Investigaciones Económicas
ISSN: 0210-1521
ie@funep.es
Fundación SEPI
España

Herrero, Carmen; Moreno-Ternero, Juan D.
HOSPITAL COSTS AND SOCIAL COSTS: A CASE STUDY OF NEWBORN HEARING SCREENING
Investigaciones Económicas, vol. XXIX, núm. 1, enero, 2005, pp. 203-216
Fundación SEPI
Madrid, España

Available in: http://www.redalyc.org/articulo.oa?id=17329107
HOSPITAL COSTS AND SOCIAL COSTS: A CASE STUDY OF NEWBORN HEARING SCREENING

CARMEN HERRERO
Universidad de Alicante

JUAN D. MORENO-TERNERO
Yale University

The economic evaluation of health care programs can be carried out from two different perspectives: the hospital’s narrow perspective and the social perspective, which includes all costs. It is well known that, depending on the perspective, the economic evaluation may lead to discrepant recommendations. In this paper we present an example of this situation by reporting an economic evaluation of the screening procedures to detect congenital hearing impairment in newborns. We obtain that from the hospital’s perspective a targeted procedure, based on a previous high-risk criterion, is preferred, whereas from the social perspective a universal procedure is preferred.

Keywords: Economic evaluation, health care programs, cost-effectiveness analysis, QALYs, newborn hearing screening.

(JEL D61, I12, I18)

1. Introduction

Cost-effectiveness Analysis (CEA hereafter) has become the dominant method for economic evaluations of health care nowadays. Studies may be carried out from the hospital’s perspective (including only health-care costs and ignoring all others) or from the social perspective (including not only health-care costs but also the costs arising

We are most grateful to the editor, Samuel Bentolila, and two anonymous referees for detailed suggestions that substantially improved the paper. Thanks are also due to Manuela López-Azorín, Dunia López-Pintado, Ignacio Ortuño-Ortín and the audience at the Department of Public Health (University of Alicante) for helpful comments. Moreno-Ternero’s research was possible thanks to a post-doctoral fellowship from the Generalitat Valenciana. Financial support from the Ministerio de Ciencia y Tecnología, under project BEC2001-0535, from the Generalitat Valenciana, under project GRUPOS 03-086, and from the Fundación BBVA, under project 1-04X, is also gratefully acknowledged.
from resources consumed in other sectors). As we show in this paper, the choice of the perspective is, by no means, irrelevant, since they might lead to discrepant recommendations. More precisely, we study the cost-effectiveness of newborn hearing screening from both perspectives and show that conclusions differ depending on the one that we consider. This is consonant with recent reflections on CEA, which suggest that accounting for future costs in CEA may substantially alter both the absolute and relative CEA of medical interventions (cf. Johannesson and Meltzer, 1998, Meltzer and Johannesson, 1999).

Our study has policy implications for a problem that the Spanish health care authorities are actually facing. Recently, the Spanish Health Ministry has presented a protocol for the early detection of hearing impairment in newborns. The protocol’s goal is the implementation of a screening program to detect congenital hearing impairment, before the development of symptoms. The protocol is based mainly on arguments expressed in medical literature, which refers rather fleetingly to the economic aspects of the problem. The CEA that we present in this paper somehow fills that gap and reinforces the evidence in favor of the implementation of a screening program. Among other things, we argue that it would be more cost-effective to test every newborn rather than just those belonging to a group at risk.

2. Newborn hearing screening

Screening is traditionally defined as testing a population of asymptomatic individuals to identify precursors of a disease. The subjects who test positive are sent on for further evaluation in a subsequent diagnostic test to determine whether they do, in fact, have the disease. Individuals can be partitioned into four groups, according to whether they do or do not have the disease and whether their screening tests are positive or negative. Thus, there are four groups of individuals: true positives, individuals whom the screening correctly indicates to have the disease; false positives, those who do not have the disease but who have a positive screening test; false negatives, those who have the disease but are mistakenly cleared by the screening; and true negatives, those who do not have the disease and are correctly identified as such by the screening.

The cells in Table 1 show the number of individuals that fall into each of these groups when a total of $n$ individuals are considered and a screening test is implemented. We can compute how likely an individ-
ual would belong to each of the four groups by using characteristics of the population (prevalence) and of the detection ability of the screening test (sensitivity and specificity). Prevalence ($\rho$) is the probability of an individual in the population being impaired. The sensitivity of the screening test ($\pi_1$) is the conditional probability that an individual with the disease is positively detected by the test. The specificity of the test ($\pi_2$) is the conditional probability of an individual without the disease being correctly detected as negative in the test. Using these definitions, the probability of an individual being a true negative ($(1 - \rho)\pi_2$), a true positive ($\rho\pi_1$), a false positive ($(1 - \rho)(1 - \pi_2)$) and a false negative ($\rho(1 - \pi_1)$) can be easily expressed. The advantage of this way of writing the screening probabilities is that it makes easier to assess the implications of variations in the parameters $\rho$, $\pi_1$ or $\pi_2$ separately.

### Table 1

<table>
<thead>
<tr>
<th>Test</th>
<th>Status: Number of cases</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
<td>Total</td>
</tr>
<tr>
<td>Positive</td>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>$\alpha + \beta$</td>
</tr>
<tr>
<td>Negative</td>
<td>$\gamma$</td>
<td>$\delta$</td>
<td>$\gamma + \delta$</td>
</tr>
<tr>
<td>Total</td>
<td>$\alpha + \gamma$</td>
<td>$\beta + \delta$</td>
<td>$n = \alpha + \beta + \gamma + \delta$</td>
</tr>
</tbody>
</table>

| Prevalence of Disease | $\rho = \frac{\alpha + \gamma}{n}$ |

| Test characteristics | | |
|----------------------| | |
| Sensitivity          | $\pi_1 = \frac{\alpha}{\rho} - \gamma$ |
| Specificity          | $\pi_2 = \frac{\delta}{\rho - \gamma}$ |

Obviously, the best scenario is to have a screening test with both the sensitivity and specificity approaching unity. This is not usually feasible, and, in general, there is a trade-off between the sensitivity and specificity of any given screening test. Nonetheless, specificity can only be meaningfully said to be high relative to the prevalence of the disease. For instance, with a prevalence of 1 case in a million even a test with specificity of 0.99 will produce 10,000 false positives for each true positive. As we shall see later in the text, this feature will be important to drive our results.¹

¹We thank a referee for raising this point.
In this paper, we focus on newborn hearing screening procedures. The importance of identifying congenital hearing-loss as early as possible has long been recognized by the medical literature (see, for instance, the Year 2000 Position Statement of the Joint Committee on Infant Hearing and the literature cited therein). There is broad agreement in the literature on the adequacy of implementing newborn hearing screening programs, although the economic aspects of this decision have been mostly ignored. The choice of the appropriate screening procedure, however, is a source of controversy in the medical literature (see, for instance, Bess and Paradise, 1994, White and Maxon, 1995, Thompson et al., 2001, and the literature cited therein). Some authors recommend a universal newborn hearing screening (UNHS hereafter), while others prefer a targeted newborn hearing screening (TNHS hereafter) in which only those infants with high-risk factors are tested. Both options have their pros and their cons. On one hand, only 50% of newborns with a hearing impairment belong to a group at risk. On the other hand, the cost of testing all newborns is considerably higher.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Sensitivity ($\pi_1$)</th>
<th>Specificity ($\pi_2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNHS</td>
<td>0.860</td>
<td>0.973</td>
</tr>
<tr>
<td>TNHS</td>
<td>0.507</td>
<td>0.999</td>
</tr>
</tbody>
</table>

Table 2 summarizes the characteristics of each screening procedure, including baseline estimates and range of values. Data are based on three different studies (Kemper and Downs, 2000, Keren et al., 2002, Kezirian et al., 2001) each of them partially building on previous experiences with newborn hearing screening programs that were in practice. Ideally, we would have included additional information of these parameters like quartiles or standard errors. Unfortunately, the medical literature is silent in this respect. To ameliorate this shortcoming, and as we shall see later in the text, we run sensitivity analyses to test the robustness of our results.2

2The reader is referred to Herrero and Moreno-Ternero (2002), for additional information on the medical aspects of screening procedures from an economic viewpoint.
3. Costs

In this section, we describe in detail the different kinds of costs associated with newborn hearing screening programs. They can be divided into two categories: health-care costs and costs from other sectors. Health-care costs include the actual cost of the screening itself (i.e., machine costs, supplies and wages of the specialists), costs of the final diagnostic test (to which every impaired and false-positive infant is referred), the costs of the treatment and (possibly) the cost of detecting a risk factor.\(^3\) Formally,

\[ c = c_r + c_s + [\rho + r(1 - \rho)(1 - \pi_2)]c_d + \rho c_t, \]

where \(c_r\) denotes the cost of detecting a risk factor, \(c_s\) the cost of the screening itself, \(r\) the return rate, \(c_d\) the cost of the diagnostic test and \(c_t\) the cost of the treatment.\(^4\) It is worth noting that we assume that all impaired individuals receive diagnostic evaluation and treatment, regardless of whether their impairment is detected early or not. Consequently, the health-care cost in the status quo (SQ hereafter), i.e., the scenario without any screening protocol, is given by \(c_{sq} = \rho(c_t + c_d)\). Thus, the incremental health-care cost of implementing a screening program is

\[ c - c_{sq} = c_r + c_s + [r(1 - \rho)(1 - \pi_2)]c_d. \]

Costs from other sectors associated with a screening program refer to educational costs and production costs - which we estimate by the earnings gap between average non-deaf and deaf individuals. It is reasonable to assume that those impaired infants who are detected early will also receive early treatment, so that the effectiveness of the treatment is more likely than in the case of a late detection. Formally, let \(\tau\) denote the probability of a cure when the treatment is carried out after an early detection and \(\tau'\) when it is done after a late detection. We, therefore, assume \(\tau \geq \tau'\). Let \(c'\) denote the costs from other sectors after implementing a screening program. Then

\[ c' = \rho[\pi_1(1 - \tau) + (1 - \pi_1)(1 - \tau')](c_r + c_p), \]

\(^3\)From here on, unless otherwise stated, it is assumed that all costs are per capita.

\(^4\)By return rate we mean the percentage of infants returned for follow-up testing.
where $c_e$ denotes the incremental cost of special education and $c_p$ is the production cost. Similarly, if $c'_{sq}$ denotes the costs from other sectors in SQ, then $c'_{sq} = \rho(1 - \tau')(c_e + c_p)$, and therefore,

$$c' - c'_{sq} = -\rho\pi_1(\tau - \tau')(c_e + c_p).$$

That is to say, the incremental costs (savings) from other sectors, after implementing a screening program, take the form of a constant factor times the difference between the probabilities of a cure after early and late detection.

Table 3 shows the baseline estimation and range of values of the parameters in the above equations. All costs were adjusted to 2001 US dollars. Future costs were discounted at a rate of 3% per year, as recommended by the Panel on Cost-Effectiveness in Health and Medicine (cf., Gold et al., 1996). The data reported in this table concerning return rate, prevalence and direct costs are based on the studies mentioned above, whereas indirect costs are reported from Cheng et al. (2000). As before, the medical literature is far from being satisfactory in providing additional statistical aspects on these parameters and we run sensitivity analyses to support our results.5

### Table 3

<table>
<thead>
<tr>
<th>Parameters associated with congenital hearing impairment</th>
<th>Baseline estimate</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return rate ($r$)</td>
<td>0.75</td>
<td>[0.70, 1]</td>
</tr>
<tr>
<td>Prevalence ($p$)</td>
<td>0.0021</td>
<td>[0.001, 0.006]</td>
</tr>
<tr>
<td>Risk factor detection cost ($c_r$)</td>
<td>$1</td>
<td>[$0.5, $15]</td>
</tr>
<tr>
<td>Diagnostic test cost ($c_d$)</td>
<td>$280</td>
<td>[$150, $540]</td>
</tr>
<tr>
<td>Treatment cost ($c_t$)</td>
<td>$60228</td>
<td>[$31860, $99680]</td>
</tr>
<tr>
<td>Educational cost ($c_e$)</td>
<td>$63558</td>
<td>[$18370, $67330]</td>
</tr>
<tr>
<td>Production cost ($c_p$)</td>
<td>$148198</td>
<td>[$0, $343080]</td>
</tr>
</tbody>
</table>

