Methodology

Econometric estimation of country-specific hospital costs
Taghreed Adam*, David B Evans and Christopher JL Murray

Address: Global Programme on Evidence for Health Policy (GPE/EQC), World Health Organization, CH-1211 Geneva 27, Switzerland
Email: Taghreed Adam* - adamt@who.int; David B Evans - evansd@who.int; Christopher JL Murray - murrayc@who.int
* Corresponding author

Abstract
Information on the unit cost of inpatient and outpatient care is an essential element for costing, budgeting and economic-evaluation exercises. Many countries lack reliable estimates, however. WHO has recently undertaken an extensive effort to collect and collate data on the unit cost of hospitals and health centres from as many countries as possible; so far, data have been assembled from 49 countries, for various years during the period 1973–2000. The database covers a total of 2173 country-years of observations. Large gaps remain, however, particularly for developing countries. Although the long-term solution is that all countries perform their own costing studies, the question arises whether it is possible to predict unit costs for different countries in a standardized way for short-term use. The purpose of the work described in this paper, a modelling exercise, was to use the data collected across countries to predict unit costs in countries for which data are not yet available, with the appropriate uncertainty intervals.

The model presented here forms part of a series of models used to estimate unit costs for the WHO-CHOICE project. The methods and the results of the model, however, may be used to predict a number of different types of country-specific unit costs, depending on the purpose of the exercise. They may be used, for instance, to estimate the costs per bed-day at different capacity levels; the “hotel” component of cost per bed-day; or unit costs net of particular components such as drugs.

In addition to reporting estimates for selected countries, the paper shows that unit costs of hospitals vary within countries, sometimes by an order of magnitude. Basing cost-effectiveness studies or budgeting exercises on the results of a study of a single facility, or even a small group of facilities, is likely to be misleading.

Introduction
Information on hospital unit costs is valuable to health decision-makers and researchers for at least three purposes: budgeting (now receiving more attention with the availability of additional funds for health in poor countries through the Global Fund to Fight AIDS, Tuberculosis and Malaria); the assessment of hospital efficiency; and the assessment, by means of either cost-benefit or cost-effectiveness analysis, of the efficiency of different health interventions. Recognizing the need to make this information available on a country-specific basis, WHO has undertaken as part of the work programme WHO-CHOICE (CHOosing Interventions that are Cost-Effective – see http://www.who.int/evidence/cea), an extensive effort to collate all sources of data on unit costs from as many countries as possible [1]. Large gaps remain, however, particularly for developing countries. Although the long-term solution is that all countries perform their own
Cost Effectiveness and Resource Allocation 2003, 1

http://www.resource-allocation.com/content/1/1/3

Page 2 of 10

(page number not for citation purposes)
currency, and methods of allocation of joint costs. The database consists of unit-cost data from 49 countries for various years between 1973–2000, totalling 2173 country-years of observations. Some studies provided information on 100% of the variables described above; at the other extreme, some provided information on less than 15%. The number of observations used in this analysis was 1171 (see Additional file 1: 1 for the percentage of missing data in the model variables and Additional file 1: 1 for the list of countries).

Data cleaning comprised consistency checks and direct derivation of some of the missing variables, when possible, from other variables from the same observation (e.g., occupancy rate was calculated from number of beds and number of bed-days). STATA software was used for data analysis [55].

Cost data were converted to 1998 International dollars by means of GDP deflators [56] and purchasing-power-parity exchange rates used for WHO’s national health accounts estimates (PPP exchange rates used in this analysis are available from the WHO-CHOICE website: http://www.who.int/evidence/cea).

Data Imputation
Most statistical procedures rely on complete-data methods of analysis: computational programs require that all cases contain values for all variables to be analyzed. Thus, as default, most software programs exclude from the analysis observations with missing data on any of the variables (list-wise deletion). This can give rise to two problems: compromised analytical power, and estimation bias. The latter occurs, for example, if the probability that a particular value is missing is correlated with certain determinants. For example, if the complete observation sets tend to be from observations with unit costs that are systematically higher or lower than average, the conclusions for out-of-sample estimation drawn from an analysis based on list-wise deletion will be biased upwards or downwards [57].

