

“NEONATE”—An Expert Application for the “HELP” System: Comparison of the Computer’s and the Physician’s Problem List

Arie Franco, M.D., Ph.D.,* Fred L. Farr, Ph.D.,*
Jerald D. King, M.D.,† Justin S. Clark, Ph.D.,‡ and
Peter J. Haug, M.D.‡

NEONATE is a prototype of an expert application for the HELP Hospital Information System. Its goal is to improve documentation in the Newborn Intensive Care Unit. The decision module of NEONATE is designed to produce an admission problem list. In this paper, the admission problem list that NEONATE generates was compared to the admission problem list of the current CETUS system for 30 patients. These were compared to a retrospectively constructed gold standard problem list. Of 101 problems in the gold standard list, 56 were on the current admission reports; 82 were found by NEONATE. NEONATE found 31 problems missed on the current admission reports; the current admission reports contained 5 problems missed by NEONATE. The current admission reports contained 9 false positives; whereas NEONATE’s reports contained 27. Of the 27, 16 were caused by a single rule in NEONATE. We conclude that an expert system has great potential for improving the documentation of the patient problem list.

INTRODUCTION

NEONATE is an expert system under development at the Primary Children’s Medical Center (PCMC) in Salt Lake City, Utah. It is part of the HELP system¹⁻³ which is a Hospital Information System (HIS) with expert system features. It is being developed as a possible alternative to the CETUS-100⁴⁻⁶ departmental desk top system which is currently in use at the Newborn Intensive Care Unit of the PCMC.

The CETUS-100 system was designed to improve the documentation of the admission/discharge report. It has been reported⁴ that automatic computer production of admission/discharge letters to physicians and agencies accounted for an 80% reduction in physician and nursing admission/discharge paperwork, a 12-fold increase in completeness

From the *Department of Information Systems, Primary Children’s Medical Center, Salt Lake City, Utah; †Department of Pediatrics, University of Utah, Salt Lake City, Utah 84112; ‡Department of Medical Informatics, University of Utah, Salt Lake City, Utah 84112.

of medical data, and a 96% reduction in time from patient discharge to receipt of patient care information within the medical community.

Data for the CETUS-100 admission report is entered into the computer by a clerk who uses an information sheet filled out by the admitting physician, notes from the transport team, the labor and delivery record, and nursery records. Information available from these sources is often sketchy. Since the CETUS-100 system does not provide decision support, the problem list on its admission report contains only those items explicitly stated on the information sheet submitted by the admitting physician or, possibly, on one of the other sources.

NEONATE is being developed to replace the CETUS-100 and expand functionality. Both systems produce an admission report that includes items such as a full history of the mother's pregnancy, delivery history, and physical examination findings. The fact that NEONATE is part of an inter-hospital HIS means that some of the input data such as demographic data and delivery history may already be entered into the system. Also, it will be easy to add new information to the admission report, such as laboratory results, since they will already be in the system.

In this article the effect on documentation of adding expert system features to a HIS in order to generate an admission problem list was studied by comparing the problem list in the current CETUS admission report to the problem list produced by NEONATE.

A comparison study between CETUS's largely clinician composed problem list (hereafter called the physician problem list) vs. NEONATE's expert system composed (clinician unedited) problem list was undertaken to test the following hypotheses:

1. NEONATE's computer inference process picks up problems that are currently missed by the physician problem list.
2. NEONATE's expert system produces more true positive items on the problem list and fewer false positives than the physician problem list.

METHODS

The development of a prototype of the NEONATE system has been accomplished. This prototype deals only with tachypnea which is the most common finding found in newborns admitted to the intensive care units. This finding is able to trigger in NEONATE a list of 35 diseases that make up the differential diagnosis of tachypnea.