4. Utility gains

The most frequently used method for CEA of health care programs is the Cost Utility Analysis (CUA hereafter). CUA measures the benefits of a program in utility terms. The most well-known utility index is

---

5 The reader is referred to Herrero and Moreno-Ternero (2002), for additional information on the estimation of costs.
the Quality Adjusted Life Years (QALYs hereafter). On the basis of the QALY index, preferences on alternative programs can be given by the cost-utility (or cost-per-QALY-gained) ratios. In particular, a standard CUA says that the program that should be implemented, from among a set of alternative ones, is the one that offers the largest number of QALYs per dollar or, what is equivalent, the one that has the lowest cost per QALY gained.

The QALY index assigns quality weights to each year provided with a health care program. To do so, one has to measure preferences for health outcomes, a rather time-consuming and complex task. In a recent paper, we provided certain conditions under which the outcome measurement of general screening programs can be drastically simplified. The following paragraphs summarize this.

Let $T$ be the expected number of QALYs enjoyed by an individual, after implementing a screening program. Then,

$$T = \rho \pi_1 Q_{TP} + (1 - \rho)(1 - \pi_2)Q_{FP} + \rho(1 - \pi_1)Q_{FN} + (1 - \rho)\pi_2 Q_{TN},$$

where $Q_{TP}$, $Q_{FP}$, $Q_{FN}$ and $Q_{TN}$ are the corresponding expected number of QALYs for an individual in each of the resulting four groups described in Section 2. Similarly, the expected number of QALYs to be enjoyed by an individual without implementing a screening program (at SQ) is given by $Q_{SQ} = \rho Q_i + (1 - \rho)Q_{IF}$, where $Q_i$ and $Q_{IF}$ are the expected number of QALYs for a newborn in the groups of impaired and impairment-free individuals, respectively. The assumptions are the following:

**Assumption 1** $Q_{TN} = Q_{IF}$, and $Q_{FN} = Q_i$

**Assumption 2** $Q_{TP} - Q_{FN} > 0$.

**Assumption 3** $Q_{FP} = Q_{TN}$.

In words, Assumption 1 says that utility does not decrease per se from being referred to a screening program, i.e., the QALYs of a true (false) negative individual after implementing a screening program coincide with the QALYs of a healthy (impaired) individual in SQ. Assumption 2 says that early detection of the disease is advantageous at an individual level. Assumption 3 says that there are no utility differences between healthy individuals with different test results.

The following result states that if the above assumptions are satisfied, then the expected QALYs gained over SQ due to the implementation of
a screening program equals its level of sensitivity, up to a multiplicative and non-negative constant factor that depends on the prevalence of the disease. Formally:

**Theorem 1.** (Herrero and Moreno-Ternero, 2003) If Assumption 1, Assumption 2 and Assumption 3 are satisfied, then there exists a constant $\lambda(\rho) \geq 0$ such that $Q - Q_{eq} = \lambda(\rho) \pi_1$.

Theorem 1 says, in particular, that the ‘cost-per-QALY-gained’ ratios and the so-called ‘cost-sensitivity’ ratios, i.e., the ratios between the cost and the sensitivity of each program, yield the same ranking of preferences among the alternative screening procedures, when the above assumptions are fulfilled. In the particular case of newborn hearing screening programs, the three assumptions are sound, given that the screening protocols are not invasive.\(^6\) Thus, we can use the ‘cost-sensitivity’ ratios as a proxy of the ‘cost-per-QALY-gained’ ratios of screening programs.

### 5. Results

In general, given any test characteristics, universal screening is desirable only when both the prevalence and the benefits of detection are high enough. The results reported here, although driven by these facts, depend crucially on the perspective we decide to assume in order to compute the costs.