There is a growing literature on how to deal with missing data in a way that does not require incomplete observation sets to be deleted, and several software programs have been developed for this purpose. If data are not missing in a systematic way, missing data can be imputed using the observed values for complete sets of observations as covariates for prediction purposes. Multiple imputation is an effective method for general-purpose handling of missing data in multivariate analysis; it allows subsequent analysis to take account of the level of uncertainty surrounding each imputed value, as described below [58–61]. The statistical model used for multiple imputation is the joint multivariate normal distribution. One of its main advantages is that it produces reliable estimates of standard errors: single imputation methods do not allow for the additional error introduced by imputation. In addition, the introduction of random error into the imputation process makes it possible to obtain largely unbiased estimates of all parameters [58].

In this study, multiple imputation was performed with Amelia, a statistical software program designed specifically for multiple imputation of missing data [57,59,62,63]. First, five completed-data sets are created by imputing the unobserved data five times, using five independent draws from an imputation model. The model is constructed to approximate the true distributional relationship between the unobserved data and the available information. This reduces potential bias due to systematic difference between the observed and the unobserved data. Second, five complete-data analyses are performed by treating each completed-data set as an actual complete-data set; this permits standard complete-data procedures and software to be utilized directly. Third, the results from the five complete-data analyses are combined [64] to obtain the so-called repeated-imputation inference, which takes into account the uncertainty in the imputed values.

Model specifications
From the tradition of using cost functions to explain observed variations in unit costs, we estimate a long-run cost-function by means of Ordinary Least Squares regression analysis (OLS); the dependent variable is the natural log of cost per bed-day [2,3,6–8,65]. The primary reason for using unit cost rather than total cost as the dependent variable is to avoid the higher error terms due to non-uniform variance (heteroscedasticity) in the estimated regression. This could arise if total cost were used as the dependent variable, as the error term could be correlated with hospital size [2,3]. The reason for using cost per bed-day rather than cost per admission is that “bed-days” are better than “admissions” as a proxy for such hospital services as nursing, accommodation and other “hotel services” [3], permitting more flexibility in the use of estimated unit costs.

As the relationship between unit costs and the explanatory variables is expected to be non-linear, the Cobb-Douglas transformation was used to approximate the normal distribution of the model variables. Natural logs were used. The Cobb-Douglas functional form can be written as follows:

\[ Y = \alpha_0 x_1^{\alpha_1} x_2^{\alpha_2} \quad \text{or} \]

Equation 1
Equation 2
\[\ln(Y) = \delta + \alpha_1 \ln(X_1) + \alpha_2 \ln(X_2)\]

where \(\delta = \ln(\alpha_0)\). This function is non-linear in the variables \(Y, X_1\) and \(X_2\), but it is linear in the parameters \(\delta, \alpha_1, \alpha_2\), and can be readily estimated using Ordinary Least Squares [66].

Log transformation has the added advantage that coefficients can be readily interpreted as elasticities [3, 66].

Therefore, the cost function specification of the OLS regression model may be written as:

Equation 3
\[UC_i = \alpha_0 + \sum_{i=1}^{n} \alpha_i X_i + e_i\]

Where \(UC_i\) is the natural log (ln) of cost per bed-day in 1998 1 $ in the \(i\)th hospital; \(X_1\) is ln of GDP per capita in 1998 1 $; \(X_2\) is ln of occupancy rate; \(X_3,4\) are dummy variables indicating the inclusion of drug or food costs (included = 1); \(X_5,6\) are dummy variables for hospital levels 1–2 (the comparator is level 3 hospital); \(X_7,8\) are dummy variables indicating facility ownership (comparator is private not-for-profit hospitals); \(X_9\) is a dummy variable controlling for USA data (USA = 1); and \(e\) denotes the error term.

The choice of explanatory variables is partly related to economic theory and partly determined by the purpose of the exercise, which is to estimate unit costs for countries where the data are not available. In this case, the chosen explanatory variables must be available in the out-of-sample countries. Country-specific – or in the case of large countries such as China, province-specific – GDP per capita in international dollars (1 $) is used as a proxy for level of technology [12–14]; occupancy rate as a proxy for level of capacity utilization; and hospital level as a proxy for case mix. Unit costs are expected to be correlated positively with GDP per capita and case mix and negatively with capacity utilization.

