HEME: A COMPUTER PROGRAM FOR DIAGNOSIS-ORIENTED ANALYSIS OF HEMATOLOGIC DISEASE

By RALPH L. ENGLE, Jr., M.D. and (*by invitation*) BETTY J. FLEHINGER, Ph.D.

NEW YORK

For 15–20 years we have heard predictions concerning the use of computers as aids to the physician in his decision making. In spite of considerable efforts, we see very little evidence of their realization. Reasonably successful programs have been developed for relatively simple situations such as the partial interpretation of an electrocardiogram[1] and the analysis of electrolyte data.[2] We are still in the early stages, however, in the application of the same principles to larger areas of medicine such as major subspecialties. For many years our group has been studying computer-aided diagnosis in the broad subspecialty of hematology.[3-8] Many investigators have contributed significant ideas to the field including Ledley and Lusted[9] who were the first to discuss the application of symbolic logic, probability, and value theory in medical decision-making; Warner, Toronto, Veasey, and Stephenson[10] who used these tools in a system for the diagnosis of congenital heart disease; Gorry and Barnett[11] who applied sequential decision theory to diagnostic problems; and Gustafson[12] who experimented with subjective judgment in the estimation of probabilities.

Our present system can best be appreciated by seeing a typical, though oversimplified, exchange between a physican or student and the computer.

At the start of the program the computer asks which function the physician wishes to exercise. The physician indicates by entering a 1 that he wishes to enter a series of findings on a patient. He then enters, by code number, one or more findings depending upon their availability.

FUNCTION?
1
ENTER SXS
7, 12, 14, 21, 56, 64, 74, 76, 89, 105, 134, 140, 150, 200, 220, 275, 280,
284, 289, 292, 294, 298, 304, 309, 318, 491, 495, 497, 501, 503
ENTER SXS

The codes are selected from a checklist of 585 findings, a portion of which

From the Departments of Medicine and Public Health, Cornell University Medical College, New York; and the I.B.M. T.J. Watson Research Center, Yorktown Heights, New York and the Department of Biomathematics, Cornell University Graduate School of Medical Sciences, New York.

is shown. For some findings the physician may pick one of several mutually exclusive and exhaustive alternatives or, for others, he may indicate a positive-negative type answer.

```
        284+ LYMPHOCYTES <20%
        285+ LYMPHOCYTES 20-39%
        286+ LYMPHOCYTES 40-59%
        287+ LYMPHOCYTES 60-79%
        288+ LYMPHOCYTES ≥ 80%
        289+ LYMPHOCYTES ATYP IN PB - NONE
        290+ LYMPHOCYTES ATYP IN PB < 10% TOTAL LYMPHS
        291+ LYMPHOCYTES ATYP IN PB ≥ 10% TOTAL LYMPHS
292- 293+ MONOCYTES >5%
294- 295+ EOSINOPHILS >3%
        296+ GRANULOCYTES (NEUT, EOSIN, BASOPHILS) <2%
        297+ GRANULOCYTES (NEUT, EOSIN, BASOPHILS) 2-49%
        298+ GRANULOCYTES (NEUT, EOSIN, BASOPHILS) 50-69%
        299+ GRANULOCYTES (NEUT, EOSIN, BASOPHILS) ≥ 70%
300- 301+ NEUTROPHILS HYPERSEGMENTED
302- 303+ GRANULOCYTES IMMATURE IN PB > 4%
```

If the physician wishes a list of the patient's findings he requests function 4. He is given a choice of a complete list or, if he enters a 1, a list of demographic and all abnormal findings. In this case the physician has asked for the latter.

```
FUNCTION?
4
ENTER "1" FOR ABNORMAL HX FORM
1
HIST
   7 AGE 40-49 YRS
  12 SEX MALE
  14 RACE WHITE
  56 FATIGUE, LETHARGY OR MALAISE
  74 PALPITATION
  76 PRECORDIAL PAIN
  89 BOWEL FUNCTION-DIARRHEA
PE
XRAY
LAB
 275 RBC INDICES HGB 7-12.9, MCV <80, MCH <30
 284 LYMPHOCYTES <20%
 309 ANISOCYTOSIS & POIKILOCYTOSIS
```

The physician then requests the listing of differential diagnoses and probabilities by entering function 5. The computer, using a version of Bayes' Theorem, computes the probability that the patient has each of the 40 diseases currently in the system and lists those with probability greater than 1 %. The physician is able to compare this list of probabilities with his own clinical judgment. In this case there is not enough information to give a high probability of any disease.