#### 5.1 CUA from the hospital’s perspective

In this section, we provide the results of a CUA from the hospital’s perspective, i.e., a CUA in which only health-care costs associated with screening programs are computed. Table 4 shows health-care costs, QALYs gained and cost per QALY gained ratios of both TNHS and UNHS. We see that TNHS exhibits the lowest ratio and, therefore,

---

\(^6\)One could argue, however, that A3 is bypassing the hypothetical anxiety or depression provoked by a false positive test to patient’s parents. Nevertheless, recent papers have shown that this effect is not so important (cf. Clemens et al., 2000, Weichbold and Welzl-Mueller, 2001).
C. Herrero, J. D. Moreno-Ternero: Newborn Hearing Screening

is the one preferred according to a standard CUA from the hospital’s perspective.

The base-case estimate of cost per QALY gained was robust to changes in most of the key probabilities and costs. The only two variables to which the model was moderately sensitive were the specificity of TNHS and the cost of detecting a risk factor. More precisely, by decreasing the specificity of TNHS from its base-case estimate ($0 = 0.999$) to the lower bound of its range ($0 = 0.970$) the cost per QALY gained of TNHS changed from $5.723\lambda(\rho)$ to $17.709\lambda(\rho)$ which is a higher value than the base-case estimate of UNHS. Similarly, as the cost of detecting a risk factor varied from its mean estimation to $\$7 = 0.05$, the cost per QALY gained of TNHS changed from $5.723\lambda(\rho)$ to $17.656\lambda(\rho)$.

5.2 CUA from the social perspective

In this section we provide the results of a CUA from the social perspective in which not only health-care costs, but also costs from other sectors associated with screening programs are computed.

Table 5 shows the ranking of preferences among the three scenarios (SQ, TNHS and UNHS), according to their overall (health care plus non-health care) costs, depending on the range of probabilities of success.\(^7\) We can observe from this table that, if the probability of success for early treatment (due to early detection) is sufficiently larger than the probability of success of late treatment (due to late detection), then UNHS produces higher savings than TNHS. Furthermore, in this case, both programs are preferred to SQ, i.e., they both produce savings in lifetime costs with respect to SQ. The precise threshold is $0.077$.

\(^7\)Note that, if \(s\) and \(\bar{s}\) denote two alternative screening programs, \(s\) is preferred to \(\bar{s}\), which we write \(s \geq \bar{s}\) if and only if \((c - c_{eq}) + (\bar{c}' - \bar{c}_{eq}) \leq (\bar{c} - c_{eq}) + (\bar{c}' - \bar{c}_{eq})\). This is equivalent to the condition \(c - \bar{c} \leq \rho(\tau - \tau')(c_{eq} + c_{eq})(\pi_1 - \pi_1)\). Taking this into account, it is straightforward to obtain the range of probabilities that are shown in Table 5.
cure due to early detection is 7.7% higher than in late detection, then UNHS dominates TNHS and both dominate SQ. Unfortunately, the medical literature does not offer conclusive evidence on whether this is the case (e.g., Thomson et al., 2001). Nonetheless, although there would be some uncertainty about the precise magnitudes of the reductions in morbidity arising from earlier detection of congenital hearing impairment, the benefits are certainly positive (see, for instance, the Year 2000 Position Statement of the Joint Committee on Infant Hearing and the literature cited therein) and \( \tau - \tau' \geq 0.077 \) is a fairly plausible assumption.\(^8\)

<table>
<thead>
<tr>
<th>Range of probabilities</th>
<th>Ranking of programs</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0 \leq \tau - \tau' \leq 0.013 )</td>
<td>SQ ( \geq ) TNHS ( \geq ) UNHS</td>
</tr>
<tr>
<td>( 0.013 \leq \tau - \tau' \leq 0.039 )</td>
<td>TNHS ( \geq ) SQ ( \geq ) UNHS</td>
</tr>
<tr>
<td>( 0.039 \leq \tau - \tau' \leq 0.077 )</td>
<td>TNHS ( \geq ) UNHS ( \geq ) SQ</td>
</tr>
<tr>
<td>( 0.077 \leq \tau - \tau' \leq 1 )</td>
<td>UNHS ( \geq ) TNHS ( \geq ) SQ</td>
</tr>
</tbody>
</table>

This says that, in general, the ranking of screening programs according to their overall costs, coincides with the ranking of programs according to the gained QALYs that they offer and there is no need to resort to the standard cost-per-QALY-gained ratios to select among them.