The inclusion of the seven control variables makes it possible to estimate unit cost for different purposes to suit different types of analysis – for example, cost per bed-day in a primary-level hospital, which does not provide drugs or food; or the cost in a tertiary level hospital, with drugs and food included.

The dummy for the USA was included because all data were charges rather than costs and because there were a large number of observations from that country. Dummies for countries other than the USA with a large number of observations, such as China and the United Kingdom, were also tested as was the use of dummy variables to capture whether the cost estimates included capital or ancillary costs. These variables were not included in the model which best fit the data. Utilization variables, such as number of bed-days or outpatient visits, and hospital indicators, such as average length of stay, were not included as explanatory variables because most out-of-sample countries do not have data on these variables, and prediction of unit costs would, therefore, be impossible.

Model-fit
Regression diagnostics were used to judge the goodness-of-fit of the model. They included the tolerance test for multicollinearity, its reciprocal variance inflation factors and estimates of adjusted R square and F statistics of the regression model.

Predicted values and uncertainty analysis
Two types of uncertainty arise from using statistical modes: estimation uncertainty arising from not knowing \(\beta\) and \(\alpha\) perfectly – an unavoidable consequence of having a finite number of observations; and fundamental uncertainty represented by the stochastic component as a result of unobservable factors that may influence the dependent variable but are not included in the explanatory variables [62]. To account for both types of uncertainty, statistical simulation was used to compute the quantities of interest, namely average cost per bed-day and the uncertainty around these estimates. Statistical simulation uses the logic of survey sampling to learn about any feature of the probability distribution of the quantities of interest, such as its mean or variance [62].

It does so in two steps. First, simulated parameter values are obtained by drawing random values from the data set to obtain a new value of the parameter estimate. This is repeated 1000 times. Then the mean, standard deviation, and 95% confidence interval around the parameter estimates are computed. Second, simulated predicted values of (the quantity of interest) are calculated, as follows: (1) one value is set for each explanatory variable; (2) taking the simulated coefficients from the previous step, the systematic component (\(g\)) of the statistical model is estimated, where \(g = f(X, B)\); (3) the predicted value is simulated by taking a random draw from the systematic component of the statistical model; (4) these steps are repeated 1000 times to produce 1000 predicted values, thus approximating the entire probability distribution of. From these simulations, the mean predicted value, standard deviation, and 95% confidence interval around the predicted values are computed. In this way, this analysis accounts for both fundamental and parameter uncertainty.
The predicted log of cost per bed day, ln \( \bar{U}C \), can then be calculated from:

**Equation 4**

\[
\ln \bar{U}C = \alpha_0 + \alpha_1 \ln X_1 + \sum_{i=1}^{n} \alpha_i X_i
\]

where \( \alpha_0 \) and \( \alpha_i \) are the estimated parameters, and \( X_{i,n} \) are the independent variables.

If \( \bar{\beta}_0 = \text{anti log} (\alpha_0) \) and \( \bar{\beta}_1 = \alpha_1 \), back-transforming Equation 4 (reduced to 1 independent log-transformed variable for simplicity) gives the power function.

**Equation 5**

\[
\bar{U}C_{\text{biased}} = \bar{\beta}_0 \bar{\beta}_1^{X_i}
\]

where \( \bar{U}C_{\text{biased}} \) denotes a biased estimate of the mean cost per bed-day due to back-transformation. This is because one of the implicit assumptions of using log-transformed models is that the least-squares regression residuals in the transformed space are normally distributed. In this case, back-transforming to estimate unit costs gives the median and not the mean. To estimate the mean it is necessary to use a bias correction technique. The smearing method described by Duan (1983) was used to correct for the back-transformation bias [67]. The smearing method is non-parametric, since it does not require the regression errors to have any specified distribution (e.g., normality). If the \( n \) residuals in log space are denoted by \( r_i \), and \( b \) is the base of logarithm used, the smearing correction factor, \( \bar{C}_{\text{bias}} \), for the logarithmic transformation is given by:

**Equation 6**

\[
\bar{C}_{\text{bias}} = \frac{1}{n} \sum_{i=1}^{n} b^{r_i}
\]

Multiplying the right side of Equation 5 by Equation 6 almost removes the bias, so that:

**Equation 7**

\[
\bar{U}C = \bar{C}_{\text{bias}} \bar{\beta}_0 \bar{\beta}_1^{X_i}
\]

The smearing correction factor (\( \bar{C}_{\text{bias}} \)) for our model was 1.25.