The Knowledge Base of NEONATE consists of three processes:

1. The Initiator which looks for abnormal findings in the database. For each abnormal finding that is found in the database, the Initiator puts all related differential diagnostic problems in the problem list with the status "suspected."
2. The Problem List which is a list of all potential problems/diseases under consideration for a given patient. Each problem has an associated status linked to it. The problem statuses are "suspected," "highly suspected," "confirmed," or "resolved."
3. The Inference Engine which consists of a series of frame-like modules—one module for each problem or disease. If a particular problem exists in the problem list, the corresponding frame will be activated and may change the status of the problem or delete the problem from the list. The problem list produced by NEONATE's inference process is only a suggested list for the admitting physician. It is not complete until the physician

has edited it to remove inappropriate problems and to add problems missed by the computer logic; this, the inference process is to be viewed as a documentation aid. Having stated this, it should be noted that for the purposes of the experiment described in this article the NEONATE problem list is composed solely by the expert system without the physician editing it would normally receive.

Fifty medical records of patients with an initial finding of tachypnea were pulled from the archive of the Primary Children's Medical Center. Ten records were selected with the admitting diagnosis of Hyaline Membrane Disease, 10 records with Transient Tachypnea of the Newborn, 10 records with Aspiration Pneumonia, 10 records with Infectious Pneumonia, and 10 records with extraventilatory air (Pneumothorax or Pneumomediastinum). However, it became apparent that separation into these groups was not possible because most of the records listed multiple problems in the admission problem list that cause respiratory distress. It is not unusual that an infant admitted to the Newborn Intensive Care Unit has multiple problems. For this reason it was decided to treat the 50 medical records as one group. Only 30 random records were picked up from this group of 50 records to prove the two hypotheses of this article. Because many of the common problems that NEONATE is capable of diagnosing are closely related and because of the relative simplicity of diagnosing the rare diseases that exist in the NEONATE knowledge base, it was impractical to choose the rare diseases for evaluating the system. For example confirmation of esophageal atresia does not pose any difficulties to the physician or to the computer if the right x-ray data are entered. On the other hand differentiation of Aspiration Pneumonia from Infectious Pneumonia or from Transient Tachypnea of the Newborn is sometimes difficult for the physician and is also a challenge to the computer.

If tachypnea is present, the initiator inserts all 35 diseases that make up the differential diagnosis of tachypnea into the problem list. The inference engine is capable of eliminating a few of them, provided adequate data are present in the computerized data base. Some of the problems can be eliminated from the problem list only by the physician. Most of the time two to three problems are promoted by the inference engine to "Highly Suspected" or "Confirmed" and about 10 diseases are left under "Suspected" waiting for more data or waiting to be eliminated by the physician. The neonatologists found this list to be inappropriate in terms of size; therefore, it was decided not to display diseases in the problem list with the attribute of "Suspected" except for Infectious Pneumonia and Sepsis which need to be listed since treatment is initiated as though they are present. The undisplayed diseases are not eliminated from the problem list, but remain for future reference. The retrospective studies were conducted by referring to the problems that are actively displayed on screen and are apparent to the user physician. The fact that the computer knows much more than what it shows would only strengthen the capabilities of the computer to request information for ruling out these problems, if necessary. If more data are available these problems can be given a higher attribute that allows them to be displayed on the screen or on the print-out without overwhelming the users with a long list of problems.

1. Methods to demonstrate that NEONATE picks up problems missed in the physician problem list are described below:

For the purpose of this study, in addition to the NEONATE admission problem list that the computer generates, and in addition to the admission problem list as it is written in the CETUS system admission report, a Gold Standard list was defined. This Gold

Standard list included all the problems (inferred from the patient's chart) that the patient should have been suspected of having at the time of admission. Problems were entered into the Gold Standard list by the following rules:

(a) All diseases/problems that are mentioned in the record's discharge report, having their onset during the day of admission, and being part of the closed list of 35 diseases NEONATE is able to diagnose.

(b) All diseases that are explicitly mentioned in the handwritten admission notes by the admitting physician, and that are part of the list of 35 diseases NEONATE is able to diagnose.

(c) All diseases that are explicitly mentioned in the first chest x-ray report, if the film was taken during the first day of admission, and the diseases are part of the closed list of 35 diseases that NEONATE is able to diagnose. (Chest x-ray findings as mentioned in the report are not interpreted as a problem. The name of the disease should be specifically written. For example: If it was written in the report "The findings suggest Aspiration Pneumonia," then the Problem Aspiration Pneumonia was inserted into the Gold Standard list.)