With respect to the results reported in Table 5, the univariate sensitivity analyses indicate that the uncertainty over production costs is the only factor that has any influence. This influence, however, is not so substantial. For instance, in the most conservative assumption, that there was no effect at all of congenital hearing impairment on future earnings, i.e., \( c_p = 0 \), the corresponding thresholds in Table 5 would only increase to 0.042, 0.128 and 0.252, respectively. In other words, assuming a 25% of increase in the probability of a cure, when the impairment is detected early, is enough to obtain the same ranking of programs according to their overall costs as according to the gained QALYs that they offer.

To conclude with this section, a comment about the differences between the results from both perspectives is in order. It is well known

\(^8\)We asked some Spanish pediatricians to this respect and they agreed on the plausibility of the assumption.
that without a monetary valuation of the effectiveness unit (e.g., QALYs), a CEA does not say anything about whether a program should be implemented or not. Indeed, it only identifies dominated alternatives that should never be implemented. If all programs produce QALY gains, but also incur in positive costs, as happens from the hospital’s perspective, then the screening options cannot be directly compared to the status quo. Under the social perspective of this problem, however, programs can be compared with the status quo. Indeed, we see that, as we vary probabilities within a reasonable range, programs simultaneously offer savings and QALY gains with respect to the status quo. In other words, the social perspective allows us to show that the status quo is a dominated alternative by the implementation of a screening procedure, whereas it is not possible to infer this with the hospital’s narrow perspective.

6. Discussion

In this paper, we have presented a CEA of health care programs whose conclusions differ depending on the perspective we decide to fix. More precisely, if we compute health-care costs alone, i.e., the hospital’s perspective, a newborn screening program that only tests infants at high risk has a lower standard cost-effectiveness ratio than a universal program. If, however, other costs like special education and production costs are also computed, i.e., the social perspective, then the latter option produces higher health benefits and higher savings. In other words, from the social perspective, the targeted alternative based on a preliminary high-risk criterion would be dominated by the universal alternative. The problem of selecting a particular newborn hearing screening procedure is therefore an example of the importance of identifying the precise viewpoint under which a CEA is to be carried out.

Our approach assumes a collectively-financed health system, like the Spanish one. Motivated by this, we have excluded computing expenses that are covered individually -like travel expenses, special equipment and time off from (parent’s) work to carry the baby to the hospital for screening or for treatment after discharge. Likewise, and for ease of exposition, we have not contemplated differences in treatment and diagnostic costs due to discounting. Since false negative patients receive diagnosis and treatment at a later point in time than true positive ones, these differences might arise. Notwithstanding, we run parallel calculations which show that if these additional aspects were considered, the
evidence supporting a universal program from the social perspective would still be highly significant.

We conclude by arguing that the social perspective should be adopted in the problem described in this paper. The reason for doing so is that in Spain (or in Europe in general) all costs associated with congenital hearing impairment are borne by the Public Authority: health care (screening and treatments), education, and disability allowances (understood as disability insurance). This calls for the social perspective to pool all these costs together when analyzing the problem. Furthermore, computing costs from health-care resources consumed and costs derived from other sectors is consonant with the widely extended belief that the objective of CEA should be the identification of efficient uses of social resources for health care. Consequently, we recommend the implementation of a universal newborn hearing screening program in Spain.
References
Resumen

La evaluación económica de programas sanitarios puede llevarse a cabo desde dos perspectivas distintas: la del hospital y la social, que abarca todos los costes. Es bien sabido que, dependiendo de la perspectiva, la evaluación económica puede proporcionar recomendaciones discrepantes. En este artículo presentamos un interesante ejemplo al respecto, realizando una evaluación económica de los procedimientos de cribado para detectar disfunciones auditivas congénitas en recién nacidos. Nuestros resultados muestran que desde la perspectiva del hospital se prefiere un protocolo selectivo, basado en un criterio preliminar de grupos de riesgo, mientras que desde la perspectiva social se prefiere un protocolo universal.

Palabras clave: Evaluación económica, programas sanitarios, análisis coste-efectividad, AVACs, cribado auditivo en recién nacidos.