) to recommend to the physician user additional findings that will resolve the differential diagnosis efficiently. Value-of-information calculations identify those findings, which, if known to be present or absent, optimally discriminate among the diseases in the differential diagnosis, where optimality takes into account not only the probability of a diagnosis, but cost-benefit considerations, such as whether a diagnosis is treatable.

As new findings become available (either as a result of following recommendations from the diagnostic expert system, consultants, or the physician’s own judgment), the physician user can rerun the diagnostic expert system to recompute the differential diagnosis. The most likely diagnosis from the first run may become less or more likely as the new information acts to rule in or rule out each diagnosis. The user can rerun the program whenever new findings about the individual become available during the course of the diagnostic work-up. The net result of this cyclic process is the diagnostic certainty about the diagnoses in the differential increases over time (i.e., probabilities of the diagnoses in the differential move towards zero, indicating certainty that a disease is not present, or one, indicating certainty that a disease is present) (Figure 13.1).

4. EXAMPLES OF DIAGNOSTIC EXPERT SYSTEMS: BOSS AND ILIAD

As indicated in the previous discussion, researchers have developed many diagnostic expert systems that vary in both the underlying technology as well as the domain of application. In this section, we briefly describe two systems that illustrate the key characteristics of probabilistic diagnostic expert systems. These systems—Iliad and BOSS—are diagnostic

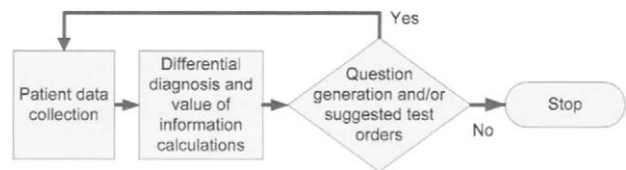


FIGURE 13.1 The process of diagnosis. Like a physician, a diagnostic expert system analyzes patient data to generate a differential diagnosis. Based on the differential diagnosis and value of information calculations, the diagnostic expert system can suggest additional questions to ask the patient, additional physical observations to make about the patient, or additional tests to consider. A physician may accept the recommendations or not. As additional information about the patient’s illness arrives over time (either as a result of tests or questions suggested by the computer; or by the physician; or simply the passage of time, this user can enter this new information and rerun the computer program to generate an updated differential diagnosis). This process of data collection and analysis can be repeated frequently until the diagnosis for a patient is established with sufficient certainty and diagnostic precision that further diagnostic evaluation is not necessary.

expert systems in the domains of diseases of humans and cattle, respectively.

4.1. Iliad

Iliad was developed by Homer R. Warner, Jr. and colleagues at the University of Utah in the 1980s (Warner, 1989, Sorenson et al., 1988, Cundick et al., 1989, Warner and Bouhaddou, 1994, Warner, 1990). Iliad is a stand-alone diagnostic expert system; that is, it requires a user (physician) to enter findings into the program. Iliad uses Bayes rules to compute a differential diagnosis for the patient's illness. It uses value-of-information calculations to recommend additional tests and observations.

Iliad performs differential diagnosis of human diseases in a variety of fields: internal medicine, sports medicine, pediatrics, dermatology, psychiatry, obstetrics/gynecology, urology, peripheral vascular diseases, and sleep disorders. Iliad covers more than 900 diseases and 1500 syndromes (which means that there are tables of information about diseases and findings for approximately 2400 diseases similar to those shown in Figures 13.1 and 13.2). Iliad computes a differential diagnosis for a patient and then displays the diseases and syndromes to the user in order of posterior probability as shown in Figure 13.2. If the user asks Iliad to suggest additional tests or questions to ask the patient, the program performs value-of-information calculations to suggest findings that would discriminate among the diseases in the differential diagnosis (Figure 13.3).

Researchers have compared Iliad's diagnostic performance on real cases with that of expert physicians and found equivalent performance (Graber and VanScoy, 2003, Friedman et al.,

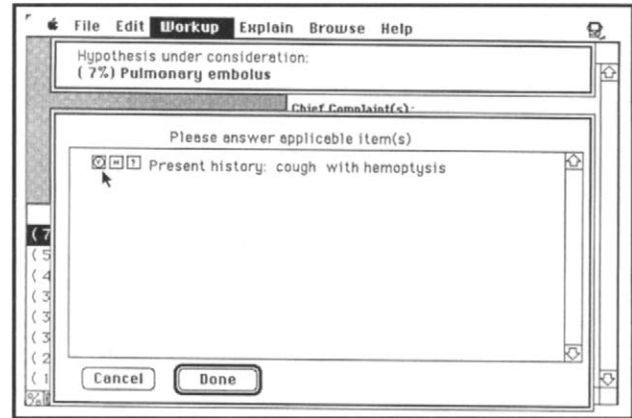


FIGURE 13.3 Iliad screen requesting additional patient information. (From Warner, H. R., Jr. (1989). Iliad: moving medical decision-making into new frontiers. *Methods Inform Med* 28:370-2, with permission.)

1999, Elstein et al., 1996, Murphy et al., 1996). Interestingly, Iliad did not achieve widespread acceptance in clinical practice because the entering of patient data was too time-consuming for busy clinicians and the program is no longer in use.

4.2. BOSSS

The Bovine Syndromic Surveillance System (BOSSS) is a web-based disease-reporting tool that incorporates a diagnostic expert system for diagnosis of diseases of cattle (Figure 13.4) (Shephard et al., 2005). BOSSS requires a user (veterinarian or cattle herd worker) to enter findings into the program (Figure 13.5). BOSSS uses Bayes rules to compute a

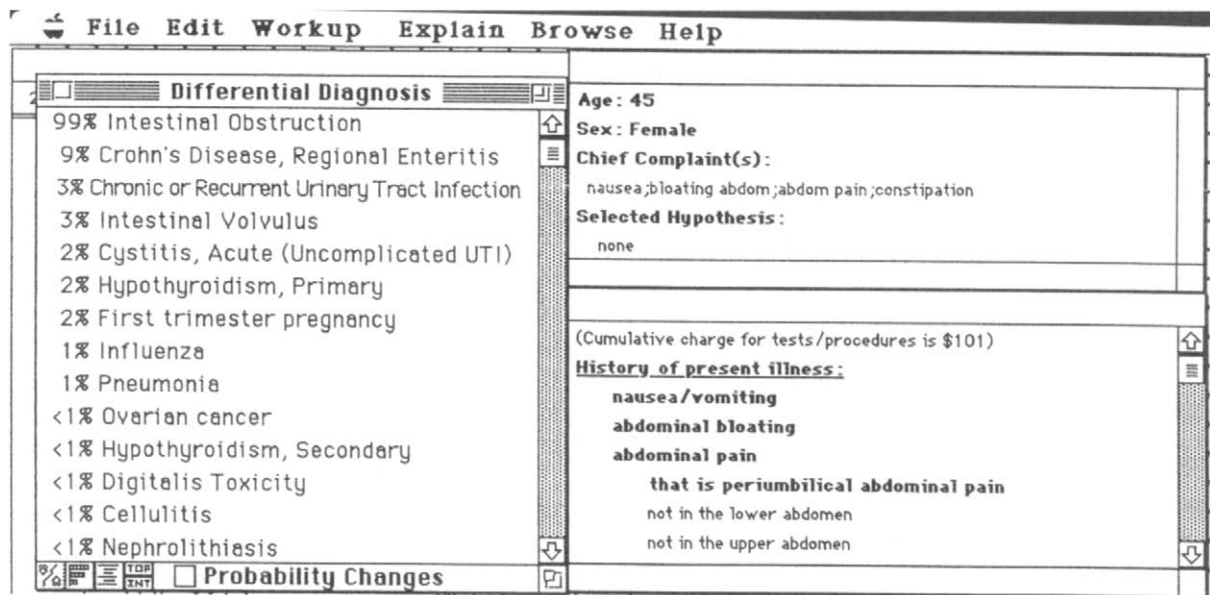


FIGURE 13.2 Consultation mode of Iliad showing the patient findings that the physician entered (lower right panel), and the differential diagnosis (panel on left) that Iliad generated from those findings. (Courtesy of LE Widman, <http://www.informaticamedica.org.br/informaticamedica/n0105/widman.htm>.)

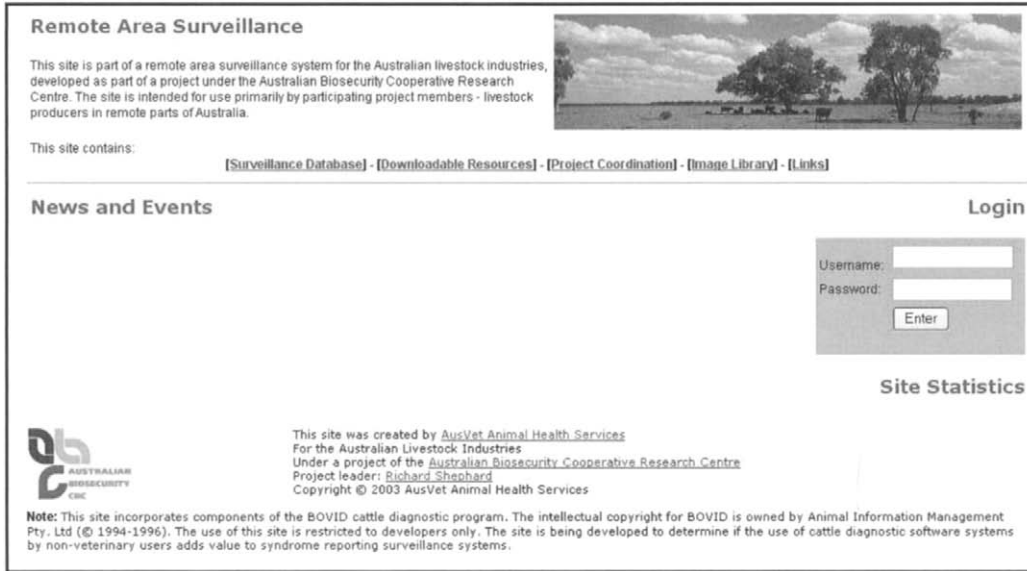


FIGURE 13.4 Log-in page of the web-based bovine syndromic surveillance system.

differential diagnosis for the patient’s illness. It also uses value-of-information calculations to recommend additional tests and observations.