**Results**

Table 1 shows the variable names, description, mean and standard error, estimated after combining the results of the five datasets of the multiple imputation estimates. Table 2 presents the results of the best-fit regression model. The adjusted R square of the combined regressions is 0.80, with an F statistic of 509 (\( p < 0.0001 \)), indicating that the model explains a large part of the variation of the cost per bed-day across countries [68]. The signs of the coefficients are consistent with the earlier hypotheses. For example, the GDP per capita is positively correlated with cost per bed-day, while the lower the occupancy rate the higher is the cost per bed-day. Unit costs are lower in level-one hospitals than in those of levels two and three. The coefficients for the two main explanatory variables (GDP per capita and occupancy rate) are highly significant (\( p < 0.0001 \)), as well as most of the control dummies, e.g., hospital level. The coefficient for food costs is not significant at the 5% level but was included in the model because it added to its explanatory power.

The tolerance test and its reciprocal variance inflation factors (VIF) showed no evidence for multicollinearity between the model variables (tolerance ranged between 0.20 and 0.89, mean VIF 1.97; tolerance less than 0.05 and VIF more than 20 indicate the presence of multicollinearity).

The only country dummy that was included in the final model was for the USA. The most plausible explanation for the positive, highly significant coefficient for the USA dummy is that USA was the only large data set where charges were reported rather than costs. In this case, the coefficient for the USA could be interpreted as a cost-to-charge ratio, estimated as 1:1.74. In other words, costs represent 57% of the charge on average. This is consistent with published national reports on the average cost-to-charge ratio for the USA such as that published by the United States General Accounting Office (63%) [69].

Figure 1 shows the three regression lines of levels one, two and three hospitals, respectively, plotted against the log of GDP per capita (the Y-axis is log of cost per bed-day). The regression lines were estimated for public hospitals, with occupancy rate of 80%, including food costs and excluding drugs. Because the original data had a lower average occupancy rate (mean 71%, SD 39%), and most observations included drug costs, it is to be expected that the regression lines will be slightly lower than the actual data points in the database. The regression lines do not pass through the USA data points situated at the upper right side of the graph because they have been calculated for the case where the US dummy was set at zero.

Overall, Figure 1 shows that the regression lines have a good fit with the data used to develop the model. They not only illustrate the relationship between cost per bed-day, hospital level and GDP per capita, but also show that
there remains substantial variation in unit costs for any given level of GDP per capita. It would be inadvisable, therefore, to base cost estimates on a single estimate of hospital costs in a particular setting, something that is a common feature of cost-effectiveness studies.

To use the equation reported in Table 2 to predict unit costs for a number of in and out-of-sample countries, with the appropriate uncertainty interval, requires consideration of the probability distributions of the predicted unit costs, given a specified level of the model variables. In order to derive these distributions, simulation techniques were used following the steps described in the Methods section. Table 3 presents for selected countries in different regions of the world the average simulated predicted values and 95% uncertainty intervals. The estimates are presented in 2000 I $, based on the 2000 GDP per capita in I $ and assuming that the estimated coefficients will remain constant over a short time period. They are specific to public hospitals, at an occupancy rate of 80%, excluding drug, but including food costs. Regional estimates of cost per bed day, with the same characteristics described above, are available from the WHO-CHOICE website: http://www.who.int\evidence\cea.

**Discussion**

This paper describes recent work on developing models to predict country-specific hospital unit costs, by level of hospital and ownership, for countries where these data are not available. The main purpose of this work was to feed into estimates of the costs and effects of many types of health interventions in different settings. Estimates are typically available for variables such as the number of days in hospital, or the number of outpatient visits, for certain types of interventions, but unit prices are not available for many countries. The model presented in this paper used all data on unit costs that could be collected after a thorough search to estimate costs for countries where this information does not exist. Data imputation techniques were used to impute missing data, which has the advantage of eliminating the bias introduced by list-wise

