Evaluating the Impact of Hospital Information Systems on the Technical Efficiency of 8 Central African Hospitals Using Data Envelopment Analysis

Frank Verbeke a,b,c *, Evelyne Ndabaniwe c, Stefaan van Bastelaere d, Ousmane Ly a,e, Marc Nyssen a,b

a VUB: Department of biostatistics and medical informatics, Faculty of Medicine and Pharmacy, Vrije Universiteit Brussels, Brussels, Belgium
b REHCE: Regional e-Health Center of Excellence, Kigali Health Institute, Kigali, Rwanda
c BHIA: Burundi Health Informatics Association, Bujumbura, Burundi
d BTC: Belgian Technical Cooperation, Brussels, Belgium
e ANTIM: Agence Nationale de Télé-santé et d’Informatique Médicale du Mali

Objectives: this study evaluates the usability of Data Envelopment Analysis (DEA) for analyzing the technical efficiency before and after hospital information system (HIS) implementation for a set of 8 Central African hospitals (6 Rwandan, 2 Burundian; 6 public and 2 private).

Methods: DEA is a method that uses linear programming techniques to produce a relative efficiency score for organizational units where the presence of multiple inputs and outputs makes straightforward comparisons difficult. DEA is non-parametric, requiring no assumptions about the (most often unknown) functional relationship between inputs and outputs (in contrast to regression based models). The method directly compares health facilities against a combination of peers. In this study post-HIS implementation health facility productivity was also compared against results obtained before HIS implementation.

Results: the average technical efficiency increase of 5.04% after HIS implementation appeared not to be statistically significant in our small dataset.

Conclusions: despite the lack of statistical significance, the results still suggest that DEA may offer interesting opportunities for measuring productivity impact of large scale implementations of health information management methods and systems using data sets from heterogeneous collections of health facilities. Further research on an extended set of sub-Saharan health facilities has been programmed for that purpose.

Keywords: Data Envelopment Analysis, Sub-Saharan Africa, Technical efficiency, Hospital Information Systems

1 Introduction

An increasing number of hospital information management system (HIS) implementations have been reported in sub-Saharan Africa in the last few years [14]-[15]. Although lots of (potential) benefits of electronic health information management have been extensively documented in the literature, it remains difficult to measure the impact of HIS deployment on a health facility’s output and productivity.

Productivity is a complex concept for which calculations are based on inputs (workforce, buildings, medical equipment, funding…) and outputs (case load, completed treatments, morbidity and mortality reduction…). No simple one-dimensional metrics (such as ratios) exist for expressing the complexity and richness of hospital productivity. As an alternative, a number of linear programming techniques, such as Data Envelopment Analysis (DEA), have gained popularity in the health domain for producing relative
efficiency scores for organizational units where the presence of multiple inputs and outputs makes comparisons difficult. Most often, DEA and related techniques have been used for benchmarking health facilities against other facilities which operate in the same context (health centres, district hospitals etc.).

The objective of our study was to also explore the usability of a DEA-based method for measuring the impact of HIS implementation on hospital productivity in a set of low resource sub-Saharan health facilities.

2 Materials and methods

As shown in Fig. 3, many different efficiency and productivity measurement approaches exist: one- and multidimensional methods, frontier- and average-based methods, parametric and non-parametric methods and stochastic or deterministic approaches. Advantages and disadvantages of these methods have been evaluated prior to this study, resulting in Data Envelopment Analysis being chosen as the best adapted solution for our purpose.

**Fig. 3.** Different evaluated productivity measurement techniques

### 2.1 Data Envelopment Analysis (DEA)

DEA is a method that uses linear programming techniques to produce a relative efficiency score for organizational units where the presence of multiple inputs and outputs makes comparisons difficult [1]-[4]. It is evident that health facilities are always using multiple inputs to produce multiple outputs, which makes a method such as DEA better adapted to our needs than the simple usage of ratios. DEA evaluates relative efficiency of each unit among a set of more or less homogeneous decision making units (DMU), e.g. health centres or district hospitals. It draws a frontier of best possible productivity combining inputs and outputs from the best performing DMUs (health facilities in our case). Health facilities that compose this best practice frontier are assigned an efficiency score of 1 and are being considered technically efficient compared to other health facilities. Health facilities that are situated below the efficiency frontier will forcibly be inefficient. Their level of inefficiency is measured in terms of their distance from the frontier and is being expressed by a score between 0 and 1, larger scores expressing higher efficiencies.

