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 (**CHO**osing **I**nterventions that are **C**ost-**E**ffective – 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

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 is to use the data collected across countries to predict unit costs in countries for which data are not yet available (both point estimates and uncertainty intervals are reported).

Health economics has a long tradition of estimating hospital-cost functions econometrically [2–10]. Econometric models explain how total costs change in response to differences in service mix, inputs, input prices, and scale of operations. They allow cost and production functions to be specified with sufficient flexibility that a non-linear relationship can be demonstrated between costs and quantity of inputs: total costs can rise at a lower rate than prices[2].

Previous studies have commonly used microeconomic data to analyse and estimate hospital-cost functions. This literature indicates two main approaches: behavioural cost functions and cost minimization functions [2,3,9,11]. Behavioural cost functions have been used to explain the variations in cost per unit of output among hospitals. They have used as determinants all variables for which a causal relationship to hospital costs is hypothesized and data are available – e.g., bed size, global indicators of hospital activity such as average length of stay and occupancy rate, dummy variables for teaching status, etc. On the other hand, the literature on cost minimization has described the minimum cost of providing a given volume of output as a function of an exogenous vector of input prices and the volume of output. The purpose is to determine whether hospitals are cost-minimizers (profit maximizers).

When testing the hypothesis of cost-minimization, the explanatory variables typically comprise only of output quantities (e.g., number of bed days) and input prices. The remaining variables used in the behavioural cost function specification are not part of the cost minimization question but can be used to explain deviation of observed unit costs from the theoretical minimum functions – e.g., possible reasons for inefficiency [3].

To our knowledge, all previous studies have used within-country data sets; we know of none that has attempted to estimate hospital-cost functions across countries. Such studies require a large number of observations from as many countries as possible.

The model described here follows the tradition of the behavioural cost function literature because its purpose is to estimate country-specific costs per bed-day, not to test the hypothesis of cost-minimization. The analysis controls for

across-country price-level differences by using unit costs adjusted for purchasing-power parity, namely in international dollars; and for differences in quantity and complexity of resource use by using macro-level indicators such as per capita GDP [12–14]. The model forms part of a series of models that can be used to predict country-specific unit costs for a number of purposes. They may be used, for instance, to estimate: (i) unit costs at different capacity levels for the purposes of efficiency analysis or economic evaluation of health interventions; (ii) the "hotel" component of average cost per bed-day for budgeting exercises; or (iii) unit costs excluding components that might be funded from other sources, such as drugs. The specific objectives of this paper are to:

• explain the observed differences in hospital inpatient cost per bed-day across and within countries; and

• use the results to predict cost per bed-day for countries for which these data are not yet available.

Methods
Data
The search sources used to obtain the data were: Medline, Econlit, Social Science Citation Index, regional Index Medicus, Eldis (for developing-country data), Commonwealth Agricultural Bureau (CAB), and the British Library for Development Studies Databases. The range of years was set at 1960 to the present. Data covering costs and charges were included.

The search terms used were: "costs and cost analysis" and hospital costs or health centre or the abbreviations HC (health centre) or PHC (primary health centre) or outpatient care. The language sources searched were English, French, Spanish and Arabic; no Arabic study was found. In addition, a number of studies were found in the grey literature, from such sources as electronic databases, government regulatory bodies, research institutions, and individual health economists known to the authors [2,15–54]. Also included were data from a number of WHO-commissioned studies on unit costs.

A standard template was used for extracting data from all sources. Database variables include: ownership; level of facility (see Additional file 1: Annex 1 for a definition of facility types as coded in the unit cost database); number of beds; number of inpatient and outpatient specialties; cost data (cost per bed-day, outpatient visit, and admission); utilization data (bed-days, outpatient visits, admissions); types of cost included in the cost analysis (capital, drugs, ancillary, food) and whether they were based on costs or charges; capacity utilization (occupancy rate, average length of stay, bed turnover, and average number of visits per doctor per day); reference year for cost data;

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 are 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:

Equation 1

$$Y = \alpha_0 X_1^{\alpha_1} X_2^{\alpha_2} \quad \text{or,}$$

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 δ, α_1, α_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 I \$ in the *ith* hospital; X_1 is ln of GDP per capita in 1998 I \$; 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 (I \$) 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 β and α 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, $\ln \overline{UC}$, can then be calculated from:

Equation 4

$$\ln\overline{UC} = \overline{\alpha_0} + \overline{\alpha_1}\ln X_1 + \sum_{i=1}^{n}\overline{\alpha_i}X_i$$

where $\overline{\alpha_0}$ and $\overline{\alpha_{i..n}}$ are the estimated parameters, and $X_{i..n}$ are the independent variables.