---

**Table 1: Descriptive statistics of the multiple imputation estimates N = 1171**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Mean</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln cost per bed day</td>
<td>Natural log of cost per bed day in 1998 I $</td>
<td>4.98</td>
<td>1.63</td>
</tr>
<tr>
<td>Ln GDP per capita</td>
<td>Natural log of GDP per capita in 1998 I $</td>
<td>8.90</td>
<td>1.06</td>
</tr>
<tr>
<td>Ln occupancy rate</td>
<td>Natural log of occupancy rate</td>
<td>-0.41</td>
<td>0.61</td>
</tr>
<tr>
<td>Drug costs</td>
<td>Dummy variable for inclusion of drug costs. Included = 1</td>
<td>0.96</td>
<td>0.18</td>
</tr>
<tr>
<td>Food costs</td>
<td>Dummy variable for inclusion of food costs. Included = 1</td>
<td>0.93</td>
<td>0.25</td>
</tr>
<tr>
<td>Level 1 hospital</td>
<td>Dummy variable for level 1 hospital (1)</td>
<td>0.33</td>
<td>0.47</td>
</tr>
<tr>
<td>Level 2 hospital</td>
<td>Dummy variable for level 2 hospital</td>
<td>0.41</td>
<td>0.49</td>
</tr>
<tr>
<td>Public</td>
<td>Dummy variable for level public hospitals (2)</td>
<td>0.84</td>
<td>0.36</td>
</tr>
<tr>
<td>Private for profit</td>
<td>Dummy variable for level private for profit hospitals</td>
<td>0.08</td>
<td>0.27</td>
</tr>
<tr>
<td>USA</td>
<td>Dummy variable for USA. USA = 1</td>
<td>0.17</td>
<td>0.37</td>
</tr>
</tbody>
</table>

(1) Dummies for levels of hospital are compared with level 3 hospitals (2) Dummies for hospital ownership are compared with public not-for-profit hospitals

**Table 2: Multiple Imputation regression coefficients and SE Dependent variable: Natural log of cost per bed-day in 1998 I $ N: 1171**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>SE</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln GDP per capita</td>
<td>0.7624</td>
<td>0.0295</td>
<td>25.813</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Ln occupancy rate</td>
<td>-0.2318</td>
<td>0.0474</td>
<td>-4.886</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Drug costs</td>
<td>0.6410</td>
<td>0.1769</td>
<td>3.624</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Food costs</td>
<td>0.2116</td>
<td>0.1394</td>
<td>1.518</td>
<td>0.152</td>
</tr>
<tr>
<td>Level 1 hospital</td>
<td>-0.5777</td>
<td>0.0742</td>
<td>-7.787</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Level 2 hospital</td>
<td>-0.3118</td>
<td>0.0594</td>
<td>-5.253</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Public</td>
<td>-0.2722</td>
<td>0.1172</td>
<td>-2.323</td>
<td>0.021</td>
</tr>
<tr>
<td>Private for profit</td>
<td>0.2444</td>
<td>0.1316</td>
<td>1.857</td>
<td>0.064</td>
</tr>
<tr>
<td>USA</td>
<td>1.7471</td>
<td>0.1022</td>
<td>17.104</td>
<td>0.000</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.5036</td>
<td>0.3264</td>
<td>-7.672</td>
<td>0.026</td>
</tr>
</tbody>
</table>
deletion of observations in cases where information for some of the variables required by the model is missing.

The goodness-of-fit of the model was tested by various regression diagnostic techniques including the tolerance test for multicollinearity, adjusted R square and F statistic. All suggested a good fit of the model with the data and that GDP per capita could be used to capture different levels of technology use across countries. Although this is the first time that costs have been compared across countries, the signs of the coefficients are consistent with results from previous microeconomic studies within countries. For example, these studies have found that occupancy rate was negatively correlated with cost per bed-day while hospital level had the opposite relationship, something also found in the model presented in this paper [70,71]. This adds confidence to the estimated results.

In addition, the estimates produced by this model were sent to health economists and researchers in different countries to check their face validity. Experts from countries in all WHO regions, covering wide differences in GDP per capita and in technologies typically found in hospitals were consulted, including Benin, Canada, Ecuador, Egypt, Kenya, Netherlands and Thailand. They were provided with a description of the estimated unit cost (e.g., which costs were included) and were asked whether they thought they approximated the average cost per bed-day in their countries. All indicated that the results had face validity.