As multiple inputs are being used to produce multiple outputs by health facilities, the technical efficiency of an individual health facility (Eff$_j$) can be expressed as the ratio of the weighted sum of outputs divided by the weighted sum of inputs:

$$Eff_j = \frac{\sum_w w \cdot y_{ij}}{\sum_w w \cdot x_{ij}}$$

© 2013 HELINA and JHIA. This is an Open Access article published online by JHIA and distributed under the terms of the Creative Commons Attribution Non-Commercial License. DOI: 10.1285/JHIA.2013-v1-i1-68
Eff\textsubscript{j} = \frac{u_1 y_{1j} + u_2 y_{2j} + \ldots}{v_1 x_{1j} + v_2 x_{2j} + \ldots}

Where:

\begin{align*}
  u_1 & = \text{the weight given to output 1} \\
  y_{1j} & = \text{amount of output 1 produced by health facility j} \\
  v_1 & = \text{the weight given to input 1} \\
  x_{1j} & = \text{amount of input 1 used by health facility j}
\end{align*}

Technically inefficient health facilities will then use more weighted inputs per weighted outputs or produce less weighted outputs per weighted inputs than their peers on the best practice frontier.

This measure of efficiency assumes that a common set of weights be applied to inputs and outputs across all DMUs. Such assumption will of course raise the problem of agreeing on a common set of weights. Different health facilities may choose to organize their care activities differently so that the relative values of their inputs and outputs may legitimately be different (e.g. one hospital might value the reduction in maternal mortality rate more than the number of admissions performed whilst another hospital might have the opposite approach).

The difficulty of finding a common set of weights had been recognized by Charnes, Cooper and Rhodes [5], accepting the legitimacy of the fact that DMUs might value inputs and outputs differently and therefore apply different weights. Therefore, they proposed that each DMU should have the opportunity of choosing the most favourable weights possible for comparing it to all other DMUs. In that case, the efficiency of a target DMU j0 can be calculated by finding the maximum possible efficiency of j0 using any combination of weight values whereby the efficiency of all DMUs remains <= 1. The solution will produce the weights most favourable to j0 and provide a resulting measure of efficiency. Algebraically, this can be represented as follows:

\[
\text{Eff}_{j0} = \max \frac{\sum_{r} u_r y_{rj0}}{\sum_{i} v_i x_{ij0}}
\]

stated that:

\[
\sum_{r} u_r y_{rj} \leq 1, \quad \sum_{i} v_i x_{ij0}
\]

\[u_r, v_i \geq \varepsilon; \quad \forall r, \forall i\]

Where:

\begin{align*}
  u_r & = \text{the weight given to output } r \\
  y_{rj} & = \text{the amount of output } r \text{ produced by health facility } j \\
  v_i & = \text{the weight given to input } i \\
  x_{ij} & = \text{the amount of input } i \text{ used by health facility } j \\
  j_0 & = \text{the health facility being assessed}
\end{align*}

The variables u and v are constrained to be greater than the constant \(\varepsilon\), which is a small positive quantity, in order to avoid that any input or output would be totally ignored in calculating the efficiency.
The above linear program is fractional and cannot be solved without converting it to a linear form. The individual values of the numerator and denominator in the above Effj0 equation are not important (they have no meaning after all); we are only interested in the ratio. Therefore it is acceptable to set the denominator equal to a constant (e.g. 1) resulting in the fact that we will remain with the numerator to be maximized. After such transformation, the linear program becomes as follows:

$$\text{Eff}_{j0} = \max \sum_r u_r y_{rj0}$$

$$u_r, v_1$$

stated that:

$$\sum_r u_r y_{rj} - \sum_i v_i x_{ij} \leq 1 ; \forall j$$

$$\sum_i v_i x_{ij0} = 1$$

$$u_r, v_1 \geq \varepsilon ; \forall r, \forall i$$

Where:

- $u_r$ = the weight given to output $r$
- $y_{rj}$ = the amount of output $r$ produced by health facility $j$
- $v_i$ = the weight given to input $i$
- $x_{ij}$ = the amount of input $i$ used by health facility $j$
- $j_0$ = the health facility being assessed

The resulting efficiency score of a health facility $j_0$ will be a value between 0 and 1 indicating how much of the weighted inputs used by $j_0$ would have been needed by an efficient health facility (score=1) to produce the same amount of weighted outputs as $j_0$. Consequently, we may also be able to calculate the effort necessary for an inefficient health facility $j_0$ to become efficient.