(d) For the problems: Anemia, Neonatal hypocalcemia, Metabolic acidosis, and Hypokalemia, the problem was inserted to the gold standard list if the corresponding laboratory value (hemoglobin, total calcium, pH and plasma potassium level) during the first 24 hr of admission was below the normal range as defined by the rules of the corresponding frame for the attribute "highly suspected." These are the same values as they appear in textbooks.

The goal of this part of the study was to determine the potential contribution of an expert system; whereas, the next part of the study compares the quality of the NEONATE problem list to the physician problem list.

2. Below we discuss methods suggesting that NEONATE produces more true positive problems and fewer false positive problems than the physician problem list:

This study tests to see if NEONATE's expert system is more complete than the physician problem list. Here, in addition to the NEONATE admission problem list, the physician admission problem list, and the Gold Standard list, four lists were defined:

(a) A list that includes all problems that exist in the physician admission problem list and in the gold standard list.

(b) A list that includes all problems that exist in the physician admission problem list, but not in the gold standard list.

(c) A list that includes all problems that exist in the NEONATE admission problem list and in the gold standard list.

(d) A list that includes all problems that exist in the NEONATE admission problem list, but not in the gold standard list.

The true positive value of NEONATE was defined as the number of problems that exist in list c. The true positive value of the physician problem list was defined as the number of problems that exist in list a.

The false-positive value of NEONATE was defined as the number of problems that exist in list d. The false-positive value of the physician problem list was defined as the number of problems that exist in list b.

The false-negative value of NEONATE was defined as the number of problems that results from the subtraction of the number of problems that determines the true-positive

value of NEONATE from the number of problems that exist in the gold standard list. In the same way the false negative value of the physician problem list was defined.

The true-negative value of NEONATE was determined from the subtraction of the true-positive value of NEONATE, the false-negative value of NEONATE and the false-positive value of NEONATE from 35 (which is the total number of problems NEONATE is able to diagnose). The same applies to the definition of the true negative value of the physician problem list.

To demonstrate that NEONATE performs better than the current physician problem list (or the current problem list that is typed into the CETUS system manually from clinician notes, since CETUS has no decision capabilities) it must be shown that the true-positive value of NEONATE is greater than the true-positive value of the physician problem list and the false-positive value of NEONATE is smaller than the false-positive value of the physician problem list.

Each of the four parameters (true-positive, false-positive, true-negative, and false-negative) was calculated for the physician problem list and NEONATE in each of 30 charts. The results were averaged to obtain a single value for each of the parameters. From these values the sensitivity was calculated in each of the 30 records as True positives/(True positives + False negatives) for NEONATE and for the physician problem list and the results were averaged. The specificity for both problem lists was calculated in each of the 30 records as True negatives/(True negatives + False positives) and the results were averaged.

RESULTS

1. Results of the study to determine if NEONATE picks up problems that are missed by the physician problem list:

Table 1 shows that of the 101 problems in the gold standard list there are 31 problems in the NEONATE admission problem list that were missed in the physician admission problem list. A mean of 1.033 problems that fits these criteria was found from the study of the 30 records. In contrast, there were only 5 problems caught by the physician or other health care workers that were missed by NEONATE.

2. Results of the study to determine if NEONATE produces more true-positive problems and fewer false-positive problems than the physician problem list:

On studying 30 charts it was shown that of 101 problems, NEONATE found 82 for a mean true-positive value of 2.733 while the physician lists contained 56 for a mean true-positive value of 1.867 ($p < 0.01$). The assumption that the mean of the false positive values of NEONATE is smaller than the mean of the false positive values of the physician problem list proved to be wrong since 27 false positives were found with NEONATE for a mean false positive value of 0.900 while only 9 false positive values were found in the physician problem list for a mean false positive value of 0.300 ($p < 0.01$).

Table 2 represents the means of the four parameters (true positive values, true negative values, false positive values and the false negative values). The difference in the values between the NEONATE problem list and the physician problem list for each of the four parameters was significant ($p < 0.01$).