It is of particular note that the model incorporates a more extensive database on unit costs by hospital level and ownership than has previously been available. Increasing the range of observations will increase the validity of ex-

---

**Figure 1**

Regression lines for level one, two and three hospitals against the natural log of GDP per capita. (The Y-axis is the dependent variable: natural log of cost per bed day) Cost in 1998 Int $ N = 1171
trapolations of cost estimates for countries in which these data are not available. Additional sources of data are being sought for this purpose and to assist countries to develop their own studies. As this body of information grows, the predictive power of unit-cost models will continue to increase.

There are other possible uses of this model such as estimating the possible costs of scaling-up health interventions for the poor, which is receiving increasing attention with the activities of such bodies as the Global Fund to Fight AIDS, Tuberculosis and Malaria. This can be done in many ways, according to the objectives of the analysis. It may be used, for instance, to estimate:

- unit costs at different capacity levels for purposes of efficiency analysis or economic evaluation of health interventions;
- the "hotel" component of average cost per bed-day;

Table 3: Predicted cost per bed-day in 2000 I$

<table>
<thead>
<tr>
<th>Country</th>
<th>GDP per capita (I$)</th>
<th>In or out-of-sample</th>
<th>Hospital level</th>
<th>Mean (I$)</th>
<th>95% uncertainty interval Low</th>
<th>95% uncertainty interval High</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mali</td>
<td>581</td>
<td>Out</td>
<td>I</td>
<td>7.39</td>
<td>5.46</td>
<td>9.80</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>II</td>
<td>9.64</td>
<td>7.07</td>
<td>12.73</td>
<td>1.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>III</td>
<td>13.14</td>
<td>9.58</td>
<td>17.70</td>
<td>2.44</td>
</tr>
<tr>
<td>Mozambique</td>
<td>720</td>
<td>Out</td>
<td>I</td>
<td>8.70</td>
<td>6.45</td>
<td>11.43</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>II</td>
<td>11.35</td>
<td>8.32</td>
<td>15.01</td>
<td>2.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>III</td>
<td>15.46</td>
<td>11.31</td>
<td>20.42</td>
<td>2.85</td>
</tr>
<tr>
<td>Algeria</td>
<td>1,449</td>
<td>Out</td>
<td>I</td>
<td>14.82</td>
<td>10.95</td>
<td>19.45</td>
<td>2.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>II</td>
<td>19.35</td>
<td>14.26</td>
<td>25.45</td>
<td>3.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>III</td>
<td>26.34</td>
<td>19.41</td>
<td>34.63</td>
<td>4.72</td>
</tr>
<tr>
<td>Indonesia</td>
<td>3,167</td>
<td>Out</td>
<td>I</td>
<td>26.90</td>
<td>19.86</td>
<td>35.19</td>
<td>4.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>II</td>
<td>35.12</td>
<td>26.05</td>
<td>46.12</td>
<td>6.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>III</td>
<td>47.80</td>
<td>35.38</td>
<td>63.28</td>
<td>8.44</td>
</tr>
<tr>
<td>Ecuador</td>
<td>3,260</td>
<td>In</td>
<td>I</td>
<td>27.50</td>
<td>20.30</td>
<td>35.95</td>
<td>4.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>II</td>
<td>35.90</td>
<td>26.63</td>
<td>47.17</td>
<td>6.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>III</td>
<td>48.87</td>
<td>36.17</td>
<td>64.71</td>
<td>8.63</td>
</tr>
<tr>
<td>Romania</td>
<td>3,634</td>
<td>Out</td>
<td>I</td>
<td>29.88</td>
<td>22.05</td>
<td>39.11</td>
<td>5.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>II</td>
<td>39.00</td>
<td>28.86</td>
<td>51.25</td>
<td>6.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>III</td>
<td>53.09</td>
<td>39.23</td>
<td>70.35</td>
<td>9.37</td>
</tr>
<tr>
<td>Greece</td>
<td>6,192</td>
<td>Out</td>
<td>I</td>
<td>44.88</td>
<td>33.02</td>
<td>59.23</td>
<td>8.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>II</td>
<td>58.58</td>
<td>43.25</td>
<td>76.84</td>
<td>10.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>III</td>
<td>79.73</td>
<td>59.01</td>
<td>104.62</td>
<td>14.10</td>
</tr>
<tr>
<td>Russian Federation</td>
<td>8,035</td>
<td>In</td>
<td>I</td>
<td>54.75</td>
<td>40.17</td>
<td>72.43</td>
<td>9.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>II</td>
<td>71.48</td>
<td>52.73</td>
<td>94.24</td>
<td>12.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>III</td>
<td>97.27</td>
<td>71.79</td>
<td>127.92</td>
<td>17.27</td>
</tr>
<tr>
<td>Bahrain</td>
<td>14,159</td>
<td>Out</td>
<td>I</td>
<td>84.41</td>
<td>61.48</td>
<td>112.10</td>
<td>15.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>II</td>
<td>110.19</td>
<td>80.27</td>
<td>146.22</td>
<td>19.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>III</td>
<td>149.93</td>
<td>110.19</td>
<td>198.97</td>
<td>27.01</td>
</tr>
<tr>
<td>United Arab Emirates</td>
<td>20,330</td>
<td>Out</td>
<td>I</td>
<td>111.30</td>
<td>80.27</td>
<td>148.59</td>
<td>20.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>II</td>
<td>145.28</td>
<td>105.96</td>
<td>191.58</td>
<td>26.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>III</td>
<td>197.66</td>
<td>144.13</td>
<td>263.01</td>
<td>36.09</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>24,348</td>
<td>In</td>
<td>I</td>
<td>127.76</td>
<td>91.83</td>
<td>170.79</td>
<td>24.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>II</td>
<td>166.77</td>
<td>121.08</td>
<td>221.00</td>
<td>30.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>III</td>
<td>226.87</td>
<td>164.90</td>
<td>302.91</td>
<td>41.76</td>
</tr>
<tr>
<td>Canada</td>
<td>28,087</td>
<td>Out</td>
<td>I</td>
<td>142.51</td>
<td>102.03</td>
<td>190.22</td>
<td>27.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>II</td>
<td>186.02</td>
<td>134.55</td>
<td>247.23</td>
<td>34.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>III</td>
<td>253.05</td>
<td>183.16</td>
<td>337.26</td>
<td>46.89</td>
</tr>
</tbody>
</table>