In order to illustrate this, let's create an example data set consisting of input & output data for a number of fictitious district hospitals over a period $t$: Every health facility produces 2 outputs (outpatient visits and in-patient admissions) from a single input being the number of employees working in the hospital. Output and input values for all health facilities are provided in Table 2:

<table>
<thead>
<tr>
<th>DMU name</th>
<th>Outpatient visits</th>
<th>Admissions</th>
<th>Staff</th>
<th>Outpatients/staff</th>
<th>Admissions/staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mulinga</td>
<td>42144</td>
<td>3112</td>
<td>188</td>
<td>224</td>
<td>17</td>
</tr>
<tr>
<td>Razonde</td>
<td>6055</td>
<td>984</td>
<td>47</td>
<td>129</td>
<td>21</td>
</tr>
<tr>
<td>Kipantu</td>
<td>254098</td>
<td>12108</td>
<td>767</td>
<td>331</td>
<td>16</td>
</tr>
<tr>
<td>Kusombo</td>
<td>108008</td>
<td>4366</td>
<td>311</td>
<td>347</td>
<td>14</td>
</tr>
<tr>
<td>Hinanji</td>
<td>86650</td>
<td>5113</td>
<td>277</td>
<td>313</td>
<td>18</td>
</tr>
<tr>
<td>Matabu</td>
<td>9806</td>
<td>1460</td>
<td>83</td>
<td>118</td>
<td>18</td>
</tr>
<tr>
<td>Timimbi</td>
<td>18120</td>
<td>1060</td>
<td>144</td>
<td>126</td>
<td>7</td>
</tr>
<tr>
<td>Rotungo</td>
<td>32077</td>
<td>1656</td>
<td>115</td>
<td>279</td>
<td>14</td>
</tr>
<tr>
<td>Fikipso</td>
<td>45050</td>
<td>804</td>
<td>76</td>
<td>593</td>
<td>11</td>
</tr>
<tr>
<td>Lamindo</td>
<td>32355</td>
<td>180</td>
<td>54</td>
<td>599</td>
<td>3</td>
</tr>
</tbody>
</table>

Using ratios such as out-patients/staff or admissions/staff, we can express the efficiency of a health facility in producing a single output based on the input. There we can clearly see that the Lamindo
hospital is more efficient than Rotungo hospital in terms of out-patient visits performance. On the other hand, the Rotungo hospital is more efficient than Lamindo in terms of the number of hospital admissions per staff member. Depending on which of both would be valued most, both of the health facilities could rightfully state that they are more efficient than the other one. The example is trivial if we only consider these 2 hospitals, but things become more confusing when we start adding more health facilities to the evaluation.

Using the DEA approach, we will be able to combine efficiencies for multiple inputs and/or outputs by first identifying the health facilities that for any combination of input and/or output weights could rightfully state that they are more efficient than the others. In our data sample, such would be the case for the Razonde, Hinanjji, Fikipso and Lamindo hospitals. The connecting lines between these efficient hospitals, as shown in Fig. 4, represent the efficiency frontier for our dataset. The frontier provides the boundaries of the best possible productivity that can be achieved based on data available and therefore can also be used as a threshold against which performance of other inefficient hospitals can be measured. Graphically, the efficiency frontier envelops the inefficient health facilities.

The efficiency score of a health facility (e.g. Mulinga) can be calculated as the ratio of its distance to the origin (black arrow) over the distance from the origin to the efficiency frontier (blue line). In the case of Mulinga, a score of 0.863 states that a hypothetic efficient hospital situated on the efficiency frontier (point A) would be able to produce the same outputs as Mulinga using only 86.3% of its inputs. In other words, in order to become efficient, Mulinga could reduce its inputs (staff employed) from 188 to 162 (input oriented approach) keeping the outputs at the same level. Another possibility would be to increase the outputs to 48817 out-patient consultations and 3605 admissions (output oriented approach) whilst keeping the input unchanged. And of course, combinations of both approaches (non-oriented approach) would also be possible.

![Efficiency Frontier](image)

Fig. 4. Sample DEA analysis

### 2.2 Returns to Scale

The above projections of inputs and/or outputs needed by a non-efficient health facility to become efficient, do only make sense in cases where a change of all inputs by a proportion $\sigma$ also leads to an increase of the outputs by the same proportion $\sigma$. Such case is called Constant Returns to Scale (CRS). In our example this would mean that increasing hospital staff by 10% would also automatically mean a 10% increase in out-patient consultations and a 10% increase in hospital admissions. Of course, such is rarely the case in real practice. A less than proportional increase of outputs would be called Decreasing Returns to Scale (DRS) and a more than proportionally increase of outputs becomes an Increasing Returns to Scale (IRS).

Health facilities that display CRS can be considered to be operating at their best productivity level. When DRS applies, the health facility should scale down and reduce inputs in order to reach its most...
productive size. In case of IRS, a health facility should increase its inputs in order to become scale efficient [6][7]. However, because hospital production processes are most often not linear, it seems appropriate to assume a default approach of Variable Returns to Scale (VRS) [8]. These can be calculated using DEA methods developed by Banker, Charnes and Cooper (BCC) in 1984, whereas the methods of the original Charnes, Cooper and Rhodes (CCR) model assumed CRS. Algebraically, the adapted model for VRS can be written as follows:

\[ \text{Eff}_{j0} = \max \sum_r u_r y_{rj0} + u_0 \]

where:

- \( u_r \) the weight given to output \( r \)
- \( u_0 \) free in sign
- \( y_{rj} \) the amount of output \( r \) produced by health facility \( j \)
- \( v_i \) the weight given to input \( i \)
- \( x_{ij} \) the amount of input \( i \) used by health facility \( j \)
- \( j_0 \) the health facility being assessed

stated that:

\[ \sum_r u_r y_{rj} - \sum_i v_i x_{ij} + u_0 \leq 1 \quad \forall j \]

\[ \sum_i v_i x_{ij0} = 1 \]

\[ u_r, v_i \geq 0 \quad \forall r, \forall i \]

2.3 DEA Strengths

DEA is multi factor, meaning that it can account for multiple inputs and outputs which is quite typical for health facilities. There are no requirements with regard to the units being used for inputs and outputs (they can be completely different). DEA is non-parametric, requiring no assumptions about the (most often unknown) functional relationship between inputs and outputs (in contrast to regression based models) [9][10]. The method directly compares health facilities against a combination of peers. Like ratios, DEA can be used to measure technical or productive efficiency. If cost data are available, differences in technical efficiency can be distinguished from differences in the costliness of the mix of productive inputs (e.g. the balance between physician and nursing labour). On the other hand, no cost-information related to inputs and outputs is required [11].

2.4 DEA Weaknesses

DEA does assume that all inputs and outputs are included in the analysis, meaning that the results may be unreliable if this assumption is not correct. DEA is typically "deterministic," that is, the method usually ignores random noise in inputs and outputs as a potential source of variation in efficiency scores. Any deviation from the best practice frontier is being attributed to inefficiency although part of it could be caused by statistical noise (measurement errors or temporary changes in the health care environment such as epidemics). The deterministic and non-parametric nature of DEA makes it difficult to perform statistical tests on the production function.

2.5 Data set

All hospitals (n=8; 6 Rwandan, 2 Burundian; 6 public and 2 private) in our study set had implemented an open source HIS (OpenClinic [12]) in the period between 2006 and 2012. For each hospital, an
assessment was performed just before (baseline) and 1 year after HIS implementation taking into account a selection of input- and output metrics. Input metrics included labour inputs (physicians, nurses, ICT-staff and others) and capital inputs (operational beds). Output metrics included out-patient and in-patient case load and mortality rate.

The resulting data were merged in a table of 16 health facility states (8 pre-implementation and 8 post-implementation states). DEA analysis was performed comparing pre-implementation to post-implementation technical efficiency. DEA software used was MaxDEA version 5.2 [13].

3 Results

DEA was used to analyze the productivity of health facilities in terms of the generated volume of out-patient encounters and in-patient admissions related to hospital staffing, called case load total labour technical efficiency (MLTL). For analyzing MLTL, the model selected was revenue/cost based, non oriented and with constant returns to scale (CRS). Although in general, variable returns to scale (VRS) better matches the health sector reality, too many facilities in our limited data set would have been projected on the efficiency border. Taking into account the limitations of the chosen CRS model, the following inputs have been fed into DEA:

**Labour**
- Physician personnel input (MD): average number of full-time equivalents of physicians employed by the health facility
- Nurse personnel input (NURS): average number of full-time equivalents of nurses
- Other personnel input (OTH): the sum of all remaining personnel categories
- ICT-personnel input (IT): average number of full-time equivalents of ICT staff

**Capital**
- Operational admission beds (ICBE)

2 case load outputs have also been considered:
- Out-patient case load (OACO)
- In-patient case load (OACI)

Input unit costs being considerably different for Rwanda and Burundi, different unit prices have been used for MD, NURS, OTH and ICBE based on market averages which have all been converted to Rwandan francs (±620 Rwandan francs = 1 USD). Output unit prices were based on average user fees collected for out-patient encounters (OACO) and in-patient admissions (OACI). Output prices were also corrected for price indices calculated yearly for the period 2007-2012.

All parameters have been provided for every health facility prior to HIS implementation (PRE) and 1 year post-implementation (POST), resulting in Table 3.

As shown in Table 4, in the post-HIS implementation group, DEA identified the private Rwandan hospital La Croix du Sud (CDS) as being technically the most efficient health facility, meaning that they used the least inputs for producing outputs, independently from whichever weight one might give to any of the input and output variables. CDS therefore gets an MLTL technical efficiency score of 100%. The second most efficient health facility was the Military Hospital of Kamenge (HMK), with a technical efficiency score of 28,54%. The score means that a technically efficient health facility (CDS) would have been able to produce the same outputs using only 28,54% of the resources consumed by HMK. The HMK could then become technically efficient by reducing its inputs keeping its outputs constant (input oriented), by increasing its outputs keeping the inputs constant (output oriented) or by a combination of both (non-oriented). Necessary change in inputs and outputs to become technically efficient are provided for every health facility in Table 4 in the slack movement columns.
Obviously, overall hospital efficiency does not exclusively depend on a limited set of input and output variables like the ones in our example. Therefore, slack movements suggested by DEA may sometimes appear very unrealistic. In our study, the Neuro-psychiatric hospital of Ndera showed to be the least technically efficient facility (17.58%). In order to become technically efficient, the hospital would have to reduce the 243 available beds by 233 units remaining with only 10 beds. Clearly, Ndera’s efficiency results are being heavily compromised by the fact that the median length of stay is considerably longer for neuro-psychiatric patients (some patients may remain hospitalized for years). This clearly demonstrates the fact that DEA can only be used to technically compare decision making units (DMUs or health management units) in order to improve resource allocation and efficiency.
facilities in our case) that sufficiently share common context: district hospitals must not be mixed with health centres, private clinics should not be compared to public hospitals etc. Homogeneous grouping could not be done (and never was the purpose) for the limited sample of health facilities in our research. On the contrary, the sole purpose of DEA was to evaluate pre- and post-implementation change in technical efficiency for individual health facilities (the pre- and post-implementation operational contexts of the same health facility being considered comparable). From this point of view, for each of the 8 analyzed hospitals, we found that MLTL had improved after implementing the OpenClinic HIS, as is demonstrated in Fig. 5.

Yet, the average increase of 5.04% appeared not to be statistically significant (single factor ANOVA test comparing pre- and post-implementation technical efficiency scores). Further research on an extended set of HIS-implementing health facilities will be needed in order to confirm or reject the obtained results and has already been planned for integration in a number of new implementation programs in Burundi, DRC, Mali, Congo-Brazzaville and Senegal.

4 Discussion

The proposed method for studying pre- and post HIS implementation change in technical efficiency looks promising when taking into account a number of important limitations. First of all, based on the preliminary results of the study, post-HIS implementation productivity improvements (based on multiple inputs and outputs) seem to be modest, requiring a sufficiently large number of study sites for demonstrating statistically significant changes. In case of different types/brands of information systems being used in the study sites, the number of required sites may further increase. Also, the study sites should preferably be located in different countries and technical efficiencies should be analyzed at different points in time in order to filter out location- or time related bias. Another problem is the fact that a more or less comprehensive set of input- and output metrics should be considered when performing DEA, which is only feasible in few low-resource hospitals in the pre-HIS implementation phase (data unavailable or unreliable, data collection to expensive etc.). Finally, DEA does not provide any information on causal relationships between input and output variables, making it hard to transform technical efficiency data into corrective actions.

For these reasons, the use of DEA for measuring HIS impact on productivity does not seem an appropriate method for evaluating progress made by individual health facilities. It may however still constitute a useful instrument for measuring global productivity impact of large scale implementations of specific health information management methods and/or systems: when performing statistical analysis solely on per-facility pre- and post HIS-implementation productivity progress, homogeneity of the health facilities in the study sample becomes irrelevant.
References

[16] Implementing electronic medical record systems in developing countries, Hamish SF Fraser, Paul Biondich, Deshen Moodley, Sharon Choi, Burke W Mamlín, Peter Szolovits, Informatics in Primary Care 2005;13:83–95