FUNCTION?
5

DIFFERENTIAL DIAGNOSIS

*	5 IRON DEFICIENCY ANEMIA	38.7%
*	30 THALASSEMIA MINOR	9.1%
*	2 ANEMIA OF MALABSORPTION	2.8%
*	37 ANEMIA OF MALIG., NON-HEM.	1.3%

Bayes' Theorem provides a method of calculating the probability of a disease in a patient after his findings are known, i.e.,

$$\text{Prob. (Disease i/Findings)} = \frac{\Phi_i \prod p_{ij}}{\Phi_i \prod p_{ij} + (1 - \Phi_i) \prod q_{ij}},$$

where Φ_i is the frequency with which the disease occurs in the population under consideration,

p_{ij} is the probability that a patient with disease i has finding j at the time the disease is diagnosed,

and q_{ij} is the probability that a patient who does *not* have disease i, but for whom the descriptor corresponding to finding j is observed during the diagnostic process, does have finding j at the time of observation.

The Φ's, p's and q's are currently judgmental estimates made by the clinicians responsible for the program based on frequency data previously collected. It is intended that they be automatically modified as patient data accumulate. In our system Bayes' Theorem is applied separately for each disease and each time it is employed it refers to a universe of hematology patients consisting of only two groups, patients who have the given disease and patients who do not have that particular disease. In the usual application of Bayes' Theorem the patient is considered to be in a universe consisting of many groups, one for each hematologic disease, and a patient

can be placed in only one of these groups. Thus, the estimated probabilities of all the diseases must add up to one. In our system, the probabilities do not add up to one, allowing for the very real possibility that a patient has more than one disease.

Since iron deficiency anemia has a relatively high score in the differential diagnosis, the physician asks the computer for the rationale behind that diagnosis by entering function 6 and the disease code number 5 for iron deficiency anemia. The computer prints out a list of the findings supporting the diagnosis and those opposing the diagnosis, in order of their significance. In this example there is only one finding in each category. However, the weight of each finding is shown as p/q (p_{ij}/q_{ij}) for making the diagnosis or q/p (q_{ij}/p_{ij}) for ruling out the diagnosis. These ratios are calculated by the computer from probabilities used in Bayes' Theorem.

```
FUNCTION?
6
ENTER DISEASE NUMBER FOR P/Q RATIOS?
5

RECORDED SYMPTOM P/Q RATIOS FOR 5 IRON DEFICIENCY ANEMIA
P/Q FOR DIAGNOSIS
    12.0 #  275 RBC INDICES HGB 7-12.9, MCV <80, MCH <30
Q/P AGAINST DIAGNOSIS
     2.5 #  12 SEX MALE
```

The p/q ratio is a significant concept for interpreting the computer analysis. If p/q for any finding is much larger than one, the observation of the finding tends to lead to a diagnosis of the disease; if p/q is much smaller than one, the finding tends to rule out the disease; and if p/q is close to unity, the finding has little relevance to the diagnosis. For convenience, when p/q is less than one we display its inverse, q/p. If either ratio is greater than 1000, its value is shown as **********.

If the physician thinks there is enough evidence to pursue the diagnosis of iron deficiency anemia he may request a list of suggested findings to investigate. He does this by entering function 9 and the code number for the disease in question. Unrecorded findings which support or oppose the diagnosis are listed in order of p/q or q/p. The physician compares this list with his own judgment and decides on the priorities for further examinations.