(i) Cost per bed day is estimated for public hospitals with 80% occupancy rate, excluding drug costs and including food costs.
- unit costs, excluding specific items such as drugs or food costs.

Finally, it must be emphasized that there is wide variation in the unit costs estimated from studies within a particular country (Figure 1). These differences are sometimes of an order of magnitude, and cannot always be attributed to different methods. This implies that analysts cannot simply take the cost estimates from a single study in a country to guide their assessment of the cost-effectiveness of interventions, or the costs of scaling-up. In some cases, they could be wrong by an order of magnitude.

Conflict of Interest
None.

Authors’ contributions
TA was responsible for data collection, management and analysis, participated in the development of the methodology and drafted the manuscript. DE contributed to the development of the methodology, as well as data analysis and reporting. CM participated in the development and coordination of the methodology. All authors read and approved the final manuscript.

Additional material

Additional File 1
Annex 1: Definition of facility types as coded in the unit cost database. Annex 2: Percentage of missing data in the model variables prior to data imputation. Annex 3: Countries and number of unit cost observations included in the model.
Click here for file [http://www.biomedcentral.com/content/supplementary/1478-7547-1-3-S1.doc]

Acknowledgements
The authors express their gratitude to Carolyn Kakundwa and Margaret Squadrani for their work in compiling and processing the data necessary for this exercise; to Bian Ying, Viroj Tangcharoensathien, Walaiporn Patcharanarumol, Jiangbo Bao, Aparnaa Somanath, Elena Potapchik and Ruth Lucio for their efforts in gathering data at the country level; Ajay Tandon, Ke Xu and Gary King for their input in the development of the methods used; and to Xingzhu Liu and Jan Oostenbrink for the valuable comments during the review process. This work represents the views of the authors and not necessarily those of the organization they represent.

References
27. Shepard DS, Carrin G and Nyandagazi P Household participation in financing of health care in government health centres in Rwanda, In: Health economics research in developing countries (Edited by: Lee K, Mills A) Oxford University Press 1990, 140-164
44. Chan S Unit cost estimation for outpatient and inpatient departments in Nakleoun District Hospital, Cambodia. Bangko, Chulalongkorn University 1997.

Publish with BioMed Central and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp