Healthcare Waste Generation Worldwide and Its Dependence on Socio-Economic and Environmental Factors

Minas Minoglou ¹, Spyridoula Gerassimidou ² and Dimitrios Komilis ¹,*

¹ Laboratory of Solid and Hazardous Waste Management, Department of Environmental Engineering, Democritus University of Thrace, Xanthi 67132, Greece; mminogl@civil.duth.gr
² Department of Civil and Environmental Engineering, University of Leeds, Leeds LS2 9JT, UK; sgerasim@env.duth.gr
* Correspondence: dkomilis@env.duth.gr; Tel.: +30-25410-79391

Academic Editor: Vincenzo Torretta
Received: 24 November 2016; Accepted: 3 February 2017; Published: 6 February 2017

Abstract: This paper examines the dependence of the healthcare waste (HCW) generation rate on several social-economic and environmental parameters. Correlations were calculated between the quantities of healthcare waste generated (expressed in kg/bed/day) versus economic indices (GDP, healthcare expenditure per capita), social indices (HDI, IHDI, MPI, life expectancy, mean years of schooling, HIV prevalence, deaths due to tuberculosis and malaria, and under five mortality rate), and an environmental sustainability index (total CO₂ emissions) from 42 countries worldwide. The statistical analysis included the examination of the normality of the data and the formation of linear multiple regression models to further investigate the correlation between those indices and HCW generation rates. Pearson and Spearman correlation coefficients were also calculated for all pairwise comparisons. Results showed that the life expectancy, the HDI, the mean years of schooling and the CO₂ emissions positively affect the HCW generation rates and can be used as statistical predictors of those rates. The resulting best reduced regression model included the life expectancy and the CO₂ emissions and explained 85% of the variability of the response.

Keywords: healthcare waste; generation rates; economic factors; social factors; correlation; regression modeling

1. Introduction

A variety of terms have been used in the literature to describe medical waste. In recent articles, the most popular term appears to be “healthcare waste” (HCW) and the most classical unit for the expression of its generation rate is mass per bed per day. There is a diversity of classifications of HCW, but the principal one is that HCW contain hazardous (mostly infectious waste) and non-hazardous (municipal solid waste) fractions. When reporting HCW generation rates (HCWGR), it is very important to specify whether the non-hazardous waste stream is included, since that fraction usually constitutes 80% of the total HCW stream, based on rough estimations provided by WHO [1]. Still, this estimation needs to be further investigated with detailed studies.

There are a variety of factors that contribute to the variability of the reported HCWGR, such as different hospital services, uncertainty on whether the non-hazardous fraction is included in the quantification of HCWGR, the units of expressing HCWGR, financial factors, etc. [2]. For example, it might be difficult to compare HCWGR between economically different countries due to different legal frameworks, differences in healthcare services and HCW management systems, and illegal dumping. The comparison may be more reliable among the most economically developing countries, for which there are similar environmental problems and strict budgets [3].
HCWGRs are usually lower in developing and poor countries than in the developed world [4]. However, it is observed that there is a steady global increment of the HCW production worldwide. In middle- and low-income countries, HCW production is sharply increasing due to improved access to healthcare services [5]. In wealthy nations, the increment of HCWGR is attributed to the rapidly aging population which leads to an increasing system usage [6].

In the present study, the healthcare waste generation rates and several socio-economic and environmental factors were recorded for 42 countries worldwide. Those factors included economic factors (gross domestic product (GDP) per capita, health expenditure per capita), social and health-related factors (human development index (HDI), inequality-adjusted human development index (IHDI), multidimensional poverty index (MPI), life expectancy at birth, mean years of schooling, HIV prevalence, deaths due to tuberculosis, deaths due to malaria, under-five mortality rate), and one environmental sustainability factor (CO$_2$ emissions). Those HCWGR included both the hazardous and non-hazardous (urban) fractions. The correlation of all aforementioned factors with the HCWGR was investigated in this work.

The aim of the work was to develop a best reduced regression model to predict the HCWGR as a function of certain indices. To the knowledge of the authors, there have been limited studies in the literature to assess the potential effect of such a number of factors on the HCWGR. Only Windfeld and Brooks (2015) have recently examined the influence of HCWGR on two factors, namely GDP and health expenditure (HE), and had observed that both had significantly positive correlations with HCWGR [7]. Karpusenkaite et al. (2016) evaluated the performance of various mathematical modeling methods during the forecasting of medical waste generation rates using Lithuania’s annual medical waste data (visits at hospital, number of children, number of visits at hospital, average life expectancy, etc.) [8].