Though not shown here, those tests mandated by accepted standards of patient care could be flagged.

```
FUNCTION?
9
ENTER DISEASE NUMBER FOR P/Q RATIOS ?
5

UNRECORDED SYMPTOM P/Q RATIOS FOR 5 IRON DEFICIENCY ANEMIA
P/Q FOR DIAGNOSIS
    980.0 # 345 BM IRON-ABSENT
     80.0 # 427 SERUM COPPER HIGH
     33.0 # 395 RESPONSE TO IRON-POSITIVE
     19.8 # 435 SERUM IRON BINDING CAP (TOTAL) HIGH
     17.4 # 340 BM CELLULARITY-INCREASED
     16.5 # 176 FINGERNAILS-SPOONED OR BRITTLE
     10.0 # 311 TARGET CELLS
     10.0 # 321 RETICULOCYTE COUNT <1%
      9.3 # 430 SERUM IRON LOW <70
      8.3 # 197 TONGUE SMOOTH OR SORE
      3.0 #  72 DYSPNEA
      2.7 # 564 ACHLORHYDRIA-PRESENT
Q/P AGAINST DIAGNOSIS
********** # 433 SERUM IRON BINDING CAP (TOTAL) LOW
********** # 432 SERUM IRON HIGH >130
********** # 348 BM IRON-INCREASED
    100.0  # 325 RETICULOCYTE COUNT ≥ 10%
     97.0  # 394 RESPONSE TO IRON-NEGATIVE
     50.0  # 338 BM CELLULARITY-DECREASED
     10.0  # 337 BM MEGALOBLASTIC
      5.0  # 426 SERUM COPPER NORMAL
```

After any or all of the additional tests suggested, or any other tests the physician wishes, have been performed, the physician may enter the additional finding codes into the computer through function 2.

```
FUNCTION?
2
ENTER SXS
435, 175, 311, 321, 430
ENTER SXS
```

A revised differential diagnosis may then be requested through function 5. The findings, now including low serum iron and high iron binding capacity, have made the diagnosis of iron deficiency anemia virtually certain. At the same time thalassemia minor has appeared on the list with relatively high probability.

```
FUNCTION?
5

DIFFERENTIAL DIAGNOSIS

 #  5 IRON DEFICIENCY ANEMIA          100.0%
 # 30 THALASSEMIA MINOR                89.9%
 # 37 ANEMIA OF MALIG., NON-HEM.       28.0%
 #  2 ANEMIA OF MALABSORPTION           2.8%
```

The physician requests the rationale for this diagnosis by entering function 6 and disease code number. The findings for and against the diagnosis are displayed together with the weights. Again, the physician compares his own clinical judgment with the ordered list of findings. Here there are no findings against the diagnosis.

```
FUNCTION?
6
ENTER DISEASE NUMBER FOR P/Q RATIOS?
30

RECORDED SYMPTOM P/Q RATIOS FOR 30 THALASSEMIA MINOR
P/Q FOR DIAGNOSIS
    12.0 # 311 TARGET CELLS
     7.4 # 321 RETICULOCYTE COUNT <1%
     2.5 # 309 ANISOCYTOSIS & POIKILOCYTOSIS
Q/P AGAINST DIAGNOSIS
```

The physician asks for additional tests to confirm or rule out this diagnosis. These are listed as before.

```
FUNCTION?
9
ENTER DISEASE NUMBER FOR P/Q RATIOS?
30

UNRECORDED SYMPTOM P/Q RATIOS FOR 30 THALASSEMIA MINOR
P/Q FOR DIAGNOSIS
    90.0 # 509 ELECTROPHORESIS HEMOGLOBIN A2 INCREASED
    10.0 # 348 BM IRON-INCREASED
    10.0 # 539 ERYTHROCYTE OSMOTIC FRAGILITY-DECREASED
     9.5 #  18 STOCK MEDITERRANEAN
     3.3 # 511 ELECTROPHORESIS HEMOGLOBIN F INCREASED
```

```
Q/P AGAINST DIAGNOSIS
********** # 267 X-RAY SKULL VERTICAL STRIATION
********** # 258 X-RAY BONE LYTIC LESIONS, FRACTURE ABSENT
    50.0   # 541 ERYTHROCYTE OSMOTIC FRAGILITY-INCREASED
    18.0   #  17 STOCK MEDITERRANEAN-ABSENT
    10.0   # 345 BM IRON-ABSENT
     9.9   # 508 ELECTROPHORESIS HEMOGLOBIN A2 NORMAL
     5.0   # 338 BM CELLULARITY-DECREASED
```

The physician orders additional studies and enters the results.

```
FUNCTION?
2
ENTER SXS
258, 17
ENTER SXS
```

The physician requests a third differential diagnosis which, as we can see, eliminates thalassemia minor but suggests that in addition to iron deficiency anemia the patient has anemia secondary to malignancy, a possibility the physician may choose to pursue.

```
FUNCTION?
5

DIFFERENTIAL DIAGNOSIS

 #   5 IRON DEFICIENCY ANEMIA          100.0%
 #  37 ANEMIA OF MALIG., NON-HEM.       79.6%
 #   6 MULTIPLE MYELOMA                  4.4%
```

From studies such as these we have concluded that HEME is useful in teaching hematology and has potential as an aid in diagnosis and as a means for studying the diagnostic process itself.

General use of this type of program has not been achieved because of a circular problem. Physicians are currently unwilling to collect and record objective data in the precise detail required for computer analysis. This will not change until they see some real benefits. On the other hand, predicted benefits cannot be realized or evaluated until a significant system has been developed, which, of course, requires a large store of carefully collected and edited data.

In HEME we have a system which, as a teaching tool, will stimulate

physicians and students to interact effectively with the computer. Furthermore, the version of Bayes' Theorem used in HEME requires far fewer probabilities of findings in diseases than does the usual version. Since each disease is analyzed separately, p's and q's need be entered only for those findings relevant to the diagnosis of that disease. Inherent in the system is the capacity to grow and improve itself in three ways. New diseases may be added without changing the rest of the system; new findings relevant to one or more diseases may be added with only minor changes; and the probabilities required for Bayes' Theorem may be modified automatically as data accumulate.

We are currently writing a new HEME program using the basic approach just described with a number of features which should further improve its usefulness. Three stages of diagnosis will be used. The patient will first be placed in a general category of hematologic disease, e.g., megaloblastic anemia, using findings that are important in defining this broad category. At the second stage only findings required to distinguish various forms of megaloblastic anemia will be used and a specific diagnosis such as pernicious anemia made. At the third stage, information of textbook type will be available in the computer and accessible to the physician. It will include information about subcategories or stages of disease, treatment, and references.

In summary, we have demonstrated the way in which HEME works to aid in the physicians' decision-making. The physician maintains control at all times and can use the information from the computer as a bench mark against which to compare his own thinking and as an accessory to his own memory. Such a system also has applicability as a teaching tool and for studying the diagnostic process.

Acknowledgments

The authors wish to acknowledge the important contributions of Dr. Scott Allen, Dr. Richard Friedman, Dr. Martin Lipkin, Dr. B. J. Davis, and Dr. Leo L. Leveridge to the development of this program.

References

1. Bonner, R. E., Crevasse, L., Ferrer, M. I., and Greenfield, J. C., Jr.: A new computer program for analysis of scalar electrocardiograms. *Comp & Biomed Res* **5:** 629–653, 1972.
2. Bleich, H. L.: Computer evaluation of acid-base disorders. *J Clin Invest* **48:** 1689–1696, 1969.
3. Lipkin, M., Engle, R. L., Jr., Davis, B. J., et al: Digital computer as aid to differential diagnosis. *Arch Int Med* **108:** 56–72, 1961.
4. Engle, R. L., Jr., and Davis, B. J.: Medical diagnosis: present, past, and future. I. Present concepts of the meaning and limitations of medical diagnosis. *Arch Int Med* **112:** 512–519, 1963.

5. ENGLE, R. L., JR.: Medical diagnosis: present, past and future. II. Philosophical foundations and historical development of our concepts of health, disease, and diagnosis. *Arch Int Med* **112:** 520–529, 1963.
6. ENGLE, R. L., JR.: Medical diagnosis: present, past, and future. III. Diagnosis in the future, including a critique on the use of electronic computers as diagnostic aids to the physician. *Arch Int Med* **112:** 530–543, 1963.
7. LIPKIN, M., ENGLE, R. L., JR., FLEHINGER, B. J., et al: Computer-aided differential diagnosis of hematologic diseases. *Proc NY Acad Sci* **161:** 670–679, 1969.
8. ENGLE, R. L., JR., LIPKIN, M., AND FLEHINGER, B. J.: Computer-aided differential diagnosis of hematologic diseases (A Bayesian probability model). In Computer Diagnosis and Diagnostic Methods. Edited by J. A. Jacquez. Springfield, Ill., Charles C Thomas, 1972, pp. 279–293.
9. LEDLEY, R. S., AND LUSTED, L. B.: Reasoning foundations of medical diagnosis. Symbolic logic, probability, and value theory aid our understanding of how physicians reason. *Science* **130:** 9–21, 1959.
10. WARNER, H. R., TORONTO, A. F., VEASEY, G., AND STEPHENSON, R.: A mathematical approach to medical diagnosis. Application to congenital heart disease. *JAMA* **177:** 177–183, 1961.
11. GORRY, G. A., AND BARNETT, G. O.: Sequential diagnosis by computer. *JAMA* **205:** 849–854, 1968.
12. GUSTAFSON, D. H., KESTLY, J. J., GREIST, J. H., AND JENSEN, N. M.: Initial evaluation of a subjective Bayesian diagnostic system. *Health Services Res* **6:** 204–213, 1971.

DISCUSSION

DR. HERBERT G. LANGFORD (Jackson): As a House Officer in Baltimore, there was a British neurologist on the Housestaff who had just come from Queen's Square, and he said that he was very distressed because combinations of signs and symptoms which usually suggested a brain tumor at Queen's Square rarely produced one when he was seeing the unfiltered selection of patients in Baltimore. That is, the frequency of a finding depends on the kind of business you are doing, how much filtering has gone on. So I ask: are your probabilities derived from the experience of your own group or from clinician's idea of the world at large.

DR. ENGLE: The p's and q's as well as the prior probabilities of the diseases are judgmental values arrived at by a panel of hematologists at our institution. The patients under consideration as we make this judgmental value are hematology patients seen on the Hematology Service at New York Hospital so that we have a limited population group. I think it is too early to say how generally applicable such estimates would be. However I did not stress, perhaps I mentioned it briefly, that the program we have is capable of learning and as cases accumulate and data are entered into the system, feed back occurs and modifies the original estimates made by the physicians. Thus, the system is capable of continually improving itself.

DR. ROBERT M. BIRD (Bethesda): Dr. Engle, congratulations! Our paths cross again. In trying to assess a new modality such as Computer Aided Instruction, could you give your assessment as to whether this is a tool toward decision making related to an immediate clinical problem or is it a tool in pursuit of a more logical intellectual analysis, a more systematic approach to the problems of hematology? Would you comment on that because I think there is need to clarify the goal. Will confusion hurt the reception by physicians? It really doesn't save a lot of time is what I am thinking at this moment.

DR. ENGLE: Evaluation of a program of this sort is extremely difficult. For example, if you use it as a diagnosis machine, which is not something we have in mind, depending on the mix of cases used, one can get fairly good results in diagnosing 60 to 95% of the cases. The variability occurs because some cases are much more difficult than others. We don't feel this is an adequate way to test the system. We have however attempted to get some idea of how well the system works in two ways. First, we have had a fourth year medical student follow cases on the hematology service from the time the patient is admitted to the hospital through the time the diagnosis is made using the computer as a diagnostic tool in the way I have just demonstrated. The computer results were compared with the decisions of the panel of physicians following the patients, as usually occurs in a teaching hospital. From studies of this sort we are encouraged since the computer did quite well in terms of keeping up with the several physicians that were following the case. The second experience was at the University of Wisconsin where Dick Friedman, who was a member of our group when he was a medical student at Cornell, has been using the program in the Department of Hematology to see how well students could use it as a learning tool. The students felt that it was a great stimulus to learning especially by pointing up what might be done and what not. They could compare their own thinking with the results from the computer. It does, in effect, what you said, mainly enable a student to tidy up his thinking about diagnosis. It is still a rather crude system; we have hopes of improving it considerably. We think we are heading in the right direction and expect that soon we will have a very useful teaching tool.