Table 1. Number of Gold Standard List Problems in the Physician List, NEONATE List, Both Lists, or Neither List in Each of the 30 Records

Record	Physician (only)	NEONATE (only)	Both	Neither	Gold Standard
1	1	1	0	0	2
2	0	0	1	1	2
3	2	2	3	0	7
4	0	2	1	1	4
5	0	1	3	0	4
6	0	1	1	0	2
7	0	1	2	0	3
8	0	1	1	2	4
9	0	1	2	0	3
10	0	2	1	0	3
11	0	0	2	1	3
12	0	0	3	0	3
13	0	1	0	0	1
14	0	2	1	0	3
15	0	0	3	1	4
16	0	1	2	0	3
17	0	0	1	3	4
18	1	1	1	1	4
19	0	1	2	0	3
20	0	1	2	0	3
21	0	1	2	1	4
22	0	0	2	0	2
23	0	2	0	0	2
24	0	1	1	1	3
25	0	2	2	1	5
26	0	1	3	0	4
27	0	2	3	0	5
28	0	1	2	0	3
29	0	1	2	1	4
30	1	1	2	0	4
Total	5	31	51	14	101
Mean	0.167	1.033	1.7	0.467	3.367

Table 3 represents the means of the sensitivity and the specificity for both problem lists. Here also the difference between the two problem lists is significant ($p < 0.01$).

DISCUSSION

As it was shown explicitly that there are problems that exist in the NEONATE admission problem list and in the gold standard list, but do not appear in the physician admission problem list, it can be concluded that NEONATE's computer inference process picks up problems that are currently not being documented by the clinicians. In this study the comparison is done with the unedited version of the admission problem list of NEONATE. In the future the problem list is expected to be edited by physicians; therefore,

Table 2. Means of True Positive, False Positive, True Negative, and False Negative Values of the NEONATE Problem List and the Physician Problem List

	NEONATE Problem List	Physician Problem List
True Positives mean	2.733	1.867
False Negatives mean	0.633	1.500
False Positives mean	0.900	0.300
True Negatives mean	30.733	31.333

it seems that the NEONATE expert system edited by clinicians may perform better than clinicians alone.

NEONATE performs better than the physician admission problem list in a sense that it diagnoses more true positive problems, but the price to be paid is an increase in false positive problems. The sensitivity of NEONATE is far better than the sensitivity of the physician problem list. It is well shown that NEONATE catches more true-positive problems than the physician admission problem list. It is difficult to conclude by the results whether the specificity of NEONATE is less than the specificity of the physician admission problem list. The computer deals only with a finite number of diseases (35 diseases in this case). In real clinical practice a much larger number of problems ought to be considered, an event that might increase the specificity value of NEONATE. Nevertheless, the values as they stand now (97% for NEONATE and 99% for the physician problem list) are similar. By relying on the experiment that demonstrates that the false-positive value of the physician problem list is smaller than the false-positive value of NEONATE, it can be concluded that the physician problem list does better in not listing problems that do not exist in the gold standard list than does NEONATE. A major contribution to this gap lies in the fact that NEONATE lists all cases of "Infectious Pneumonia" that get the status "suspected" if tachypnea is present. Fourteen cases of "Infectious Pneumonia" were true-positive in NEONATE and 16 cases were false positive in NEONATE. Sepsis, the other "suspected" status problem that makes NEONATE's problem list, is also listed whenever tachypnea is present, but in this case 28 were true-positive and only 2 were false positive. Figure 1 displays a histogram that compares the number of the true positives between NEONATE and the physician problem list by problem, whereas Figure 2 displays a histogram that compares the number of the false-positives between NEONATE and the physician problem list by problem. From these histograms it can be concluded that the difference in the true positive values, which

Table 3. Sensitivities and Specificities of the NEONATE Problem List and the Physician Problem List

	NEONATE Problem List	Physician Problem List
Sensitivity	82.82%	53.77%
Specificity	97.19%	99.07%

