Factors That Determine the Cost and Performance of Early Identification Protocols

Robert G. Turner*

Abstract
In this paper, a simple model is used to examine the relation between important factors, such as prevalence of hearing loss, and protocol performance and cost. Protocol hit rate is determined primarily by the hit rate of the screening protocol and follow-up percentage, the percentage of infants who return for diagnostic testing. Many factors influence protocol cost. Not only does absolute cost change as certain model parameters are varied, but the relative cost of different protocols can change as well. The false alarm rate of the screening protocol and the follow-up percentage have the greatest impact on protocol cost.

Key Words: Hearing loss, hearing screening, early identification, infant, model, cost-benefit analysis

In the first paper of this series of three, a strategy was suggested for the evaluation and selection of early identification (EID) protocols (Turner, 1991). This strategy combines quantitative measures of protocol performance and cost with qualitative consideration of other important factors. A simple model was described that provides reasonable estimates of cost and performance. This approach should provide a more rigorous, defensible process for the design and selection of EID protocols.

In the second paper (Turner, 1992), the model was used to evaluate four EID protocols that are similar to the most common strategies used today. The model used base parameters that were derived from the clinical literature. For some of the parameters, however, the reported values varied significantly. The base parameters reflected the most typical value, or a value that was in the middle of the reported range. The base parameters were adequate for an initial comparison of EID protocols, but different parameter values may change the relative cost and performance of the protocols. There is a need to better understand this relation between model parameters and protocol characteristics. In this paper, we determine how factors, such as prevalence of hearing loss, impact on protocol cost and performance. The information provided in this paper should aid in the design and development of EID protocols and provide a better theoretical foundation for work in this area.

The four EID protocols considered in this paper have the same general design (Fig. 1). This design has been described in detail in Turner (1991). The protocols differ only in their screening strategies.

Protocol 1 (P1): This protocol consists of no screening tests. All infants are scheduled for diagnostic testing.

Protocol 2 (P2): This protocol consists of the high-risk register (HRR). Infants that fail the HRR are scheduled for diagnostic testing. Infants that pass are no longer followed. Use of the HRR does not require the physical presence of the infants; thus, all infants can be tested.

Protocol 3 (P3): This protocol consists of auditory brainstem response (ABR) screening. Some infants may be discharged before testing. These infants and infants that fail are scheduled for diagnostic testing.

Protocol 4 (P4): This protocol consists of the HRR combined in series-positive with the ABR. This means that infants who fail the HRR receive ABR testing. Those that fail the ABR screening are scheduled for diagnostic testing.

*University of California, San Francisco, San Francisco, California
Reprint requests: Robert G. Turner, 400 Parnassus, Room A706, UCSF, San Francisco, CA 94143-0430
Figure 1 Basic early identification protocol design. Also shown are the four specific screening protocols that are evaluated. The screening protocols consist of no screening test or the high-risk register (HRR) and/or auditory brainstem response (ABR) screening. When ABR screening is used, some infants cannot be tested before discharge. "Out" means that an infant is no longer tested or followed by the protocol.

This is similar to the screening protocol recommended by ASHA (ASHA Committee on Infant Hearing, 1989). Some of the infants that fail the HRR may be discharged before ABR screening.

CALCULATIONS FOR VARIABLE PARAMETERS

Several measures of protocol performance and cost will be used. The most basic measures of performance are protocol hit rate (HTp) and protocol false alarm rate (FAp). HTp is the percentage of hearing-impaired infants in the nursery who are identified by the protocol. FAp is the percentage of normal hearing infants in the nursery who are incorrectly called hearing impaired by the protocol. We will be concerned primarily with HTp, which indicates our success at identifying hearing loss. Because we have assumed the diagnostic protocol to be definitive (HT/FA = 100/0%), FAp will always equal zero (Turner, 1992).

We will restrict our calculations of cost to two basic measures. These are the financial cost of the EID protocol per infant in the nursery (CPIN) and the financial cost of the EID protocol per hearing-impaired infant identified (CPHL). These two measures reflect different aspects of protocol expense. CPIN is a measure of what it costs to implement the program; whereas, CPHL is a measure of cost-effectiveness.