If $\overline{\beta_0} = \text{anti}\log\left(\overline{\alpha_0}\right)$ and $\overline{\beta_1} = \overline{\alpha_1}$, back-transforming Equation 4 (reduced to 1 independent log-transformed variable for simplicity) gives the power function.

Equation 5

$$\overline{UC}_{biased} = \overline{\beta_0}X^{\overline{\beta_1}}$$

where $\overline{UC}_{biased}$ 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, $\overline{C}_{bias}$, for the logarithmic transformation is given by:

Equation 6

$$\overline{C}_{bias} = \frac{1}{n}\sum_{i=1}^{n}b^{r_i}$$

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

Equation 7

$$\overline{UC} = \overline{C}_{bias}\overline{\beta_0}X^{\overline{\beta_1}}$$

The smearing correction factor ($\overline{C}_{bias}$) 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

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

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

(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

| Adjusted R^2 = 0.80	F statistic = 509		p of F statistic <0.00001	
Variable	Coefficient	SE	T	P
Ln GDP per capita	0.7624	0.0295	25.813	<0.0001
Ln occupancy rate	-0.2318	0.0474	-4.886	<0.0001
Drug costs	0.6410	0.1769	3.624	<0.0001
Food costs	0.2116	0.1394	1.518	0.152
Level 1 hospital	-0.5777	0.0742	-7.787	<0.0001
Level 2 hospital	-0.3118	0.0594	-5.253	<0.0001
Public	-0.2722	0.1172	-2.323	0.021
Private for profit	0.2444	0.1316	1.857	0.064
USA	1.7471	0.1022	17.104	0.000
Constant	-2.5036	0.3264	-7.672	0.026

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
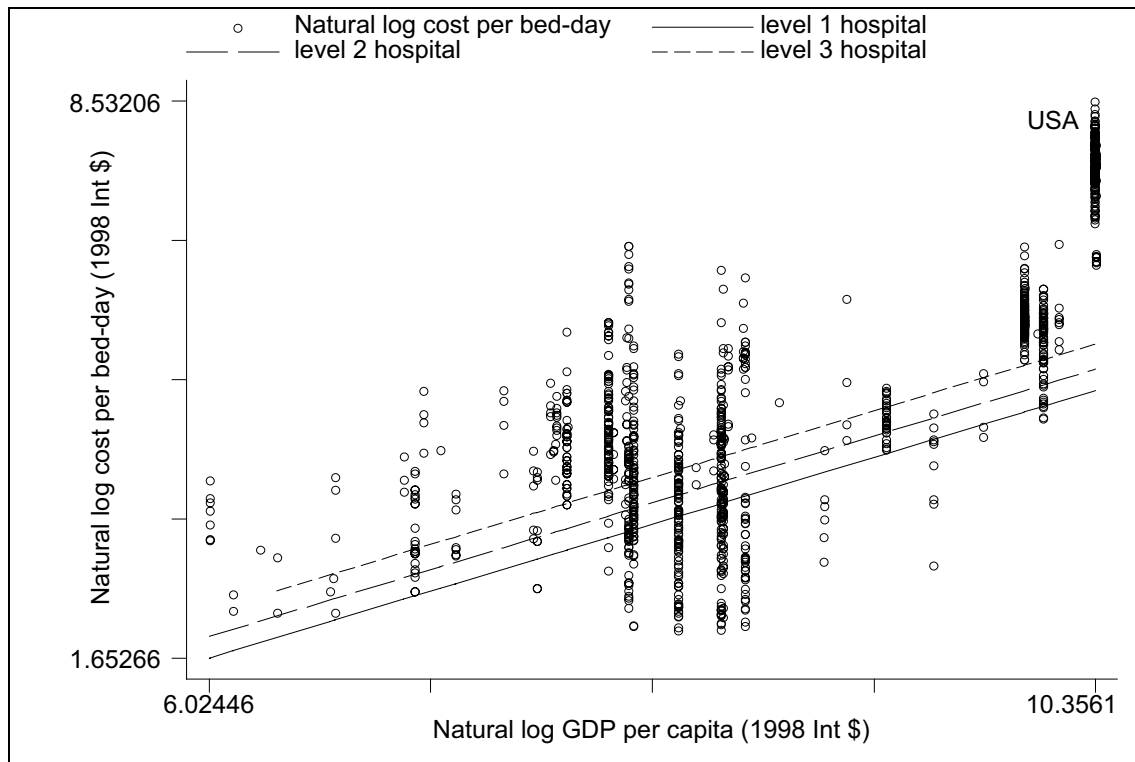
Figure 1
Regression lines for level one, two and three hospitals against the natural log of GDP per capita. (The Y-axis is the dependent variable: natural log of cost per bed day) Cost in 1998 I$ N = 1171

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-

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

Country	GDP per capita (I$)	In or out-of-sample	Hospital level	Cost per bed day			
				Mean (I $)	95% uncertainty interval Low	95% uncertainty interval high	SD
Mali	581	Out	I	7.39	5.46	9.80	1.36
			II	9.64	7.07	12.73	1.74
			III	13.14	9.58	17.40	2.44
Mozambique	720	Out	I	8.70	6.45	11.43	1.58
			II	11.35	8.32	15.01	2.03
			III	15.46	11.31	20.42	2.85
Algeria	1,449	Out	I	14.82	10.95	19.45	2.65
			II	19.35	14.26	25.45	3.39
			III	26.34	19.41	34.63	4.72
Indonesia	3,167	Out	I	26.90	19.86	35.19	4.77
			II	35.12	26.05	46.12	6.11
			III	47.80	35.38	63.28	8.44
Ecuador	3,260	In	I	27.50	20.30	35.95	4.88
			II	35.90	26.63	47.17	6.25
			III	48.87	36.17	64.71	8.63
Romania	3,634	Out	I	29.88	22.05	39.11	5.30
			II	39.00	28.86	51.25	6.79
			III	53.09	39.23	70.35	9.37
Greece	6,192	Out	I	44.88	33.02	59.23	8.03
			II	58.58	43.25	76.84	10.28
			III	79.73	59.01	104.62	14.10
Russian Federation	8,035	In	I	54.75	40.17	72.43	9.86
			II	71.48	52.73	94.24	12.62
			III	97.27	71.79	127.92	17.27
Bahrain	14,159	Out	I	84.41	61.48	112.10	15.52
			II	110.19	80.27	146.22	19.85
			III	149.93	110.19	198.97	27.01
United Arab Emirates	20,330	Out	I	111.30	80.27	148.59	20.81
			II	145.28	105.96	191.58	26.62
			III	197.66	144.13	263.01	36.09
United Kingdom	24,348	In	I	127.76	91.83	170.79	24.11
			II	166.77	121.08	221.00	30.85
			III	226.87	164.90	302.91	41.76
Canada	28,087	Out	I	142.51	102.03	190.22	27.11
			II	186.02	134.55	247.23	34.69
			III	253.05	183.16	337.26	46.89

(i)Cost per bed day is estimated for public hospitals with 80% occupancy rate, excluding drug costs and including food costs.

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;

- 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]

References
1. Hutubessy RCW, Baltussen RMPM, Tan Torres-Edejer T and Evans DB **Generalised cost-effectiveness analysis: an aid to decision making in health.** *Applied Health Economics and Health Policy* 2002, **1:**89-95
2. Barnum H and Kutzin J **Public hospitals in developing countries: resource use, cost, financing.** *Baltimore, The Johns Hopkins University Press for the World Bank* 1993,
3. Breyer F **The specification of a hospital cost function. A comment on the recent literature.** *J Health Econ* 1987, **6:**147-157
4. Conrad RF and Strauss RP **A multiple-output multiple-input model of the hospital industry in North Carolina.** *Appl Econ* 1983, **15:**341-352
5. Cremieux PY and Ouellette P **Omitted variable bias and hospital costs.** *J Health Econ* 2001, **20:**271-282
6. Grannemann TW, Brown RS and Pauly MV **Estimating hospital costs. A multiple-output analysis.** *J Health Econ* 1986, **5:**107-127
7. Lave JR and Lave LB **Hospital cost functions.** *Annu Rev Public Health* 1984, **5:**193-213
8. Vitaliano DF **On the estimation of hospital cost functions.** *J Health Econ* 1987, **6:**305-318
9. Li T and Rosenman R **Estimating hospital costs with a generalized Leontief function.** *Health Econ* 2001, **10:**523-538
10. Wagstaff A and Barnum H **Hospital cost functions for developing countries.** *WPS 1044. Washington, D.C, World Bank* 1992,
11. Lopez-Casasnovas G and Saez M **The impact of teaching status on average costs in Spanish hospitals.** *Health Econ* 1999, **8:**641-651
12. Liu X and Hsiao WC **The cost escalation of social health insurance plans in China: its implication for public policy.** *Soc Sci Med* 1995, **41:**1095-1101
13. Newhouse JP **Medical care costs: how much welfare loss?** *J Econ Perspect* 1992, **6:**3-21
14. Peden EA and Freeland MS **Insurance effects on US medical spending (1960–1993).** *Health Econ* 1998, **7:**671-687
15. Garattini L, Giuliani G and Pagano E **A model for calculating costs of hospital wards: an Italian experience.** *J Manag Med* 1999, **13:**71-82
16. Kirigia JM, Snow RW, Fox-Rushby J and Mills A **The cost of treating paediatric malaria admissions and the potential impact of insecticide-treated mosquito nets on hospital expenditure.** *Trop Med Int Health* 1998, **3:**145-150
17. Mitchell M, Thomason J, Donaldson D and Garner P **The cost of rural health services in Papua New Guinea.** *P N G Med J* 1991, **34:**276-284
18. Patcharanarumol W **A study of unit cost of out-patient and inpatient service of Khon Kaen Hospital in the fiscal year 1996.** *Bangkok, Chulalongkorn University* 1997,
19. Huff-Rousselle M **Dzongkhag costing study for Tashigang Dzongkhag.** *Royal Government of Bhutan, Department of Health Services, Ministry of Social Services* 1992,
20. Rannan-Eliya R and Somanathan A **Bangladesh facility efficiency survey.** *Working Paper No. 16. Dhaka, Health Economics Unit, Ministry of Health and Family Welfare, Government of the People's Republic of Bangladesh and Health Policy Programme, Institute of Policy Studies of Sri Lanka* 1999,
21. Musau S, Kilonzo M and Newbrander W **Development of a revised FIF user fee schedule. Report July 1996.** *Kenya Health Care Financing Project: Contract No. 623-0245-C-00-0040-00. Boston, Management Sciences for Health* 2000,
22. Robertson RL **Review of literature on costs of health services in developing countries.** *PHN Technical Note 85-21. Washington, Population, Health and Nutrition Department, World Bank* 1985,
23. Hospital services in Australia **Access and financing.** *National Health Strategy Issues Paper No. 2. Canberra, Department of Health, Housing and Community Services* 1991,
24. Ojo K **Economic and management perspectives of Windhoek state hospital complex.** *Windhoek, Ministry of Health and Social Services* 1995,
25. Fox-Rushby JA and Foord F **Costs, effects and cost-effectiveness analysis of a mobile maternal health care service in West Kiang, The Gambia.** *Health Policy* 1996, **35:**123-143
26. Brooks RG **Cost of selected health institutions in the Central Region 1972–1973.** *Ghana Med J* 1975, **14:**209-214
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
28. Flessa S **The costs of hospital services: a case study of Evangelical Lutheran Church hospitals in Tanzania.** *Health Policy Plan* 1998, **13:**397-407
29. Ministry of Health The Gambia World Health Organization **Cost analysis of the health care sector in The Gambia. Volume 1.** *Ministry of Health* 1995, **1:**

30. Puglisi R and Bicknell WJ **Functional expenditure analysis. Final Report for Queen Elizabeth II Hospital, Maseru, Lesotho.** *Boston University, Health Policy Institute* 1990, **I:**

31. Wong H **Cost analysis of Niamey hospital.** *USAID Project No. 683-0254. Bethesda, MD, Abt Associates Inc* 1989,

32. Ojo K, Foley J, Renner A and Kamara FM **Cost analysis of health services in Sierra Leone. A case study of Connaught hospital and Waterloo Community Health Centre. Annex III.** *Sierra Leone, Ministry of Health* 1995,

33. De Virgilio G, Haile M, Lemma A and Mariani D **Technical and economic efficiency of the Asella regional hospital in Ethiopia.** *La Medicina Tropicale nelle Cooperazione allo Sviluppo* 1990, **6:**1-7

34. Omar AO, Komakech W, Hassan AH, Singh CH and Imoko J **Costs, resource utilisation and financing of public and private hospitals in Uganda.** *East Afr Med J* 1995, **72:**591-598

35. Hansen K, Chapman G, Chitsike I, Kasilo O and Mwaluko G **The costs of HIV/AIDS care at government hospitals in Zimbabwe.** *Health Policy Plan* 2000, **15:**432-440

36. Pepperall J, Garner P, Fox-Rushby J, Moji N and Harpham T **Hospital or health centre? A comparison of the costs and quality of urban outpatient services in Maseru, Lesotho.** *Int J Health Plann Manage* 1995, **10:**59-71

37. Carey K and Burgess JF Jr **On measuring the hospital cost/quality trade-off.** *Health Econ* 1999, **8:**509-520

38. Raymond SU, Lewis B, Meissner P and Norris J **Financing and costs of health services in Belize.** *HCFLAC Research Report no. 2. Stony Brook, State University of New York* 1987,

39. Lewis MA, La Forgia GM and Sulvetta MB **Measuring public hospital costs: empirical evidence from the Dominican Republic.** *Soc Sci Med* 1996, **43:**221-234

40. Olave M and Montano Z **Unit cost and financial analysis for the hospital 12 de Abril in Bolivia.** *Small Applied Research Report No. 11. Bethesda, Abt Associates Inc* 1993,

41. Gill L and Percy A **Hospital costing study Glendon Hospital – Montserrat.** *Report No.15. Organisation of Eastern Caribbean States, Health Policy & Management Unit* 1994,

42. Russell SS, Gwynne G and Trisolini M **Health care financing in St Lucia and costs of Victoria Hospital.** *HCFLAC Research Report no. 5. Stony Brook, State University of New York* 1988,

43. Snow J **Papua New Guinea: health sector financing study project. Final Report – volume II hospital cost study** *Prepared for the Papua New Guinea Department of Health under contract with the Asian Development Bank, TA No. 1091-PNG* 1990,

44. Chan S **Unit cost estimation for outpatient and inpatient departments in Nakleoung District Hospital, Cambodia.** *Bangkok, Chulalongkorn University* 1997,

45. Banks DA, As-Sayaideh ASK, Shafei ARSH and Muhtash A **Implementing hospital autonomy in Jordan: an economic cost analysis of Princess Raya Hospital.** *Bethesda, MD, The Partners for Health Reformplus Project, Abt Associates Inc* 2002,

46. Department of Health **The new NHS – 2001 reference costs.** *London, Department of Health* 2001,

47. Department of Planning Ministry of Health and Population Data for Decision Making, Harvard School of Public Health, University of California, Berkeley, School of Public Health **Cost analysis and efficiency indicators for health care: report number 1 summary output for Bani Suef General Hospital, 1994–1994.** *Boston, Harvard University Press* 1997,

48. Department of Planning, Ministry of Health and Population, Data for Decision Making, Harvard School of Public Health, University of California, Berkeley, School of Public Health **Cost analysis and efficiency indicators for health care: report number 2 summary output for Suez General Hospital, 1993–1994.** *Boston, Harvard University Press* 1997,

49. Department of Planning, Ministry of Health and Population, Data for Decision Making, Harvard School of Public Health, University of California, Berkeley, School of Public Health **Cost analysis and efficiency indicators for health care: report number 3 summary output for El Gamhuria General Hospital, 1993–1994.** *Boston, Harvard University Press* 1997,

50. Department of Planning, Ministry of Health and Population, Data for Decision Making, Harvard School of Public Health, University of California, Berkeley, School of Public Health **Cost analysis and efficiency indicators for health care: report number 4 summary output for 19 primary health care facilities in Alexandria,** **Bani Suef and Suez, 1993–1994.** *Boston, Harvard University Press* 1997,

51. Waters H, Abdallah H and Santillan D **Application of activity-based costing (ABC) for a Peruvian NGO healthcare provider.** *Int J Health Plann Manage* 2001, **16:**3-18

52. Robertson RL, Barona B and Pabon R **Hospital cost accounting and analysis: the case of Candelaria.** *J Community Health* 1977, **3:**61-79

53. Jorbenadze A, Zoidze A, Gzirirshvili D and Gotsadze G **Health reform and hospital financing in Georgia.** *Croat Med J* 1999, **40:**221-236

54. Mills AJ **The cost of the district hospital. A case study from Malawi.** *WPS 742. Washington,D.C, World Bank* 1991,

55. StataCorp **Stata Statistical Software: Release 7.0.** *College Station, TX, Stata Corporation* 2001,

56. World Bank **World Development Indicators 2000.** *Washington, DC, World Bank* 2000,

57. King G, Honaker J, Joseph A and Scheve K **List-wise deletion is evil: what to do about missing data in political science.** *Paper presented at the Annual Meeting of the American Political Science Association, Boston* 1998,

58. Allison PD **Multiple imputation for missing data. A cautionary tale.** *Sociological Methods & Research* 2000, **28:**301-309

59. Honaker J, Joseph A, King G, Scheve K and Singh N **Amelia: A Program for Missing Data (Windows Version).** *Cambridge, MA, Harvard University* 1999,

60. Lu K and Tsiatis AA **Multiple imputation methods for estimating regression coefficients in the competing risks model with missing cause of failure.** *Biometrics* 2001, **57:**1191-1197

61. Patrician PA **Multiple imputation for missing data.** *Res Nurs Health* 2002, **25:**76-84

62. King G, Tomz M and Wittenberg J **Making the most of statistical analyses: improving interpretation and presentation.** *Am J Pol Sci* 2000, **44:**341-355

63. King G, Honaker J, Joseph A and Scheve K **Analyzing incomplete political science data: An alternative algorithm for multiple imputation.** *Am Polit Sci Rev* 2000,

64. Rubin DB **Multiple imputation for survey nonresponse.** *New York, Wiley* 1986,

65. Cowing TG, Holtmann AG and Powers S **Hospital cost analysis: a survey and evaluation of recent studies.** *Adv Health Econ Health Serv Res* 1983, **4:**257-303

66. Gujarati DN **Basic econometrics.** *New York, McGraw-Hill, Inc* 1995,

67. Duan N **Smearing estimate: a nonparametric retransformation method.** *J Amer Statistical Assoc* 1983, **78:**605-610

68. King G **How not to lie with statistics: avoiding common mistakes in quantitative political science.** *Am J Pol Sci* 1986, **30:**666-687

69. US General Accounting Office **Military health care. Savings to CHAMPUS from using a prospective payment system.** *Fact Sheet for Congressional Requesters GAO HRD-90-136FS. Washington D.C., United States General Accounting Office* 1990,

70. Wouters A **The cost and efficiency of public and private health care facilities in Ogun State, Nigeria.** *Health Econ* 1993, **2:**31-42

71. Anderson DL **A statistical cost function study of public general hospitals in Kenya.** *Journal of Developing Areas* 1980, **14:**223-235