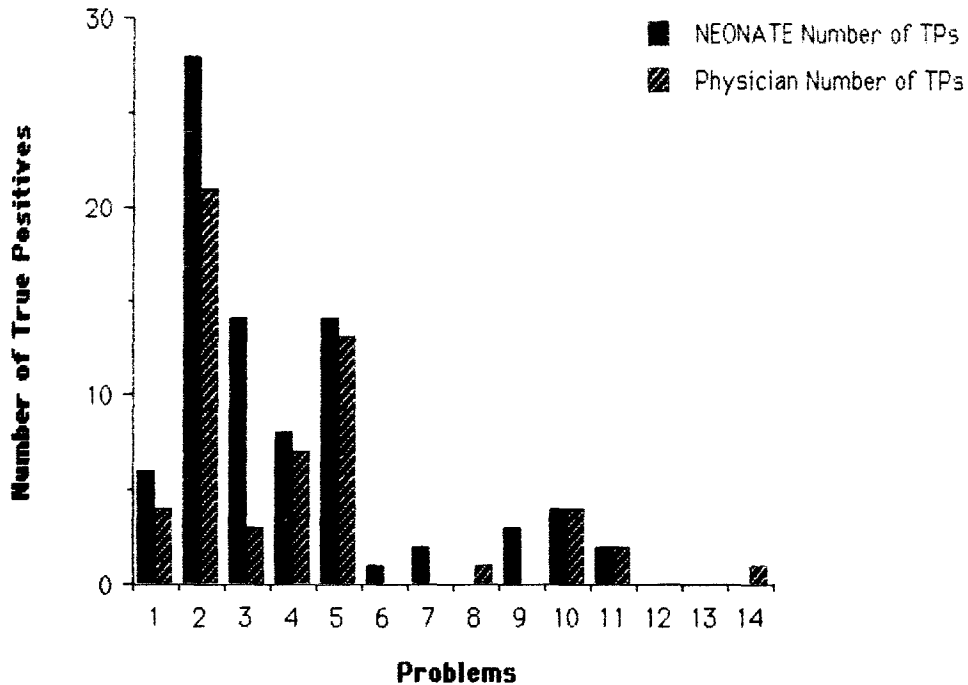


Figure 1. A histogram that displays the difference between the number of true positives of NEONATE and the Physician Problem List by problem.

Problems:

1. Transient Tachypnea of the Newborn
2. Sepsis
3. Infectious Pneumonia
4. Aspiration Pneumonia
5. Hyaline Membrane Disease
6. Metabolic Acidosis
7. Hypokalemia
8. Septic Shock
9. Anemia
10. Pneumothorax
11. CNS Depression
12. Subarachnoid Hemorrhage
13. Intraventricular Hemorrhage
14. Persistent Pulmonary Hypertension

is spread among several problems, contributes to the improvement of NEONATE over the physician problem list.

The significantly larger number of true-positives between NEONATE and the physician problem list probably results from several variables. First, and perhaps most importantly, it may be from inadequate documentation on the part of the admitting neonatologist. Common problems such as hypoglycemia are so frequently managed as part of routine care that they may be overlooked on the problem list even though there is laboratory proof, which the NEONATE system is designed to document. Second, the approach to some problems may not conform to a "standard." Choosing a hematocrit to