Techniques for calculating protocol cost and performance are given in Turner (1991) and will not be described here. Calculations were made for a range of parameter values that are based on published data (see Turner, 1991). The base values and the range of parameter values are given in Table 1. It was not practical to consider every possible combination of all parameter values. Instead, protocol performance and cost were calculated for the specified range of values of one parameter while remaining parameters were held at base value.

Prevalence

Estimates of prevalence of hearing loss (Pr) vary in the literature, particularly for the ICN. Does prevalence have a major impact on the relative performance and cost of different EID protocols? Protocol performance is measured in terms of hit rate and false alarm. By definition, these two measures are independent of the number of hearing-impaired (IH) and normal-hearing (NH) infants that are tested. Variations in prevalence will have no effect on the calculations of protocol hit rate or false alarm rate.

Prevalence may have a significant impact on protocol costs. We will consider first the ICN. The two measures of cost, CPIN and CPHL, were calculated as prevalence varied from 1 to 7 percent. This range includes most reported
values for the ICN. Surprisingly, prevalence has minor effect on CPIN, at least within this prevalence range (Fig. 2). This means that the cost per infant and, thus, the total cost of the EID protocol is not significantly affected by prevalence of hearing loss in the nursery.

CPHL is a strong function of prevalence (Fig. 3). The higher the prevalence, the more IH infants available to be detected. The more IH infants identified, the lower the CPHL. Of greater interest is the result that the relative costs of the four protocols do not change significantly with prevalence. This is best illustrated by calculating the ratio of CPHL for the three more expensive protocols to the least expensive protocol (P4). This ratio changes little as prevalence is increased seven-fold (Fig. 4). For example, P1 is approximately 1.3 times as expensive as P4, for any reasonable prevalence. The results for the WBN are qualitatively similar to

Table 1 Parameters Used With Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Performance</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICN</td>
<td>WBN</td>
</tr>
<tr>
<td>Pr: Prevalence (%)</td>
<td>Base 3.0</td>
<td>0.1</td>
</tr>
<tr>
<td>CNT: Cannot Test Before Discharge (%)</td>
<td>Base 10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Range 0-30</td>
<td>0-30</td>
</tr>
<tr>
<td>HRR: High Risk Register</td>
<td>Base 95</td>
<td>60</td>
</tr>
<tr>
<td>Hit rate (%)</td>
<td>Base 65</td>
<td>10</td>
</tr>
<tr>
<td>False alarm rate (%)</td>
<td>Range 65-95</td>
<td></td>
</tr>
<tr>
<td>ABR: Auditory Brainstem Response Screening</td>
<td>Base 95</td>
<td>95</td>
</tr>
<tr>
<td>Hit rate (%)</td>
<td>Base 15</td>
<td>10</td>
</tr>
<tr>
<td>False alarm rate (%)</td>
<td>Range 5-25</td>
<td></td>
</tr>
<tr>
<td>Follow-up</td>
<td>Base 50</td>
<td>50</td>
</tr>
<tr>
<td>FU: Success (%)</td>
<td>Base 30-90</td>
<td>30-90</td>
</tr>
<tr>
<td>Diagnostic Protocol</td>
<td>Base 100</td>
<td>100</td>
</tr>
<tr>
<td>Hit rate (%)</td>
<td>Base 0</td>
<td>0</td>
</tr>
<tr>
<td>False alarm rate (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Correlation</td>
<td>Base Zero</td>
<td>Zero</td>
</tr>
</tbody>
</table>

Figure 2 Cost per infant in nursery (CPIN) versus prevalence of hearing loss (Pr) for the intensive care nursery (ICN). Relation between CPIN and Pr is qualitatively similar for the well baby nursery. Base value for P4 is 3 percent.

Figure 3 Cost per hearing-impaired infant identified (CPHL) versus prevalence of hearing loss (Pr) for the intensive care nursery (ICN). Relation between CPHL and Pr is qualitatively similar for the well baby nursery. Base value for Pr is 3 percent.
Figure 4 Ratio of CPHL (cost per hearing-impaired infant identified) for Protocols 1, 2, and 3 to that for Protocol 1. Note that the relative costs of the protocols change very little with prevalence (Pr). Results for the well baby nursery are similar. Base value for Pr is 3 percent.

those for the ICN; therefore, these data are not presented.

From these results, we can conclude that prevalence only affects the absolute value of CPHL. It has little influence on CPIN or the relative costs of the protocols.

Cannot Test

Experience has shown that it may not be possible to screen all infants with ABR before discharge. For protocols that use ABR screening (P3 and P4), some infants will be referred directly to the diagnostic protocol. The HRR does not require access to the infants and can be completed on all newborns. CNT is the percentage of infants who are discharged before ABR screening. What impact does this factor have on EID cost and performance. Calculations were made as CNT varied from 0 to 30 percent. This range is consistent with that reported in the literature; however, the information is very limited.

For the ICN, protocol hit rates for P3 and P4 increased only about 1 percent (e.g., 45 to 46% for P4) (Fig. 5). P1 and P2 did not vary because these do not include ABR screening. The two measures of cost, CPIN and CPHL, increase only slightly with increasing CNT (CPHL is shown in Fig. 6; CPIN is not shown). The results for the WBN are qualitatively similar to those for the ICN and are not shown.

These results indicate that the number of infants discharged before screening does not have a major effect on the performance or cost of protocols, at least for reasonable values of CNT. The value of CNT specified is not particularly critical for the comparison of protocols. For very large values of CNT, P3 would essentially become P1 and P4 would become P2 with corresponding performance and cost.

Screening Protocol

The hit rate of the screening protocol is one of several factors that directly determine the hit rate of the total protocol. There is a direct, linear relation between screening hit rate and HTp. If screening hit rate doubles, so will HTp. Also, screening hit rate sets an upper limit on protocol hit rate. For example, if screening hit rate equals 90 percent, then maximum possible HTp is 90 percent.
There are several other issues relating to screening to consider. The first concerns the HRR. In general, screening with the HRR will reduce the cost of identification. We would expect P2 (HRR) to cost less than P1 (No Tests) and P4 (HRR+ABR) to cost less than P3 (ABR). This, in fact, was the case for the base parameters (see Turner, 1992) where the HRR was modeled as having a hit rate of 95 percent and a false alarm rate of 65 percent in the ICN, corresponding to a failure rate (FR) of 66 percent. Failure rate is the percentage of infants who fail a particular test. Higher failure rates have been reported for the HRR. What happens to the relative costs of protocols with a higher failure rate? CPHL has been calculated for a failure rate of 65 to 90 percent (Fig. 7); hit rate has been maintained at 95 percent. False alarm rate was calculated from the failure rate, using a prevalence of 3 percent. The two protocols that use the HRR increase in cost as failure rate increases. At 75 to 85 percent, the protocols that use the HRR actually cost more than the corresponding protocols without the HRR. That is, P2(HRR) costs more than P1(No Tests), and P4(HRR+ABR) costs more than P3(ABR). In addition, P2 and P4 have lower hit rates than P1 and P3, respectively, because of the additional screening with the HRR. Normally, an additional screening test will reduce costs; however, when failure rate is sufficiently high, the use of the HRR will actually increase costs. This is not an issue in the WBN because the failure rate for the HRR is low.

Failure rates for ABR screening in the ICN vary significantly. Since false alarm rate is only slightly less than the failure rate, FA varies significantly as well. How does FA for ABR screening affect the relative costs of protocols? CPHL was calculated for a FA of 5 to 25 percent (Fig. 8). This is the approximate range of most reported values. HT was held constant at 95 percent. At the base value of 15 percent, the use of ABR screening reduces cost below that of the corresponding protocol without ABR. That is, P3(ABR) costs less than P1(No Tests) and P4(HRR+ABR) costs less than P2(HRR). This is to be expected since additional testing reduces false alarm rate which, in turn, decreases cost. As FA decreases, the cost of P3 and P4 decreases, improving the cost advantage of these protocols relative to P1 and P2. The cost of P1 and P2 does not change with ABR false alarm rate because ABR is not part of the screening protocol. As FA increases, the cost of P3 and P4 increases. At FA = 25%, there is little cost advantage to using ABR screening. P3 costs about the same as P1, and P4 costs the same as P2. In addition, P3 and P4 have lower hit rates than P1 and P2, respectively, because of the additional screening with ABR. When FA is high, the use of ABR in the screening protocol can increase cost and reduce hit rate. If a facility experiences a high failure rate with ABR screening, then screening with ABR may not be cost-effective.
Follow-Up

There is little information in the literature on follow-up success or the factors that affect follow-up. The indication is that follow-up success (FU) can vary significantly and is influenced by many factors. What is the impact of follow-up on protocol performance and cost? The situation is much like that with the hit rate of the screening protocol. There is a linear relation between FU and protocol hit rate with FU setting an upper limit on maximum possible performance. This relation is shown for the ICN in Figure 9, but is also true for the WBN. The differences in hit rate for the different protocols is small compared to the variation caused by follow-up. The relative performance of the protocols does not change with FU; P1 always has the best hit rate regardless of follow-up success.

Follow-up percentage also has a significant impact on protocol costs. With a higher FU, more infants receive diagnostic testing. This actually increases the total cost of the protocol and, therefore, CPIN in both the ICN (Fig. 10) and WBN (Fig. 11). On the other hand, CPHL will decrease with increasing FU (Figs. 12 and 13). This is because more IH infants will be identified with better follow-up, making the EID protocol more cost-effective. In addition to changing the absolute cost, variation in FU can also change the relative cost of different protocols. In the ICN (see Fig. 12), P1 has the highest CPHL and P4 the lowest for base parameters; yet, for small values of FU, P4 can actually cost more than P1.

Diagnostic Protocol

As with screening hit rate and follow-up success, protocol hit rate is linearly related to
diagnostic performance. Because we have assumed the diagnostic protocol to be definitive (HT = 100%), protocol hit rate is essentially determined by the screening hit rate and follow-up success.

Any increase in diagnostic cost will increase the total cost of a protocol. Thus, CPIN and CPHL will both increase with increasing diagnostic expense. Because diagnostic cost does not affect protocol hit rate, CPHL and CPIN will have the same relation to diagnostic expense. Therefore, only CPHL is shown for the ICN (Fig. 14) and WBN (Fig. 15).

Of more interest is the relative impact of diagnostic cost on protocol cost. In the ICN, the two protocols with high false alarm rates, P1 and P2, increase in cost at a much greater rate than P3 and P4. Near base value ($160), there is little difference in CPHL for the four protocols. At high diagnostic cost, P3 and P4 have a clear cost advantage. In the WBN, CPHL is relatively independent of diagnostic expense except for P1, which has a high false alarm rate. High diagnostic cost makes P1 significantly more expensive.

DISCUSSION

It should now be clear that the various factors associated with an EID program can interact in a complex way to determine the cost and performance of the protocol. While general principles can be identified, there are some circumstances that will produce surprising results. One example is with nursery screening. We screen to reduce the cost of identification. In general, more screening makes a protocol more cost-effective. This relation was evident in the calculations for the base parameters (Turner, 1992). Protocol P4, which uses two screening tests, is more cost-effective than P2 or P3, which use one test; P2 and P3 are more cost-effective than P1, which uses no screening tests. We have seen in this paper, however, several situations where more screening is not more cost-effective. For example, the relative costs of different screening protocols can vary significantly with follow-up success (see Figs. 10–13). When follow-up is poor, P1(No Tests) can actually be more cost-effective than P4 (HRR+ABR) in the ICN. While it is important to understand the general relations that exist between protocol parameters and protocol performance or cost, we must understand that these relations may not always be true. This is one reason why it is extremely difficult to recommend one EID strategy as optimum for all facilities. The EID strategy will depend on local experience with such factors as follow-up success, diagnostic cost, etc.

Performance and costs are the two most basic measures when evaluating an EID proto-
col. Let us consider the various factors and how they influence these measures. Two factors that have little influence are the percentage of infants discharged before testing (CNT) and prevalence (Pr). CNT has so little effect that this factor can be removed from the model. Somewhat surprisingly, prevalence is also of little consequence with one exception. CPHL is inversely related to prevalence. Higher prevalence means that more hearing-impaired infants will be identified, which results in greater cost-effectiveness and, therefore, a lower CPHL.

There are two factors that have a significant impact on the hit rate of the EID protocol. These are the hit rate of the screening protocol and the follow-up percentage. If we ignore the infants who cannot be screened before discharge, then protocol hit rate equals the screening hit rate times the follow-up percentage. We need not consider diagnostic hit rate because we have assumed the diagnostic protocol to be definitive. Thus, total protocol hit rate can be no better than screening hit rate or follow-up percentage.

The hit rate of the screening protocol is determined by the hit rates of the individual screening tests and how these tests are combined. The use of no screening test actually provides the highest hit rate because there is no opportunity for IH infants to be missed by screening. When several screening tests are used, they have traditionally been combined in series-positive, as in P4. Combining tests in this way reduces hit rate of the screening protocol and degrades the performance of the EID protocol. If the hit rates of the screening tests are high, however, the resultant hit rate will still be relatively high.

The second important factor is follow-up percentage. Even though this is often overlooked, follow-up has a powerful influence on both performance and cost. In practice, FU probably has more impact on EID performance than the design of the screening protocol. In many programs, far more infants will be lost to follow-up than will be missed by screening. It is important to optimize follow-up success; unfortunately, there is little information in the literature as to the factors that influence follow-up. Controlled studies are needed to evaluate the different variables that determine follow-up success.

Many factors influence protocol cost, but the most important are the false alarm rate of the screening protocol, follow-up percentage, and post-screening expense. Because the prevalence of hearing loss is small in both intensive care and well baby nurseries, the number of infants to be followed is determined primarily by the false alarm rate of the screening protocol. More infants followed means more protocol expense. Screening false alarm rate can be reduced by combining screening tests in series-positive. Of course, this also reduces hit rate, resulting in a trade-off between performance and cost.

Both follow-up percentage and post-screening expense have a powerful influence on absolute and relative protocol costs (see Figs. 10–15). From these results, we can identify an important general principle. If experience indicates good success following infants, then it is cost-effective to perform more screening before discharge (e.g., P4) to reduce screening false alarms. On the other hand, if follow-up is poor, then it is better to postpone most of the testing until the diagnostic evaluation (e.g., P1). Likewise, the greater the cost of follow-up and/or diagnostic testing, the more cost-effective it is to reduce false alarms by additional screening.

One final issue deserves consideration. This is the use of the HRR for hearing screening. We have already seen that there is a problem using the HRR in the WBN because of its poor hit rate (Turner, 1992). There are also potential problems in the ICN where failure rates of 20 to 90 percent have been reported for the HRR (see Turner, 1991, p. 198). We have seen that if failure rate is high, it may be more cost-effective to forego screening with the HRR (see Fig. 7). If failure rate is low, as has been reported for some facilities, we cannot assume a high hit rate. In the model we have used HT/FA = 95/65 percent for the HRR, but this corresponds to a failure rate of 66 percent. The HRR may only have a high hit rate when failure rate is high (i.e., when most infants fail the screening). Unfortunately, there is little information in the literature on the hit rate of the HRR when failure rate is low. If the HRR is to be used for hearing screening, then limitations and potential problems must be clearly understood.

**SUMMARY**

A model for early identification protocols was used to examine the relation between important factors, such as prevalence of hearing loss, and protocol performance and cost. Four protocols were considered that differed only in their hearing screening strategies. The four screening protocols consisted of (1) no
Factors That Determine Cost and Performance/Turner

screening tests; (2) high-risk register (HRR); (3) auditory brainstem response (ABR) screening; and (4) HRR + ABR (infants who fail the HRR receive the ABR). The protocols were evaluated separately for the intensive care nursery and the well baby nursery. Protocol hit rate is determined primarily by the hit rate of the screening protocol and follow-up percentage, the percentage of infants who return for diagnostic testing. Many factors influence protocol cost. Not only does absolute cost change as certain model parameters are varied, but the relative cost of different protocols can change as well. The false alarm rate of the screening protocol and the follow-up percentage have the greatest impact on protocol cost.

This work exploits a particular strategy for the evaluation and selection of protocols for the early identification of hearing loss. This strategy combines quantitative measures of protocol performance and cost with qualitative consideration of important factors that are not included in the quantitative model. With this approach, it is also possible to study the factors that determine protocol characteristics. Hopefully, this effort will provide a better theoretical foundation for the development of early identification protocols.


REFERENCES