Shephard and colleagues developed BOSSS to meet a need to capture disease and syndrome observation data from field observations made by veterinarians, producers, and lay observers. These observations are important to population-based surveillance of cattle herds and to the business of cattle production, but are largely unrecorded and, therefore,

unavailable for biosurveillance purposes. BOSSS encourages use by rewarding a user with diagnostic support in his efforts to determine the cause of illness or death in cattle.

Like Iliad, BOSSS contains information on the prevalence of diseases (for approximately 1000 diseases of cattle). It also contains the sensitivity and specificity of each finding for each disease. (We discuss sensitivity and specificity in Sect. 5.1.) The development team compiled this information from veterinary literature and the opinions of veterinary experts. Like Iliad, it

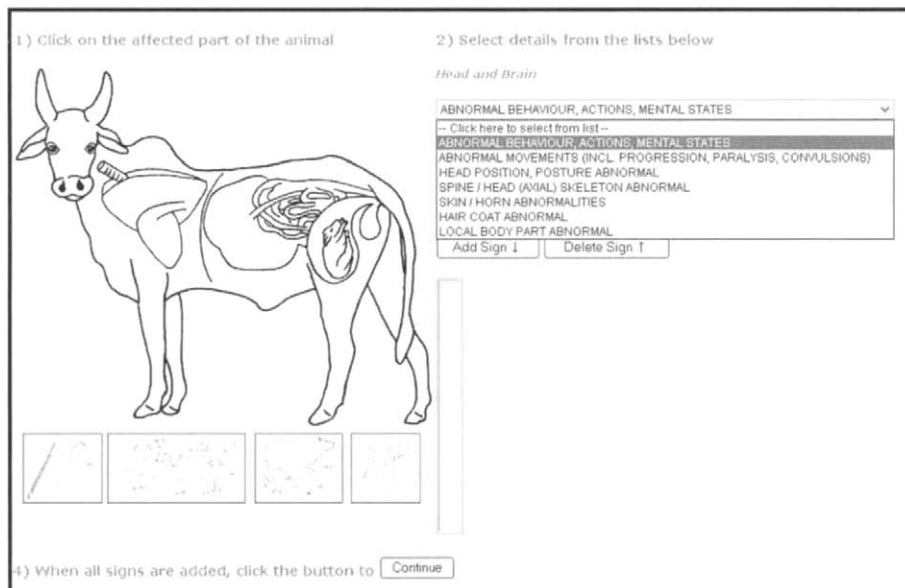


FIGURE 13.5 A graphic user interface for entering findings of a sick animal.

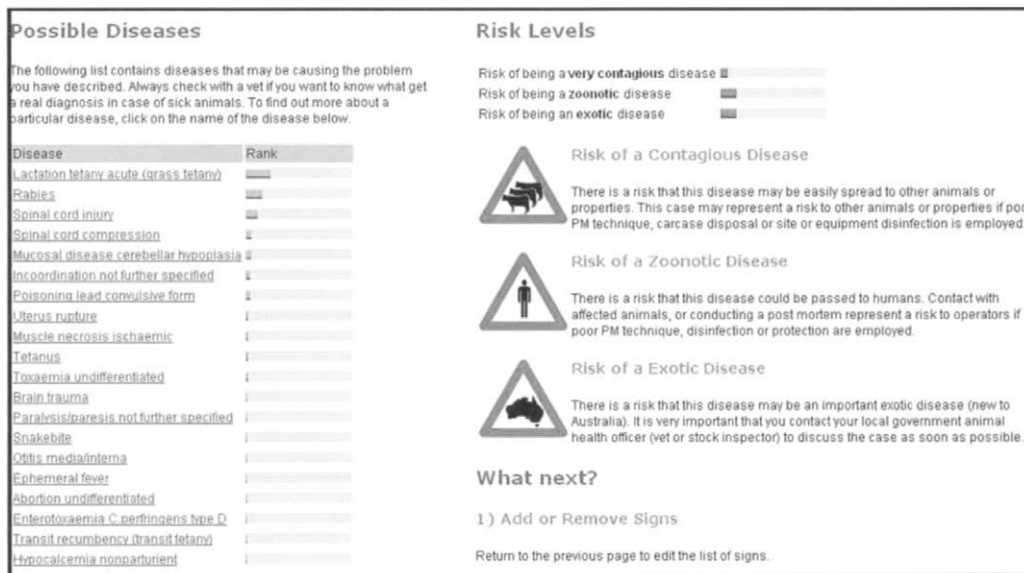


FIGURE 13.6 Differential diagnosis results ranked by posterior probabilities.

uses Bayes rules to determine the posterior probability of each disease given the findings (Figure 13.6). The program promotes the capture of extra information on each case through the use of an interrogation module that presents questions to the user (Figure 13.7). These questions are key differentiating signs for the most likely diseases identified by the system that have not already been recorded by the user. Unlike Iliad, BOSSS has an explicit biosurveillance mission. The “syndromic-surveillance” component of BOSSS is illustrated by the mapping of cases shown in Figure 13.8.

The developers of BOSSS, learning lessons from Iliad and other diagnostic expert systems about the importance of fitting

into the workflow of busy clinicians, are developing a palm-computer version of BOSSS. This portable version will fit into the practice patterns of cattle veterinarians, who spend their days at different locations and most definitely do not work in an office setting. A training manual is available for download at http://www.ausvet.com.au/bosss/resources/BOSSS_Manual.pdf.

5. KNOWLEDGE REPRESENTATION AND INFERENCE IN DIAGNOSTIC EXPERT SYSTEMS

As previously discussed, a diagnostic expert system contains an internal store of facts about diseases, including (1) the prevalence of each disease, (2) their findings, and (3) the statistical relationships between the findings and the diseases. We call this collection the knowledge base. In addition to the knowledge base, a diagnostic expert system contains an inference engine, which performs diagnostic reasoning (Figure 13.9).

To demonstrate how a diagnostic expert system uses Bayes rules to compute a differential diagnosis, we created mini-BOSSS, a tiny version of BOSSS. In particular, mini-BOSSS has a knowledge base with only two diseases and two findings. The diseases are foot and mouth disease (FMD) and mad cow disease (MCD). The findings are drooling saliva and whether there is more than one cow in the herd with this symptom.

5.1. Probabilistic Knowledge Bases

We note that there are several kinds of expert systems, including the probabilistic diagnostic expert systems that we have been discussing, and rule-based expert systems. We discuss rule-based expert systems later in this chapter.

A probabilistic knowledge base uses probabilities to represent the prevalence of disease and the relationships between findings and diseases. A development team creates a knowledge

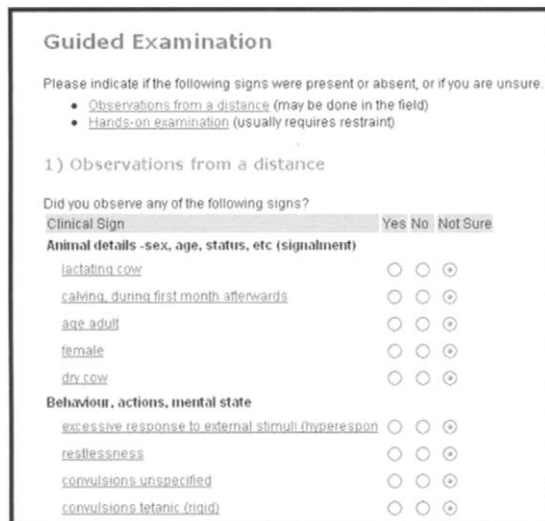


FIGURE 13.7 A graphic user interface for guided examination.

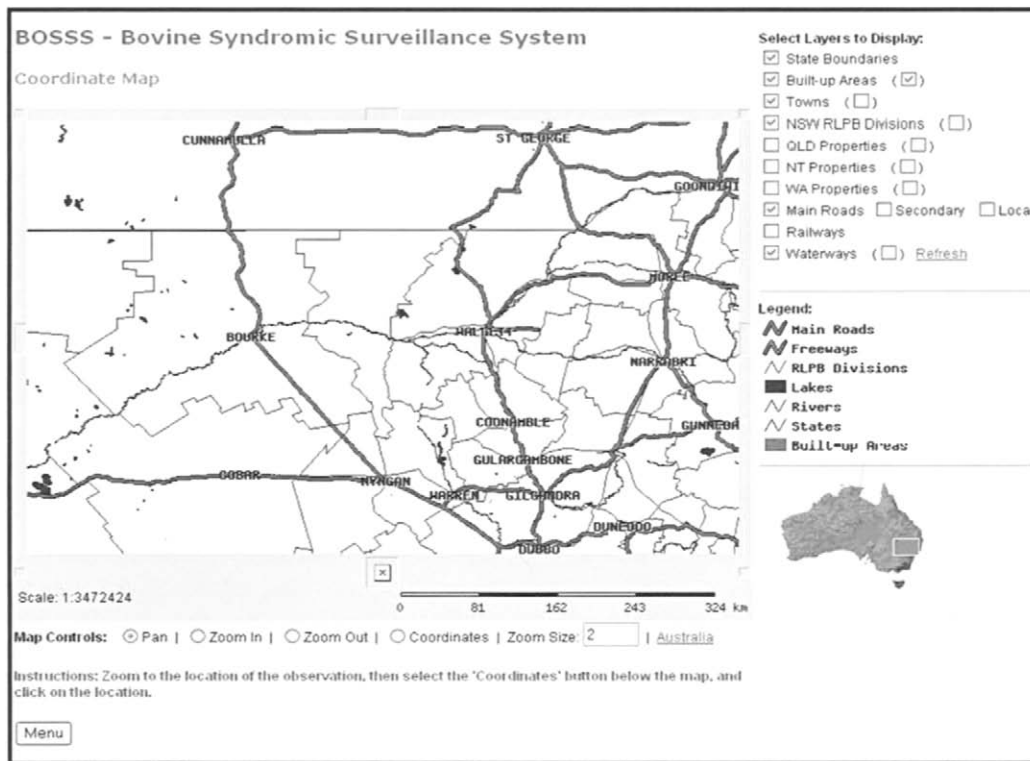


FIGURE 13.8 A mapping tool that displays geographic locations of reported cases.

base by a labor-intensive literature review supplemented by interviews with experts. Increasingly, developers use large data sets collected by hospital information systems (especially for disease prevalence) to develop knowledge bases. The process that developers use to elicit knowledge from experts and to

convert information in the literature into a knowledge base is referred to as knowledge acquisition or knowledge engineering (Feigenbaum, 1977).

5.1.1. Prior Probabilities (Disease Prevalence)

The prior probability is the prevalence of disease in the population. We represent prior probabilities using the notation $P(\text{Disease})$. For example, $P(\text{FMD})$ represents the prior probability of foot and mouth disease. Table 13.1 shows the prior probabilities of FMD and MCD that we use in our example. Table 13.1 also shows the prior odds of the diseases in mini-BOSS. Odds are a simple mathematical transform of probabilities, which are essentially equal to probabilities when probabilities are less than 0.1. The reason that we show odds in Table 13.1 is that we use the odds-likelihood form of Bayes rules in our example. We will define odds and the odds-likelihood form of Bayes shortly.

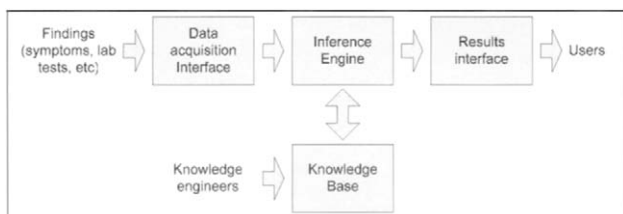


FIGURE 13.9 Components and process flow in an expert system. The inference engine uses both patient data and medical knowledge to compute a differential diagnosis and to generate suggestions for additional collection of data. The inference engine obtains data from an inbound data-acquisition interface and it outputs the differential diagnosis and suggestions for additional data through an outbound results interface. In a freestanding diagnostic expert system such as Iliad or BOSS, both the data-acquisition and results interfaces are screens that a physician interacts with. In an embedded system like Antibiotic Assistant, the data-acquisition interface is with hospital information systems (and to some extent with the user, but only for selected items of information that the user wishes to provide). The results interface could also be with another computer system such as a point-of-care system, which would present the differential diagnosis and suggestions through its own screens.

5.1.2. Conditional Probabilities

A conditional probability is the chance of one event occurring given the occurrence of another event. In diagnostic expert systems, we use conditional probabilities to describe the probability of seeing a certain finding (one event) when a disease is present (another event). If the conditional probability is high, it means that the particular finding is often associated with the disease. The mathematical notation for the conditional

probability of a finding, given a disease, is $P(\text{Finding}|\text{Disease})$. For example, $P(\text{drooling of saliva is present}|\text{FMD is present})$ is the probability of observing drooling of saliva in a cow with foot and mouth disease. Table 13.2 lists the conditional probabilities for the findings and diseases in mini-BOSSS.

If you are familiar with the concepts of sensitivity and specificity (discussed in detail in Chapter 20), you may recognize that the conditional probability $p(\text{drooling of saliva is present}|\text{FMD is present})$ is the sensitivity of the symptom *drooling of saliva* for the disease *FMD*. Similarly, $p(\text{drooling of saliva is absent}|\text{FMD is absent})$ is the specificity of drooling of saliva for FMD. Many readers will be quite familiar and comfortable with the concept of sensitivity and specificity of laboratory tests, but perhaps not with the idea of sensitivity and specificity of other findings. In fact, symptoms, travel history, results of physical examination, and results of laboratory tests are all nothing more than observations that we make about an individual that may help us discriminate between individuals with disease and those without disease. It is possible to measure a sensitivity, specificity (or, alternatively, likelihood ratios) for any of these diagnostic observations. Consider the diseases presented in mini-BOSSS. Cattle affected with FMD develop very severe mouth vesicles and ulcers. This mouth pain results in excessive salivation; therefore, it causes drooling of saliva. Cattle affected with MCD do not develop mouth wounds; therefore, they are not likely to be observed drooling saliva. The conditional probabilities in Table 13.2 reflect the observations above; $P(\text{drooling of saliva}|\text{FMD})$ is equal to 0.95 and $p(\text{drooling of saliva}|\text{MCD})$ is equal to 0.001.

If our knowledge base included laboratory tests for FMD and MCD, these conditional probabilities would be the sensitivities and specificities of the laboratory tests for FMD and MCD.

Researchers working in the field of diagnostic expert systems sometimes refer to conditional probabilities of findings for given diseases as textbook knowledge because these probabilities are often available in textbooks of medicine (human or veterinary). For example, a chapter in a textbook of veterinary medicine on FMD will contain many statements about the frequency with which different findings occur in animals with FMD. These frequencies are basic facts, highly relevant to diagnosis.

5.2. Probabilistic Inference Engines

A probabilistic inference engine is an algorithm that computes a differential diagnosis for a sick individual. In particular, it computes the posterior probability of each disease in its knowledge base, given the findings for a sick individual. In our example, the algorithm will compute the posterior probability $p(\text{FMD is present}|\text{drooling of saliva is present, more than one animal is affected})$ and the posterior probability $p(\text{MCD is present}|\text{drooling of saliva is present, more than one animal is affected})$.

A probabilistic inference engine uses Bayes rules to compute the posterior probability from the prior probability (or prior odds) and the sensitivities and specificities of the observed findings. Once it has computed the posterior probability of every disease in its knowledge base, the inference engine outputs a differential diagnosis, that is, a list of all diseases in the knowledge base sorted from most probable to least probable.

Bayes rules are sometimes referred to as Bayesian inversion because it inverts the conditional probability that a given finding will be observed in an individual with disease (textbook knowledge) into the probability that an individual has the disease, given that we observe a finding in that individual (diagnostic knowledge). That is, it inverts $p(\text{drooling of saliva is present}|\text{FMD is present})$ to $p(\text{FMD is present}|\text{drooling of saliva is present})$, which is exactly what a veterinarian needs to know. A veterinarian needs to know the likelihood that the cow has FMD or MCD, given the findings.

A complete discussion of the various algorithms that a probabilistic inference engine can use to compute posterior probabilities using Bayes rules would fill a book, such as the excellent textbook by Richard Neapolitan (2003). For teaching purposes, we will use a simple form of Bayes rules called the *odds-likelihood form of Bayes rules*. This simplified form rests on an assumption that findings are independent, given the disease, which is why researchers called this formulation (and similar formulations) *naive Bayes*. This simple form of Bayes rules works surprisingly well in many diagnostic expert systems, including the systems developed by Homer Warner (congenital heart advisor) and Homer Warner Jr. (Iliad), as well as the BOSSS. We will use the odds-likelihood version of Bayes rules to illustrate how BOSSS computes a differential diagnosis for two diseases, given two findings about a sick cow.

5.2.1. Definition of Odds

For clarity and convenient reference, we here provide the definition of *odds*:

$$\text{odds} = \frac{p}{1-p} \quad (1)$$

Odds are simply a rescaling of probability from a range of 0 to 1 to a range of 0 to infinity (which you can prove to yourself by substituting probabilities of zero and one into Eq. 1). For probabilities less than 0.1, probabilities and odds are roughly equal. A probability of 0.1, for example, equals an odds of $\frac{0.1}{1-0.1} = 0.11$. You can safely think “probability”

whenever we use the term “odds” and vice versa, as we will be dealing with small probabilities. Upon reading Appendix D, which contains the simplest probabilistic formulation of Bayes rules, you will see why we use the odds-likelihood form in this chapter—it is simpler to learn Bayes rules using this form, and more illuminating.

5.2.2. Odds-Likelihood Form of Bayes Rule

Equation 2 is the odds-likelihood form of Bayes rules when we have only one finding.

$$\text{Odds}(D|f) = LR_{f|D} \times \text{Odds}(D) \quad (2)$$

This equation says that if we know the prior odds $\text{Odds}(D)$ (read prevalence) of a disease D and we observe one finding f of that disease in an individual, we can compute the posterior odds (read probability) by multiplying the prior odds times the likelihood ratio of that finding for that disease.

Likelihood ratios are nothing more than an alternative way of expressing the sensitivity and specificity of a test or observation for a disease. In fact, the likelihood ratio is defined in terms of sensitivities and specificities. The likelihood ratio positive (positive means that the finding *drooling saliva* is known to be present) of drooling saliva for the disease FMD follows:

$$LR_{\text{drooling saliva|FMD}}^+ = \frac{p(\text{drooling saliva present|FMD present})}{p(\text{drooling saliva present|FMD absent})} \quad (3)$$

In Equation 3, the numerator is sensitivity, and the denominator is the false positive rate (equal to $1 - \text{specificity}$).

Note that the beauty of the likelihood ratio $LR_{\text{drooling saliva|FMD}}^+$ is that it is a very direct measure of how well drooling saliva discriminates between animals with FMD and animals without FMD. If drooling of saliva occurs much more frequently in animals with FMD than animals without FMD, $LR_{\text{drooling saliva|FMD}}^+$ will be a large number. For a finding that is pathognomic (meaning that the finding in itself is sufficient to diagnose a disease), the denominator will be zero and the LR^+ will be infinity, meaning that no matter how small the prior odds are, the posterior odds will be infinity (which converts to a probability of 1.0, and means that the individual has the disease with certainty). If, on the other hand, drooling of saliva occurs with equal frequency in animals with FMD and animals without FMD, the numerator and the denominator will be equal and the $LR_{\text{drooling saliva|FMD}}^+$ will be equal to one, which when multiplied times the prior odds of FMD will result in a posterior odds that is equal to the prior odds. This result makes sense. If a finding cannot discriminate between animals with FMD and without FMD, it contains no diagnostic information for the disease FMD, and should not increase or decrease our belief that the animal has FMD.

The likelihood ratio negative (negative means that the finding *drooling saliva* is known to be absent) follows:

$$LR_{\text{drooling saliva|FMD}}^- = \frac{p(\text{drooling saliva absent|FMD present})}{p(\text{drooling saliva absent|FMD absent})} \quad (4)$$

In Eq. 4, the numerator is the false negative rate (equal to $1 - \text{sensitivity}$), and the denominator is specificity. In medicine,

the absence of a finding that we expect to see if the patient actually has the disease in question is useful information. There is even a term for it—*significant negative finding*. A significant negative finding, such as a negative laboratory test, helps to rule out a diagnosis. A likelihood ratio negative is always a number that is less than one (but greater or equal to zero) for findings that we expect to see more often in individuals with the disease than in individuals without the disease.

Equation 2 expresses the essence of Bayes rules. Diagnosticians use Bayes rules to update their prior belief in a diagnosis in light of new information. When a diagnostician has no information whatsoever about an individual, her belief that the individual has a disease should be the prevalence of the disease. If she makes an observation (whether positive or negative), she should update her belief in the diagnosis by multiplying the likelihood ratio for the test or observation for that disease times the prior odds of the disease (think “prevalence”).

If we know nothing about a cow whatsoever, then our belief that the cow has FMD is simply the prevalence of FMD. If we subsequently observe that the cow is drooling saliva, Bayes rules instructs us to update our belief that the cow has FMD using the information in Table 13.1 and 13.2 using the following calculation:

$$\begin{aligned} \text{Odds}(\text{FMD} | \text{drooling saliva}) &= LR_{\text{drooling saliva|FMD}}^+ \times \text{Odds}(\text{FMD}) \\ &= \frac{p(\text{drooling saliva|FMD present})}{p(\text{drooling saliva|FMD absent})} \times \frac{0.001}{1-0.001} \\ &= \frac{0.95}{0.05} \times \frac{0.001}{0.999} \\ &= 19 \times 0.001 \\ &= 0.019 \end{aligned} \quad (5)$$

(If we observe that the cow is not drooling saliva, we would use the likelihood ratio negative in Bayes rules.)

If we subsequently observe that a second cow is sick, we can apply Bayes rules a second time. In effect, we are treating the posterior odds from the prior calculation as the new prior odds:

$$\begin{aligned} \text{Odds}(\text{FMD} | \text{drooling saliva, more than one animal affected}) &= LR_{\text{more than one animal affected|FMD}}^+ \times \text{Odds}(\text{FMD} | \text{drooling saliva}) \\ &= \frac{p(\text{more than one animal affected|FMD present})}{p(\text{more than one animal affected|FMD absent})} \times 0.0019 \\ &= \frac{0.95}{0.20} \times 0.019 \\ &= 4.75 \times 0.019 \\ &= 0.09 \end{aligned} \quad (6)$$

In general, if we have N diagnostic facts about a cow (or person), the odds-likelihood form of Bayes rules has the following form:

$$\text{Odds}(D|f_1, f_2, \dots, f_n) = \text{LR}_{f_1|D} \times \text{LR}_{f_2|D} \times \dots \times \text{LR}_{f_n|D} \times \text{Odds}(D) \quad (7)$$

The result of Eq. 6, Odds(FMD|drooling saliva, more than one animal affected) + 0.09, is the odds (think probability) that the cow has FMD given that we have observed drooling of saliva in this cow and at least one other cow. The probabilistic inference engine would repeat the same calculation for the disease MCD, using the prior odds of MCD and the likelihood ratios for the two findings for MCD. The result of this calculation (not shown) is 9.8×10^{-8} . Cows with MCD stagger, but rarely drool saliva.

5.2.3. Differential Diagnosis

Table 13.3 shows the differential diagnosis that our extremely simple diagnostic expert system would show to a user. The two diseases are sorted in order of posterior odds. Note that we converted the posterior odds back to posterior probabilities using the formula $p = \frac{\text{odds}}{1 + \text{odds}}$, which we obtained by solving Eq. 2 for p.

Note that the posterior odds (and probabilities) in this example are very low. The low posterior odds are also partly due to the low prior prevalence of FMD and MCD that we arbitrarily assign (note that zero prior probability is not acceptable for the likelihood ratio computation). Countries that are currently free from these diseases or within the final stage of a disease eradication program for the diseases use low prior prevalence for exotic diseases (e.g., FMD and MCD). The LR+s for these findings for the disease FMD are only 19 and 4.75, respectively. If we had a third finding, such as blisters on feet, the LR+ for this finding would be a third multiplier in the equation, possibly increasing the posterior odds for one or both diagnoses. If we had a positive result from a highly sensitive and specific test for FMD (very high LR+), the posterior odds for FMD might be quite high. The key question for bio-surveillance systems is at what level of certainty of a diagnosis in a single individual or a set of individuals different response actions are warranted. The answer to this question depends on treatability and other cost-benefit considerations that we discuss in Part V of this book.

TABLE 13.3 Differential Diagnosis for Cow Drooling Saliva in a Herd with At Least One Other Cow Drooling Saliva

	Posterior Odds	Posterior Probability
FMD	0.09	0.083
MCD	9.8×10^{-8}	$\sim 9.8 \times 10^{-8}$

FMD, foot and mouth disease; MCD, mad cow disease.

5.3. Computing Posterior Probabilities Using Bayesian Networks

Readers should be aware that there are other forms of Bayes rules that do not rest on an assumption of conditional independence, given a diagnosis. A developer of a probabilistic expert system may choose to use these alternative forms to improve the diagnostic accuracy of a system. Although mathematically too complex to cover in a brief introductory tutorial to Bayesian inference, the differences among these more complex forms of Bayes rules are simple to explain using a graphical representation called a Bayesian network (Figure 13.10). A Bayesian network comprises a set of nodes and arcs where each node represents a conditional probability distribution for the variable that the node represents, conditioned on its parents (the nodes from which directed arcs connect to the node). For the benefit of statisticians, a Bayesian network is a (compact) factorization of the complete joint probability distribution over all variables represented by the nodes (Neapolitan, 2003). The arcs in a Bayesian network represent the statistical dependence and independence relationships among the variables in the model. Figure 13.10 is a Bayesian network diagram for our mini-BOSSS diagnostic expert system.

The fact that there is no arc between *drooling saliva* and *more than one animal affected* indicates that the probability of observing drooling of saliva in a cow, once we know whether the cow has FMD or MCD, is not affected by knowing that other animals have these symptoms and vice versa (knowing that other animals have drooling once we know that this cow has FMD does not change the probability that this animal is drooling). If *drooling saliva* and *more than one animal affected* were not independent, given that we know that the cow has FMD, then we could add an arc between them.

Figure 13.11 shows a portion of the Bayesian network that underlies the Pathfinder system, which is a diagnostic expert system for pathologists who are interpreting biopsies of lymph

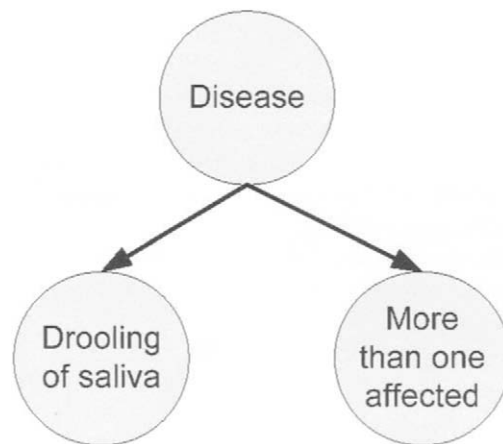


FIGURE 13.10 A Bayesian network of mini-BOSSS.

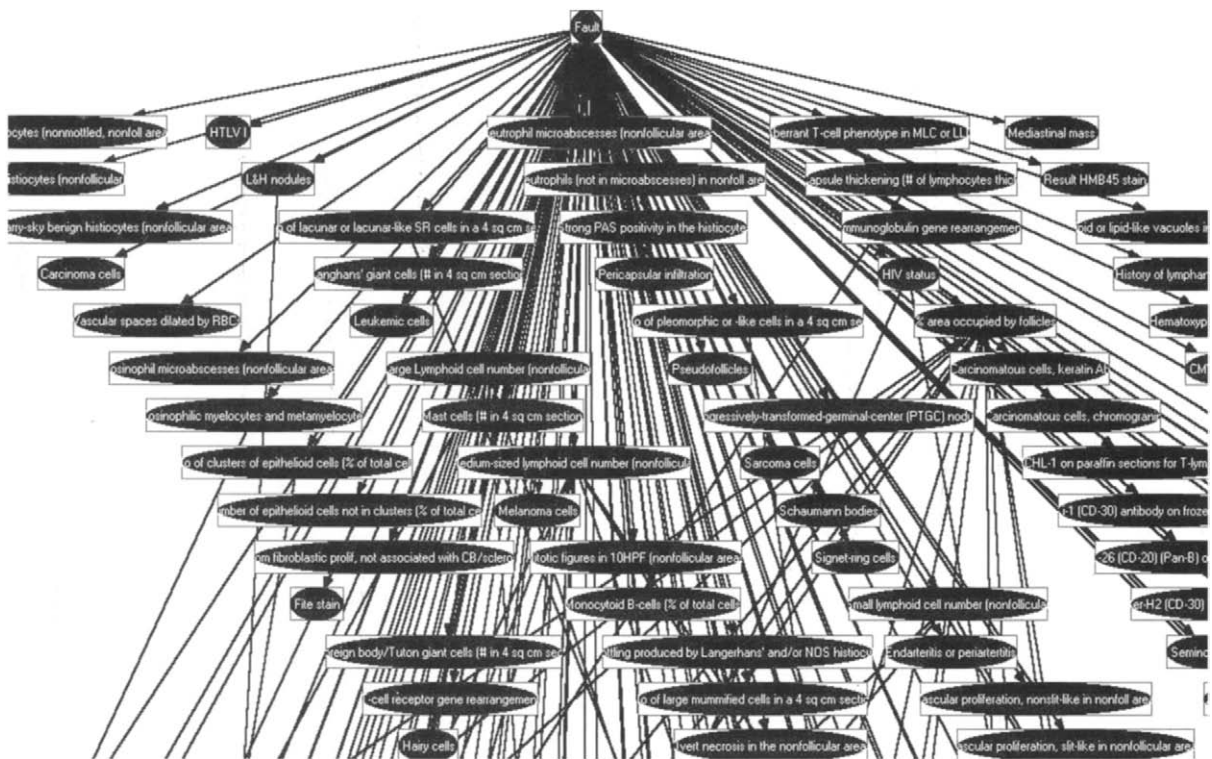


FIGURE 13.11 A portion of the Bayesian network in the Pathfinder expert system for pathology. (From Stuart Turner, lecture notes, 2001, with permission.)

nodes (Heckerman et al., 1992). The arcs between findings indicate that they are not independent, given diagnoses. Note that the Pathfinder network is a more realistic illustration of the size and complexity of the models used in diagnostic expert systems. In chapter 18, we discuss Bayesian networks in the context of Bayesian biosurveillance. That chapter discusses a really big model.

6. RULE-BASED EXPERT SYSTEMS

Rule-based expert systems may also find application in biosurveillance. A rule-based expert system (also known as a deterministic expert system) encodes medical knowledge in a set of *if-then* rules. Rule-based expert systems are simpler, conceptually, than probabilistic systems, but they are only appropriate for problems that do not involve reasoning under uncertainty. A rule-based formalism is appropriate when diagnostic or other knowledge can be represented as Boolean statements (e.g., if A and B then C). A hospital or biosurveillance organization can use a rule-based expert system to automate case detection when the findings are diagnostically precise (e.g., when case-detection is based on diagnostically precise data such as a positive tuberculosis culture [Miller et al., 1997]).

Figure 13.12 is an anthrax case detection rule from the Clinical Event Monitor, an embedded rule-based expert system developed by several of the authors (Wagner et al., 1997).

This rule is in a format that can be interpreted by a rule-based inference engine called CLIPS (National Aeronautics and Space Administration). The rule contains variables that represent findings of the disease anthrax. The part of the rule before the \geq symbol is the IF part and the part after the symbol is the THEN part. If all the conditions defined in the IF segment are true—wide mediastinum is found in a chest radiograph report and gram positive rods are reported in laboratory results within 10 days of each other—then the rule will conclude that there is a possible anthrax case. A typical rule-based expert system contains hundreds of such rules.

7. EMBEDDED EXPERT SYSTEMS

In medical applications, the requirement for manual input of data by busy clinicians is a significant barrier to the use of diagnostic expert systems (Graber and VanScoy, 2003). In the 1970s, several research groups took this barrier seriously. They began to explore methods to embed diagnostic expert systems—providing diagnostic support and, more generally, decision support—to clinicians without asking them to enter patient findings. Their research demonstrated that it is possible to obtain needed findings directly from existing clinical information systems (Evans, 1991, Gardner et al., 1999, Rind et al., 1993, Rind et al., 1995, McDonald, 1976c, McDonald et al., 1992a, McDonald, 1976b, McDonald et al., 1992b, Overhage et al., 1996,

```

(defrule detect_anthrax
  (event Radiology_report)
  (report (patient_id ?patient_id))
  (freetext ?freetext)
  (test (str-index "wide mediastinum" (lowercase ?freetext)))
  =>
  (if (> (cCheckGrams ?patient_id "GPR" 10) -1) then
    (event ANTHRAX)
    (printout t "Found gram positive rods and widened mediastinum in chest x-ray"
     crlf)
    (printout t "==== Detected a possible ANTHRAX case =====" crlf)
  )
)

```

FIGURE 13.12 An anthrax-case detection rule in the Clinical Event Monitor, a rule-based expert system. The rule is written in CLIPS, a language and expert system shell developed by the National Aeronautics and Space Administration.

Wagner et al., 1997) and from alterations in work flow, such as replacing unstructured paper records with computer-generated encounter forms (from which findings can be optically scanned or manually extracted) (McDonald, 1976a, McDonald et al., 1992a, McDonald, 1976b, McDonald et al., 1992b). The research also demonstrated how diagnostic expert systems can be embedded in clinical information systems, such as physician order entry systems (Tierney et al., 1993, Dexter et al., 2001) and electronic medical records (Gardner et al., 1999, Warner et al., 1997). Several of these systems functioned in the domain of hospital infection control (Hripcsak et al., 1997, Hripcsak et al., 1999, Kahn et al., 1996a, Kahn et al., 1996b, Kahn et al., 1993).

The secret to successful deployment of a diagnostic or other expert system in medicine is designing a system in which the benefits to clinicians (from decision support and increases in efficiency) outweigh any additional effort required by clinicians to enter data by a considerable margin.

8. DIAGNOSTIC EXPERT SYSTEMS FOR BIOSURVEILLANCE

In addition to the diagnostic expert systems for hospital infection control just outlined, there are two other projects developing systems in the domain of biosurveillance. Shannon and colleagues at the Children's Hospital in Boston are developing a web-based diagnostic expert system to assist emergency room clinicians with diagnosis of approximately 20 diseases caused by biological agents known to have been weaponized or that are otherwise of concern as potential bioterrorist threats. The system has developed and evaluated on-line educational modules about the biological agents (Chung et al., 2004), and is intended to support the reporting of cases to local, state, and federal agencies (Shannon et al., 2002). Similar to stand-alone

expert systems, the web-based tool requires clinicians to manually enter patient information.

The National Library of Medicine has developed a system called WISER (Wireless Information System for Emergency Responders) to assist first responders when they arrive at a hazardous material (Hazmat) incident, such as a chemical spill (<http://wiser.nlm.nih.gov/>). WISER operates on a personal digital assistant (PDA) and can send and receive data with a central location and other PDAs running the WISER program through a wireless network. WISER contains a diagnostic expert system that provides assistance in identification of an unknown substance and, once the substance is identified, provides guidance on actions to save lives and protect the environment. The WISER framework could be expanded to include a diagnostic expert system for the analysis of patient or animal findings.

9. PERFECTING CASE AND OUTBREAK DETECTION

Recall that the best case detection system imaginable would be one in which every individual in a community is examined every morning by the best diagnostician in the world. Since she would be examining everyone every day, she would be aware of patterns of illness in a community and this awareness would appropriately influence her diagnostic thinking (and treatment) of individual patients.³ This diagnostician would also never fail to report immediately each fever, early syndromic presentation, or reportable disease to governmental public health.

Diagnostic expert systems are the key to building such a system. The research we reviewed in this chapter has already solved most, if not all, of the technical problems. What is needed is the will to create such a system.

3 Physicians are taught (and reminded incessantly) "when you hear hoof beats, don't think of zebras". This adage is an informal statement that when the evidence available about a particular patient supports equally a diagnosis of either influenza or SARS (e.g., the patient has constitutional symptoms and no history of exposure to SARS), they should conclude that the diagnosis of influenza is far more likely than SARS.

If and when diagnostic expert systems are embedded in the clinical information systems of every hospital (animal and human), long-term care facility, clinic, and laboratory in a region, they will be able to notify a health department or other biosurveillance organization of every fever, syndromic presentation, and reportable disease in individuals receiving medical or veterinary care. If diagnostic expert systems are made available to the public or to selected high-risk populations (e.g., postal employees or patients with preexisting conditions such as asthma or diabetes), case finding would be extended to an even larger fraction of the population, approximating the “every-patient-every-day” capabilities of an ideal case detection system.

If a biosurveillance system located in a region’s health department were to receive the differential diagnosis for each individual in the region in real time (anonymously, of course, and perhaps selectively based on diseases and probability thresholds), it could compute the current incidences of

these conditions. It could monitor the region for increases in incidence of findings, syndromes, and diseases of concern.

Note that the outputs of a probabilistic diagnostic expert system are posterior probabilities of diseases for one patient, so the central monitoring would be monitoring of the sums of the probabilities for all reported patients. Figure 13.13 illustrates this summation. It plots the daily sums of the posterior probabilities of “flu-like” illness of all patients seen in emergency departments on each day. If the diagnostic expert systems are well calibrated, the sums of the posterior probabilities should equal the actual number of patients with the disease in the population being evaluated by the system.

Finally, if the biosurveillance system would then communicate the current fever, syndrome, and disease incidences back to the diagnostic expert systems being used by clinicians and citizens, we would realize the “diagnosticians-are-aware-of-patterns-of-illness-in-the-community” capability of an ideal case detection system.

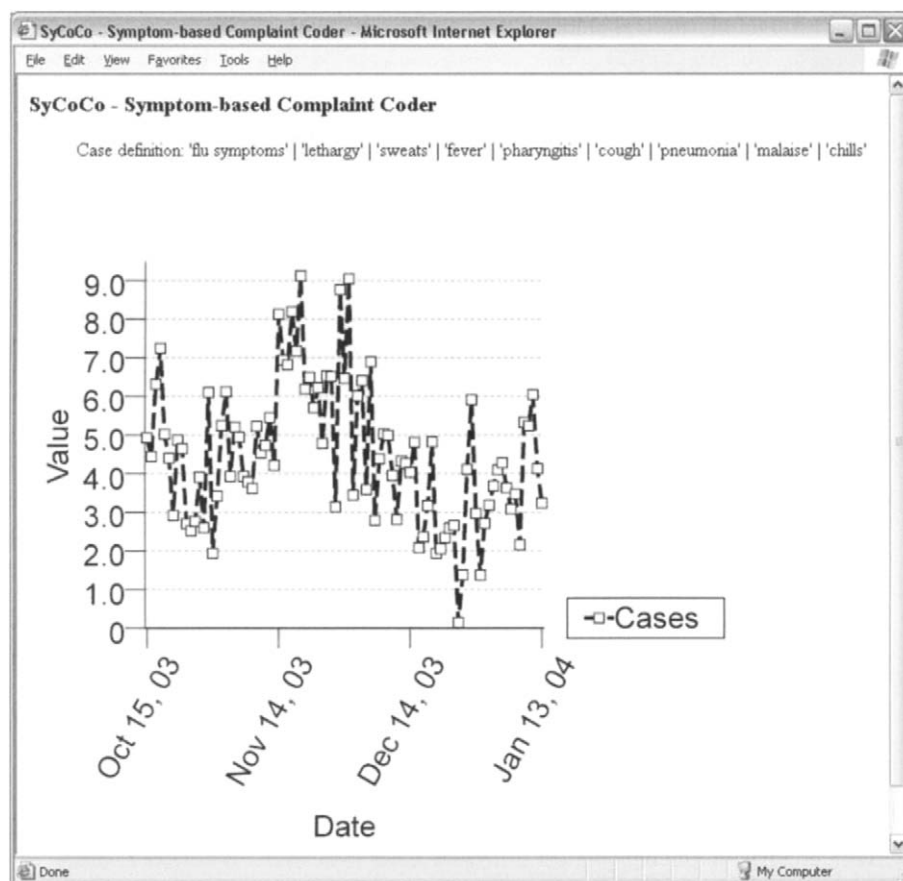


FIGURE 13.13 Daily sum of syndrome probabilities produced by SyCO². SyCO² computes the posterior probability that a patient has a flu-like illness from his chief complaint. (From Espino, J., Dara, J., Dowling J, et al. (2005). SyCO²: A Bayesian Machine Learning Method for Extracting Symptoms from Chief Complaints And Combining Them Using Probabilistic Case Definitions. Pittsburgh, PA: Center for Biomedical Informatics, University of Pittsburgh, with permission.) An outbreak-detection system would sum the posterior probabilities of flu-like illness from all patients seen in 24-hour periods to form a time series of expected daily counts of patients with respiratory illness. Readers familiar with Bayesian statistics will recognize this sum as the expectation for the number of individuals with a given diagnosis.

This ideal approach underlies the Bayesian approach to outbreak detection described in Chapter 18 Bayesian Methods for Diagnosing Outbreaks. PANDA (Population-wide ANomaly Detection and Assessment)—the research system described in that chapter—actually merges many individual diagnostic Bayesian networks into a very large network that also includes a subnetwork that draws inferences about the presence or absence of an outbreak. PANDA is pursuing this idea on a citywide scale. Conceptually, a PANDA network comprises millions of person-specific diagnostic Bayesian networks, each of whose probabilities of disease (e.g., anthrax) are being influenced by population-level observations (e.g., aggregate sales of over-the-counter medications) and population-level inferences (e.g., the likelihood that other individuals who may have inhalational anthrax are present in a population). PANDA also includes a prior probability distribution over outbreak diseases (e.g., the prior probability of inhalational anthrax, based on the national terror alert level). By integrating person-specific diagnostic submodels into a population-wide super model, approaches like PANDA are able to make inferences about the probability of a disease outbreak in the population as a whole, as well as the probability of disease in individual people (or subgroups of people) within the population.

10. COMPUTER-INTERPRETABLE CASE DEFINITIONS

An ideal case detection system would also support case finding during outbreak investigations. Case definitions (described in Chapter 3) are the basis for case-control studies and investigations of emerging diseases. A computer-interpretable case definition is a prerequisite for providing computer-support to case

finding during investigations. As discussed in Chapter 3, case definitions are Boolean (logical) statement of findings (Figure 13.14).

Case definitions, as currently written, are not well suited for automation. The authors of the SARS case definition intended it for use by physicians and epidemiologists, not computers. The clause “findings of lower respiratory illness (e.g., cough, shortness of breath, difficulty breathing)” does not enumerate all findings of lower respiratory illness. A computer requires a complete enumeration of all findings that it should count as evidence of *lower respiratory illness* (e.g., cough, shortness of breath, difficulty breathing, wheezing, cyanosis, tachypnea, dullness to percussion, fremitus, whispered pectoriloquy, rales, and rhonchi). The findings would also have to be described more precisely. For example, a physician or an epidemiologist would not count chronic cough or cough associated with asthma as a finding of lower respiratory illness when applying this case definition, but a computer would (unless told otherwise). Note that it is difficult, if not impossible, to enumerate all of the possible exceptions to the counting of a finding as evidence of a disease. This difficulty is the reason that diagnostic expert systems in medicine are probabilistic. They quantify the number of exceptions to a categorical statement about the relationship between findings and disease using probabilities. For example, 70% of patients with cough have an acute lower respiratory illness, but 30% (the exceptions) have some other cause. This observation suggests that computer-interpretable case definitions will employ Bayesian networks, as illustrated by Figure 13.15.

For readers interested in the topic of knowledge representations for computer-interpretable case definitions, there is

Case definition for confirmed case of SARS-CoV disease: (early OR mild-to-moderate OR severe illness) AND laboratory confirmation

Definitions of Terms used in Case Definition:

Early illness: two or more of the following findings: fever (might be subjective), chills, rigors, myalgia, headache, diarrhea, sore throat, rhinorrhea.

Mild-to-moderate respiratory illness: Temperature of >100.4° F (>38° C) AND one or more findings of lower respiratory illness (e.g., cough, shortness of breath, difficulty breathing)

*Severe illness**

Laboratory confirmation: serum antibody to SARS-CoV by a test validated by CDC (e.g., enzyme immunoassay [EIA]), OR isolation in cell culture of SARS-CoV from a clinical specimen, OR SARS-CoV RNA by a reverse-transcription-polymerase chain reaction (RT-PCR) test validated by CDC and with subsequent confirmation in a reference laboratory (e.g., CDC).

FIGURE 13.14 Excerpt from the CDC case definition for confirmed SARS-CoV disease (<http://www.cdc.gov/ncidod/sars/guidance/b/app1.htm>). Only the definitions relevant for classifying a patient as a probable case of SARS CoV are shown. (We omitted the definition of severe respiratory illness for clarity. The illness clause in the case definition for confirmed case of SARS-CoV disease is a disjunction, and all patients that satisfy the definitional criteria for severe disease also satisfy the criteria for mild to moderate respiratory illness.

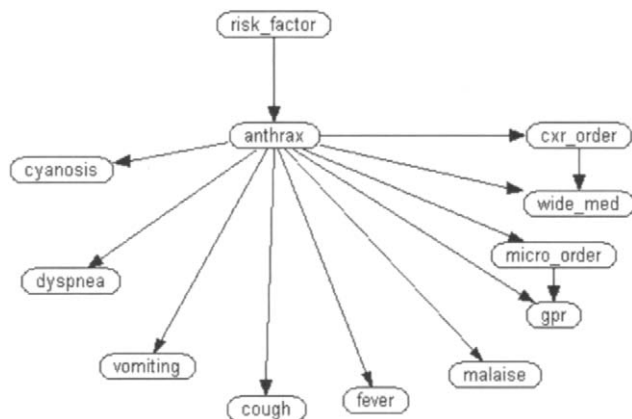


FIGURE 13.15 A Bayesian network “case definition” for anthrax. The network computes the posterior probability for a patient admitted to the emergency department. *Cxr order*, electronic record of an order for a chest radiograph; *wide med*, chest radiograph finding of wide mediastinum; *Gpr*, gram-positive rods in blood or cerebrospinal fluid (CSF) smear; *micro order*, order for a blood or CSF culture. (From Espino J, Tsui, F-C. (2000). A Bayesian network for detecting inhalational anthrax outbreaks. Pittsburgh, PA: Center for Biomedical Informatics, University of Pittsburgh, with permission.)

literature from research on computer-interpretable patient care guidelines that is relevant to this topic (Shiffman et al., 2004, Tu and Musen, 2001, Peleg et al., 2003, Boxwala et al., 2004, Wang et al., 2003, Seyfang et al., 2002, Fox et al., 1997, Terenziani et al., 2003, de Clercq et al., 2004, Johnson et al., 2001, Ciccarese et al., 2004, Parker et al., 2004).

11. SUMMARY

Despite the unfamiliarity of most readers with diagnostic expert systems, we chose to begin our discussion of algorithms for biosurveillance with this topic because case detection provides the case data needed for outbreak detection. Outbreak detection cannot function without case detection (unless the surveillance data are aggregated data, such as daily sales of over-the-counter thermometers).

Diagnostic expert systems have the potential to improve the quality and completeness of the case data available to analytic methods designed to detect and characterize outbreaks, which we discuss in the following chapters. They can profoundly improve the reporting of syndromes. Improvements in case detection will translate directly into improvements in the earliness of outbreak detection and characterization.

McDonald did not conclude in his seminal paper that the solution to the non-perfectibility of man was to admit only women to medical schools. Rather, he stated, “Thus, I conclude that though the individual physician is not perfectible, the system of care is, and that the computer will play a major part in the perfection of future care systems.” His point—that technology can be used to create a system (involving both

humans and computers) that can then be perfected—also seems to apply to biosurveillance systems.

12. ADDITIONAL RESOURCES

12.1. Expert Systems in Plant Pathology

For readers who are interested in diagnostic expert systems for diagnosing plant diseases, Travis and Latin (1991) briefly reviewed several diagnostic expert systems in plant pathology, including PLANT/ed, Apple Pest and Disease Diagnosis, CALEX/Peaches, Muskmelon disorder management system, and Penn State Apple Orchard Consultant, in “Development, Implementation, and Adoption of Expert Systems in Plant Pathology.”

13. ACKNOWLEDGMENTS

This work was supported in part by grants from the National Science Foundation (IIS-0325581), the Defense Advanced Research Projects Agency (F30602-01-2-0550), and the Pennsylvania Department of Health (ME-01-737). We thank William Hogan and Gregory Cooper for helpful discussions and comments.

14. REFERENCES

- Adams, I. D., Chan, M., Clifford, P. C., et al. (1986). Computer aided diagnosis of acute abdominal pain: a multicentre study. *BMJ (Clin Res Ed)* 293:800–4.
- American College of Emergency Physicians. (1994). Clinical policy for the initial approach to patients presenting with a chief complaint of nontraumatic acute abdominal pain. American College of Emergency Physicians. *Ann Emerg Med* 23:906–22.
- Aronsky, D., Fisman, M., Chapman, W. W., et al. (2001). Combining decision support methodologies to diagnose pneumonia. In: Proceedings of American Medical Informatics Association Annual Fall Symposium, 12–16.
- Boxwala, A. A., Peleg, M., Tu, S., et al. (2004). GLIF3: a representation format for sharable computer-interpretable clinical practice guidelines. *J Biomed Inform* 37:147–61.
- Buchanan, B. G., Shortliffe, E. H. (1984). Rule-based expert systems: the MYCIN experiments of the Stanford Heuristic Programming Project. Reading, MA: Addison-Wesley.
- Burnside, E. S., Rubin, D. L., Shachter, R. D., et al. (2004). A probabilistic expert system that provides automated mammographic-histologic correlation: initial experience. *AJR Am J Roentgenol* 182:481–8.
- Chung, S., Mandl, K. D., Shannon, M., et al. (2004). Efficacy of an educational web site for educating physicians about bioterrorism. *Acad Emerg Med* 11:143–8.
- Ciccarese, P., Caffi, E., Boiocchi, L., et al. (2004). A guideline management system. *Medinfo* 11:28–32.
- Cundick, R., et al. (1989). Iliad as a patient care simulator to teach medical problem solving. Paper presented at 13th Symposium on Computer Applications in Medical Care, Washington, DC.

- de Clercq, P. A., Blom, J. A., Korsten, H. H., et al. (2004). Approaches for creating computer-interpretable guidelines that facilitate decision support. *Artif Intell Med* 31:1–27.
- de Dombal, F., de Baere, H., van Elk, P., et al. (1993). Objective medical decision making: acute abdominal pain. In: Beneken J, Thevenin V, eds. *Advances in Biomedical Engineering: Results of the 4th EC Medical and Health Research Programme*. Vol. 7 of Studies in Health Technology and Informatics, 65–87. Burke, VA: IOS Press.
- de Dombal, F., Leaper, D., Staniland, J., et al. (1972). Computer-aided diagnosis of acute abdominal pain. *BMJ* 2:9–13.
- de Dombal, F. T. (1975). Computer-aided diagnosis and decision-making in the acute abdomen. *J R Coll Physicians Lond* 9:211–8.
- de Dombal, F. T. (1984). Computer-aided diagnosis of acute abdominal pain. The British experience. *Rev Epidemiol Sante Publique* 32:50–6.
- de Dombal, F. T. (1990). Computer-aided decision support in acute abdominal pain, with special reference to the EC concerted action. *Int J Biomed Comput* 26:183–8.
- de Dombal, F. T., Leaper, D. J., Horrocks, J. C., et al. (1974). Human and computer-aided diagnosis of abdominal pain: further report with emphasis on performance of clinicians. *BMJ* 1:376–80.
- Dexter, P. R., Perkins, S., Overhage, J. M., et al. (2001). A computerized reminder system to increase the use of preventive care for hospitalized patients. *N Engl J Med* 345:965–70.
- Elstein, A. S., Friedman, C. P., Wolf, F. M., et al. (1996). Effects of a decision support system on the diagnostic accuracy of users: a preliminary report. *J Am Med Inform Assoc* 3:422–8.
- Espino, J., Dara, J., Dowling, J., et al. (2005). SyCo2: A Bayesian Machine Learning Method for Extracting Symptoms from Chief Complaints And Combining Them Using Probabilistic Case Definitions. Pittsburgh, PA: Center for Biomedical Informatics, University of Pittsburgh.
- Espino J., Tsui, F.-C. (2000). A Bayesian network for detecting inhalational anthrax outbreaks. Pittsburgh, PA: Center for Biomedical Informatics, University of Pittsburgh.
- Evans, R. S. (1991). The HELP system: a review of clinical applications in infectious diseases and antibiotic use. *MD Computing* 8:282–8, 315.
- Feigenbaum, E. A. (1977). The Art of Artificial Intelligence: Themes and Case Studies of Knowledge Engineering. Paper presented at 5th International Joint Conference on Artificial Intelligence.
- Fox, J., Johns, N., Lyons, C., et al. (1997). PROforma: a general technology for clinical decision support systems. *Comput Methods Programs Biomed* 54:59–67.
- Friedman, C. P., Elstein, A. S., Wolf, F. M., et al. (1999). Enhancement of clinicians' diagnostic reasoning by computer-based consultation: a multisite study of 2 systems. *JAMA* 282:1851–6.
- Gardner, R. M., Pryor, T. A., Warner, H. R. (1999). The HELP hospital information system: update 1998. *Int J Med Inform* 54:169–82.
- Gerberding, J. L., Hughes, J. M., Koplan, J. P. (2002). Bioterrorism preparedness and response: clinicians and public health agencies as essential partners. *JAMA* 287:898–900.
- Graber, M. A., VanScoy, D. (2003). How well does decision support software perform in the emergency department? *Emerg Med J* 20:426–8.
- Haynes, R. B., Davis, D. A., McKibbon, A., et al. P. (1984). A critical appraisal of the efficacy of continuing medical education. *JAMA* 251:61–4.
- Heckerman, D. E., Horvitz, E. J., Nathwani, B. N. (1992). Toward normative expert systems. Part I. The Pathfinder project. *Methods Inform Med* 31:90–105.
- Hripscak, G., Knirsch, C., Jain, N., et al., (1999). A health information network for managing innercity tuberculosis: bridging clinical care, public health, and home care. *Computers Biomed Res* 32:67–76.
- Hripscak, G., Knirsch, C. A., Jain, N. L., et al. (1997). Automated tuberculosis detection. *J Am Med Inform Assoc* 4:376–81.
- Hughes, J. M., Gerberding, J. L. (2002). Anthrax bioterrorism: lessons learned and future directions. *Emerg Infect Dis* 8:1013–4.
- Johnson, P., Tu S., Jones, N. (2001). Achieving reuse of computable guideline systems. *Medinfo* 10:99–103.
- Kahn, M. G., Bailey, T. C., Steib, S. A., et al. (1996a). Statistical process control methods for expert system performance monitoring. *J Am Med Inform Assoc* 3:258–69.
- Kahn, M. G., Steib, S. A., Dunagan, W. C., et al. (1996b). Monitoring expert system performance using continuous user feedback. *J Am Med Inform Assoc* 3:216–23.
- Kahn, M. G., Steib, S. A., Fraser, V. J., et al. (1993). An expert system for culture-based infection control surveillance. In: *Proceedings of Annual Symposium on Computer Applications in Medical Care*, 171–5.
- Kohn, L. T., Corrigan, J., Donaldson, M. S., et al. (2000). *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academy Press.
- Kolata, G. (2001). Florida Man Is Hospitalized with Pulmonary Anthrax. *New York Times*, October 5.
- Leavitt, M. (2001). Medscape's response to the Institute of Medicine Report: crossing the quality chasm: a new health system for the 21st century. *Med Gen Med* 3:2.
- Ledley, R. S., Lusted, L. B. (1959). Reasoning foundation of medical diagnosis: symbolic logic, probability, and value theory aid our understanding of how physicians reason. *Science* 130:9–12.
- Leist, J. C., Kristofco, R. E. (1990). The changing paradigm for continuing medical education: impact of information on the teachable moment. *Bull Med Libr Assoc* 78:173–9.
- London, S. (1998). DXplain: a web-based diagnostic decision support system for medical students. *Med Ref Serv Q* 17:17–28.
- McAdam, W. A., Brock, B. M., Armitage, T., et al. (1990). Twelve years' experience of computer-aided diagnosis in a district general hospital. *Ann R Coll Surg Engl* 72:140–6.
- McDonald, C. (1976a). Use of a computer to detect and respond to clinical events: its effect on clinician behavior. *Ann Intern Med* 84:162–7.
- McDonald, C., Tierney, W., Overhage, J. (1991). H.I.S. and the physician: direct inpatient order entry by physicians through medical

- gopher workstations—problems and promises. In: *Proceedings IMIA Working Conference on Trends in Modern Hospital Informatics Systems, Gottingen, Germany*, 71–80.
- McDonald, C., Tierney, W., Overhage, J., et al. (1992a). The Regenstrief Medical Record System: 20 years of experience in hospitals, clinics, and neighborhood health centers. *MD Computing* 9:206–17.
- McDonald, C. J. (1976b). Protocol-based computer reminders, the quality of care and the non-perfectability of man. *N Engl J Med* 295:1351–5.
- McDonald, C. J. (1976c). Use of a computer to detect and respond to clinical events: Its effect on clinician behavior. *Ann Intern Med* 84:162–7.
- McDonald, C. J. (1983). Computer technology and continuing medical education: a scenario for the future. *Conn Med* 47:159–62.
- McDonald, C. J., Hui, S. L., Tierney, W. M. (1992b). Effects of computer reminders for influenza vaccination on morbidity during influenza epidemics. *MD Computing* 9:304–12.
- Miller, P. L., Frawley, S. J., Sayward, F. G., et al. (1997). Combining tabular, rule-based, and procedural knowledge in computer-based guidelines for childhood immunization. *Comput Biomed Res* 30(3):211–31.
- Miller, R., Masarie, F. E., Myers, J. D. (1986). Quick medical reference (QMR) for diagnostic assistance. *MD Computing* 3:34–48.
- Murphy, G. C., Friedman, C. P., Elstein, A. S., et al. (1996). The influence of a decision support system on the differential diagnosis of medical practitioners at three levels of training. In: *Proceedings of American Medical Informatics Association Annual Fall Symposium*, 219–23.
- National Aeronautics and Space Administration A. C language integrated production system (CLIPS). <http://www.www.ghg.net/clips/WhatIsCLIPS.html>.
- Neapolitan, R. (2003). *Learning Bayesian Networks*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Overhage, J. M., Tierney, W. M., McDonald, C. J. (1996). Computer reminders to implement preventive care guidelines for hospitalized patients. *Arch Intern Med* 156:1551–6.
- Parker, C. G., Rocha, R. A., Campbell, J. R., et al. (2004). Detailed clinical models for sharable, executable guidelines. *Medinfo* 11:145–8.
- Peleg, M., Tu, S., Bury, J., et al. (2003). Comparing computer-interpretable guideline models: a case-study approach. *J Am Med Inform Assoc* 10:52–68.
- Pynsent, P. B., Fairbank, J. C. (1989). Computer interview system for patients with back pain. *J Biomed Eng* 11:25–9.
- Rind, D. M., Davis, R., Safran, C. (1995). Designing studies of computer-based alerts and reminders. *MD Computing* 12:122–6.
- Rind, D. M., Safran, C., Russell, S. P., et al. (1993). The effect of computer-based reminders on the management of hospitalized patients with worsening renal function. In: *Proceedings of the 15th Annual Symposium on Computer Applications in Medical Care, Washington, D.C.*, 28–32.
- Seyfang, A., Miksch S., Marcos, M. (2002). Combining diagnosis and treatment using ASBRU. *Int J Med Inform* 68:49–57.
- Shannon, M., Burstein, J., Mandl K, et al. (2002). Usage of a web-based decision support tool for bioterrorism detection. *Am J Emerg Med* 20:384–5.
- Shephard, R., Cameron, A., Toribio J., et al. (2005). Bovine syndromic surveillance system. <http://www.www.ausvet.com.au/bosss/>.
- Shiffman, R. N., Michel, G., Essaihi A., et al. (2004). Bridging the guideline implementation gap: a systematic, document-centered approach to guideline implementation. *J Am Med Inform Assoc* 11:418–26.
- Shwe, M. A., Middleton, B., Heckerman, D. E., et al. (1991). Probabilistic diagnosis using a reformulation of the INTERNIST-1/QMR knowledge base. I. The probabilistic model and inference algorithms. *Methods Inform Med* 30:241–55.
- Sorenson D, et al. (1988). Generation and maintenance of ILIAD medical knowledge in a hypercard environment. In: *Proceedings of 12th Symposium on Computer Applications in Medical Care, Washington, D.C.*
- Terenziani, P., Montani, S., Torchio, M., et al. (2003). Temporal consistency checking in clinical guidelines acquisition and execution: the GLARE's approach. In: *Proceedings of American Medical Informatics Association Annual Symposium*, 659–63.
- Tierney, W. M., Miller, M. E., Overhage, J. M., et al. (1993). Physician inpatient order writing on microcomputer workstations. Effects on resource utilization. *JAMA* 269:379–83.
- Travis J., Latin, R. (1991). Development, implementation, and adoption of expert systems in plant pathology. *Ann Rev Phytopathol* 29:343–360.
- Tu, S. W., Musen, M. A. (2001). Modeling data and knowledge in the EON guideline architecture. *Medinfo* 10:280–4.
- Turner, S. (2001). Bayesian Networks & More. <http://www.www.medInformatics.org/mdi207/lab3.html>.
- Wagner, M. M., Pankaskie, M., Hogan, W., et al. (1997). Clinical Event Monitoring at the University of Pittsburgh. In: *Proceedings of American Medical Informatics Association Annual Fall Symposium*, 188–92.
- Wald, J. S., Rind, D., Safran, C., et al. (1995). Patient entries in the electronic medical record: an interactive interview used in primary care. In: *Proceedings of Annual Symposium on Computer Applications in Medical Care*, 147–51.
- Wang, D., Peleg, M., Bu, D., et al. (2003). GESDOR—a generic execution model for sharing of computer-interpretable clinical practice guidelines. In: *American Medical Informatics Association Annual Symposium*, 694–8.
- Warner, H., Jr., Blue, S. R., Sorenson, D., et al. (1997). New computer-based tools for empiric antibiotic decision support. In: *Proceedings of American Medical Informatics Association Annual Symposium*, 238–42.
- Warner, H. R., Jr. (1989). Iliad: moving medical decision-making into new frontiers. *Methods Inform Med* 28:370–2.

- Warner, H. R., Jr. (1990). Iliad: diagnostic tools for general medicine. Interview by Bill W. Childs. *Healthcare Inform* 7:38.
- Warner, H. R., Jr., Bouhaddou, O. (1994). Innovation review: Iliad—a medical diagnostic support program. *Top Health Inform Manag* 14:51–8.
- Warner, H. R., Toronto, A. F., Veasey, G. (1964). Experience with Bayes theorem for computer analysis of congenital heart disease. *Ann N Y Acad Sci* 115:558–67.
- Warner, H. R., Toronto, A. F., Veasey, G., et al. (1961). A mathematical approach to medical diagnosis: application to congenital heart disease. *JAMA* 177:177–83.
- Weinstein, M. C. (1980). The value of information. In: *Clinical Decision Analysis*, 131–67. Philadelphia: Saunders.
- Williamson, J. W., German, P. S., Weiss, R., et al. (1989). Health science information management and continuing education of physicians: a survey of U.S. primary care practitioners and their opinion leaders. *Ann Intern Med* 110:151–60.
- Wilson, D. H., Wilson, P. D., Walmsley, R. G., et al. (1977). Diagnosis of acute abdominal pain in the accident and emergency department. *Br J Surg* 64:250–4.
- Wilson, P. D., Horrocks, J. C., Lyndon, P. J., et al. (1975). Simplified computer-aided diagnosis of acute abdominal pain. *BMJ* 2:73–5.
- Yasnoff, W. A., Humphreys, B. L., Overhage, J. M., et al. (2004). A consensus action agenda for achieving the national health information infrastructure. *J Am Med Inform Assoc* 11:332–8.