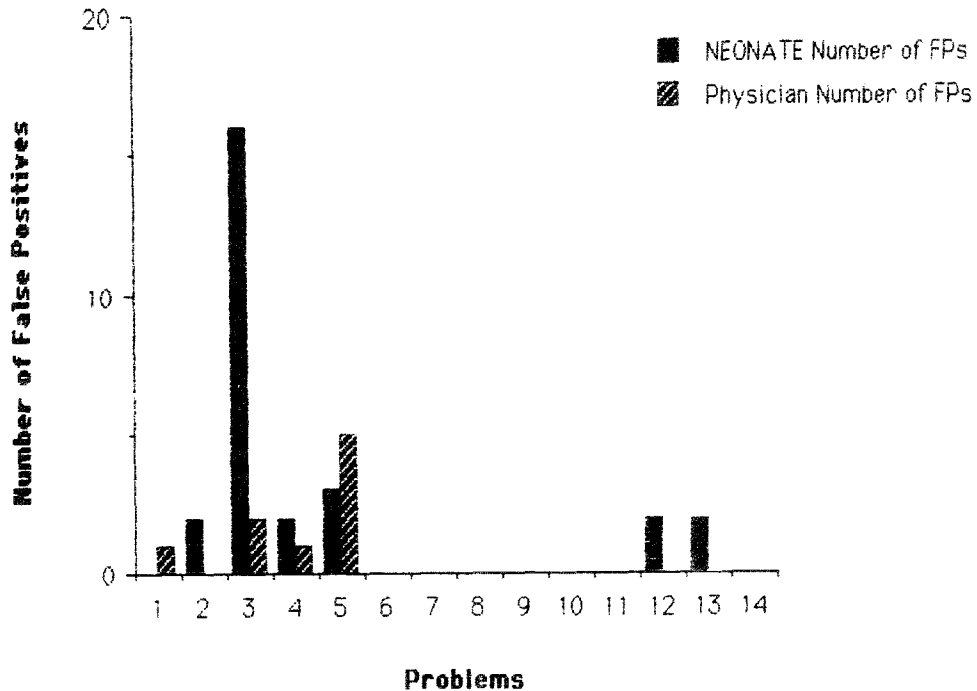


Figure 2. A histogram that displays the difference between the number of false positives of NEONATE and the Physician Problem List by problem.

Problems:

1. Transient Tachypnea of the Newborn
2. Sepsis
3. Infectious Pneumonia
4. Aspiration Pneumonia
5. Hyaline Membrane Disease
6. Metabolic Acidosis
7. Hypokalemia
8. Septic Shock
9. Anemia
10. Pneumothorax
11. CNS Depression
12. Subarachnoid Hemorrhage
13. Intraventricular Hemorrhage
14. Persistent Pulmonary Hypertension

define anemia is an example. The standard number is 40, below which a neonate is often transfused. However, in the setting of having an infant who is sick on a ventilator, one neonatologist may choose to transfuse when the hematocrit drops below 45, never to reach the "standard" number of 40. Another neonatologist may define anemia by whether a transfusion is needed. He or she may tolerate a hematocrit down to 35 without transfusing, and because this hematocrit is not reached, and no transfusion is given, anemia is not entered. Finally, neonates may often have one or more problems that are manifested by the same symptoms. For example, an infant may be delivered prematurely because of maternal infection and have sepsis with pneumonia, but because of the prematurity, also

have hyaline membrane disease. There are two diseases causing pulmonary disease. Often, the mothers are treated with antibiotics prior to delivery, neonatal blood cultures never grow, and sepsis and pneumonia are eventually dropped as a diagnoses even though the patient received a 7- to 10-day course of antibiotics, and may have indeed had both. NEONATE is designed to list both diagnoses, but some neonatologists may elect to list only the "proven" diagnosis of hyaline membrane disease. This may explain why sepsis and infectious pneumonia generate a particularly large number of differences between NEONATE and the physician problem list.

This study is supportive of the concept of embedding an expert system in a HIS. The CETUS system does not have any expert features in it. An expert system such as NEONATE might improve its output considerably. The study demonstrates the improvement of the documentation. Further studies in real clinical settings should be done to explore the implications of feedback from NEONATE on physicians' behavior.

REFERENCES

1. Warner, H.R., Olmsted, C.M., and Rutherford, B.D., HELP—A program for medical decision-making. *Comput. Biomed. Res.* 5:65-74, 1972.
2. Pryor, T.A., Gardner, R.M., Clayton, P.D., and Warner, H.R., The HELP system. *J. Med. Sys.* 7(2):87-101, 1983.
3. Pryor, T.A., The HELP medical record system. *MD Comput.* 5(5):22-33, 1988.
4. Janik, D.S., Swarner, O.W., Henriksen, K.M., and Wyman, M.L., A computerized single entry system for recording and reporting data on high-risk newborn infants. *J. Pediatr.* 93:519-523, 1978.
5. Janik, D.S., Swarner, O.W., Henriksen, K.M., and Wyman, M.L., Computerized newborn intensive care data recording and reporting II: An online system. *J. Pediatr.* 94:326-330, 1979.
6. Janik, D.S., Sharp, E.M., Forbush, L., Wyman, M.L., and Jung, A.L., Computerized newborn intensive care data recording, reporting and research. III. A practical microcomputer system. *J. Pediatr.* 97:497-500, 1980.