The practical application of the work is that certain widely used indices per country can be used to mathematically predict HCWGR using empirical models. Thus, direct weight measurements of HCW, which are usually costly, can be avoided since the regression models can be used instead to make such evaluations. The knowledge of the HCWGR is important when designing HCW management systems, which commonly comprise collection, treatment, and disposal.

2. Methodology

2.1. Materials

The HCWGR for 42 countries were recorded based on a literature search, the results of which are included in Table 1. The basic problem during that recording was that, in a few cases, it was not clear if the generation rates referred to hazardous wastes only or to the total fraction of HCW; that is, whether HCW contained both the hazardous and non-hazardous fractions. This distinction was particularly difficult in some developing and low-income countries. In most cases, however, it was clear that HCW included both fractions. In those unclear cases (of which there were seven), we assumed that the reported HCWGR contained both fractions.

The HCWGR reported in Table 1 include the hazardous and non-hazardous fractions of HCW. It is mentioned, however, that China, Korea, Lebanon, Argentina, El-Salvador, Bulgaria, and Greece provided information only for the hazardous fraction of HCW. According to WHO estimation, approximately 80% of the total HCW stream consists of the non-hazardous (urban) fraction [1]. The total HCW (hazardous and non-hazardous) for those seven countries was calculated by taking into account that percentage so as to be able to compare all 42 values together.
Table 1. Healthcare waste generation rates in selected countries worldwide (Africa).

<table>
<thead>
<tr>
<th>Country</th>
<th>HCWGR (kg/Bed/Day)</th>
<th>References</th>
<th>Country</th>
<th>HCWGR (kg/Bed/Day)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Algeria</td>
<td>0.96</td>
<td>[9]</td>
<td>Mauritius</td>
<td>0.44</td>
<td>[10]</td>
</tr>
<tr>
<td>Cameroon</td>
<td>0.55</td>
<td>[11]</td>
<td>Morocco</td>
<td>0.53</td>
<td>[12]</td>
</tr>
<tr>
<td>Egypt</td>
<td>1.03</td>
<td>[13,14]</td>
<td>Sudan</td>
<td>0.87</td>
<td>[15]</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>1.1</td>
<td>[16]</td>
<td>Tanzania</td>
<td>0.75</td>
<td>[17,18]</td>
</tr>
<tr>
<td>Asia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bangladesh</td>
<td>1.24</td>
<td>[3,19,20]</td>
<td>Malaysia</td>
<td>1.9</td>
<td>[20]</td>
</tr>
<tr>
<td>China</td>
<td>4.03</td>
<td>[21,22]</td>
<td>Pakistan</td>
<td>2.07</td>
<td>[7]</td>
</tr>
<tr>
<td>India</td>
<td>1.55</td>
<td>[20,23]</td>
<td>Palestine</td>
<td>2.02</td>
<td>[24]</td>
</tr>
<tr>
<td>Indonesia</td>
<td>0.75</td>
<td>[25]</td>
<td>Thailand</td>
<td>2.05</td>
<td>[26]</td>
</tr>
<tr>
<td>Iran</td>
<td>3.04</td>
<td>[20,27]</td>
<td>Turkey</td>
<td>4.55</td>
<td>[17,28]</td>
</tr>
<tr>
<td>Japan</td>
<td>2.15</td>
<td>[10,17]</td>
<td>Nepal</td>
<td>0.5</td>
<td>[20]</td>
</tr>
<tr>
<td>Jordan</td>
<td>2.69</td>
<td>[17]</td>
<td>Lebanon</td>
<td>5.7</td>
<td>[29]</td>
</tr>
<tr>
<td>Korea</td>
<td>2.4</td>
<td>[30]</td>
<td>Kazakhstan</td>
<td>5.34</td>
<td>[31]</td>
</tr>
<tr>
<td>Laos</td>
<td>0.51</td>
<td>[28]</td>
<td>Vietnam</td>
<td>1.57</td>
<td>[20,32]</td>
</tr>
<tr>
<td>America</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>3</td>
<td>[4]</td>
<td>Ecuador</td>
<td>2.09</td>
<td>[33]</td>
</tr>
<tr>
<td>Brazil</td>
<td>2.94</td>
<td>[34,35]</td>
<td>El Salvador</td>
<td>1.85</td>
<td>[36]</td>
</tr>
<tr>
<td>Canada</td>
<td>8.2</td>
<td>[33]</td>
<td>USA</td>
<td>8.4</td>
<td>[7,17,35]</td>
</tr>
<tr>
<td>Europe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bulgaria</td>
<td>2</td>
<td>[7]</td>
<td>Netherlands</td>
<td>1.7</td>
<td>[37]</td>
</tr>
<tr>
<td>Italy</td>
<td>4</td>
<td>[17]</td>
<td>Norway</td>
<td>3.9</td>
<td>[7]</td>
</tr>
<tr>
<td>Germany</td>
<td>3.6</td>
<td>[37]</td>
<td>Latvia</td>
<td>1.18</td>
<td>[31]</td>
</tr>
<tr>
<td>Greece</td>
<td>3.6</td>
<td>[38]</td>
<td>UK</td>
<td>3.3</td>
<td>[7]</td>
</tr>
</tbody>
</table>

The presence of potential correlation of the HCWGR and selected economic, social, and environmental indices of each country was examined. Specifically, the 12 indices that were used in this study are explained below:

- **GDP per capita (US $/capita).** This is the gross domestic product (GDP) converted to dollars using purchasing power parity rates. GDP at purchaser’s prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products [39]. Data were based on the 2015 calendar year.

- **Health expenditure (HE) or healthcare spending per capita (US $/capita).** Total health expenditure is the sum of public and private health expenditures. It is a percentage of the GDP and was expressed here in $ per capita [39]. Data were based on the 2014 calendar year.

- **Human Development Index (HDI).** The HDI was created to emphasize that people and their capabilities should be the ultimate criteria for assessing the development of a country, not economic growth alone. The Human Development Index (HDI) is a summary measure of average achievement in key dimensions of human development, i.e., a long and healthy life, being knowledgeable, and having a decent standard of living. The HDI does not reflect on inequalities, poverty, human security, empowerment, etc. [40]. Data were based on the 2014 calendar year.

- **Inequality-adjusted Human Development Index (IHDI).** The IHDI combines the country’s average achievements in health, education and income with how those achievements are distributed among the country’s population by “discounting” each dimension’s average value according to its level of inequality. Under perfect equality, the IHDI is equal to the HDI, but falls below the HDI when inequality rises [40]. Data were also based on the 2014 calendar year.

- **Multidimensional Poverty Index (MPI).** The index identifies the number of people who are multi-dimensionally poor and the number of deprivations with which poor households typically strive [40]. Note that MPI refers to developing countries only, since there are no relevant data for developed countries. Data were based on available values from different calendar years and were available for 21 countries only (European countries had no MPI).
• **Life expectancy (LE) at birth (years).** This is the number of years that a newborn infant could expect to live if prevailing patterns of age-specific mortality rates at the time of birth stay the same throughout the infant’s life [40]. Data were based on the 2014 calendar year.

• **Mean years of schooling.** Average number of years of education received by people of ages 25 and older, converted from education attainment levels using official durations at each level [39]. Data were based on the 2014 calendar year.

• **HIV prevalence, adult (% ages 15–49).** Percentage of the population (at ages 15–49) who are living with HIV [40]. Data were based on the 2013 calendar year.

• **Deaths due to tuberculosis (per 100,000 people).** Number of deaths due to tuberculosis from confirmed and probable cases, expressed per 100,000 people [40]. Data were based on the 2012 calendar year.

• **Deaths due to malaria (per 100,000 people).** Number of deaths due to malaria from confirmed and probable cases, expressed per 100,000 people [40]. Data were based on the 2012 calendar year.

• **Under-five mortality rate (per 1000 live births).** Probability of dying between birth and the age of five, expressed per 1000 live births [40]. Data were based on the 2013 calendar year.

• **CO\textsubscript{2} emissions (annual metric tonnes per capita).** Carbon dioxide emissions (CDE) are those stemming from the burning of fossil fuels and the manufacturing of cement. They include carbon dioxide produced during consumption of solid, liquid, and gas fuels and gas flaring [40]. According to Human Development Reports, there are eight indices which indicate the environmental sustainability for each country worldwide. One of them is CO\textsubscript{2} emissions per capita, which was chosen as a representative environmental sustainability index. Data were based on the 2011 calendar year.

### 2.2. Methods

The dependent variable in this work was the HCW generation rate (in kg/bed/day), whilst the socio-economic and environmental indices were treated as independent variables. Linear multiple regression modeling was employed to investigate the correlations among HCWGR and the selected indices. The most complicated model was first fitted to the data and then parameters were sequentially removed until the best reduced model was reached based on the methodology of Berthouex and Brown ([41], pp. 337–342). Additionally, the normal distribution of the response and of the independent variables was checked. Pearson and Spearman correlation coefficients were also calculated to further study the pairwise correlations. The statistical analysis was done using Minitab® v.17 (Coventry, UK).

### 3. Results

Figures 1 and 2 illustrate HCWGR and GDP for the 42 countries considered in this work. The data are clustered by continent. It can be roughly observed from these two figures that there seems to be some type of correlation between HCW and GDP. Countries with high GDP seem to have high HCW as well. For example, in Africa, there are low HCWGR, while those same rates are clearly much higher in Europe and in two countries of North America (USA and Canada).
3.1. Descriptive Statistics and Test of Normality

The descriptive statistics related to HCWGR are shown in Table 2.

![HCWGR per country](image1)

**Figure 1.** HCWGR for selected countries worldwide.

![GDP per capita per country](image2)

**Figure 2.** GDP per capita for selected countries worldwide.

<table>
<thead>
<tr>
<th>Continent</th>
<th>Mean ± St. Dev.</th>
<th>Sample Size</th>
<th>Anderson-Darling (AD) Value</th>
<th>p of the AD Normality Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>0.80 ± 0.23</td>
<td>8</td>
<td>0.223</td>
<td>0.738</td>
</tr>
<tr>
<td>America</td>
<td>4.41 ± 3.0</td>
<td>6</td>
<td>0.722</td>
<td>0.028</td>
</tr>
<tr>
<td>Asia</td>
<td>2.44 ± 1.5</td>
<td>18</td>
<td>0.664</td>
<td>0.069</td>
</tr>
<tr>
<td>Europe</td>
<td>3.10 ± 1.1</td>
<td>10</td>
<td>0.577</td>
<td>0.099</td>
</tr>
<tr>
<td>All data</td>
<td>2.57 ± 1.89</td>
<td>42</td>
<td>1.338</td>
<td>&lt;0.005</td>
</tr>
</tbody>
</table>

Table 2. Descriptive statistics of the HCWGR.
The normality of the dependent variable (HCWGR) was examined using both a graphical depiction (boxplots) and by applying the Anderson-Darling (AD) normality test per continent (see Table 2). Figure 3 illustrates the boxplots of the data per continent. The mean values are indicated with the symbol (circled cross). According to Figure 3, the HCWGR values from America and Asia are slightly skewed to the right. Using a threshold $p$ of 0.05, the results from the AD test indicate that values from Africa, Asia, and Europe are normally distributed, whilst values from America and the overall values are not. The lack of normality for the HCWGR from America is practically attributed to the outliers of USA and Canada (since both countries had HCWGR above 8 kg/bed/d). These outliers were not removed from the overall data during subsequent analysis.

![Boxplot for the HCWGR values per continent](image)

**Figure 3.** Boxplot for the HCWGR values per continent (total $n = 42$). The box size is proportional to sample size (Africa: $n = 8$, America: $n = 6$, Asia: $n = 18$, Europe: $n = 10$); symbol indicates the mean value.

### 3.2. Correlations

Using the data described above, correlations were examined between the HCWGR (measured in kg/bed/day) and the 12 indices mentioned in Section 2.1. The correlations were studied both by following a graphical procedure (Figures 4–6), as well as by calculating the Pearson and Spearman correlation coefficients (Table 3). Linear regression was employed to examine the relationship between the HCWGR and the independent indices, separately. Finally, the best reduced model was developed to accurately express the correlation of the HCW generation rates with the statistically significant indices from the ones mentioned above. The graphical depiction of the data is included in Figures 4–6.

Figures 4–6 group the values per continent using a different color. However, the regression line shown was constructed from all data values ($n = 42$). The linear correlations per continent are included in Table 3, where both Pearson (linear) and Spearman (ranking based) correlations are shown.

Figures 4–6 reveal that there is a positive correlation between HCW generation rates and GDP, HE, HDI, IHDI, life expectancy, mean years of schooling, and CDE ($R^2 = 0.35, 0.359, 0.375, 0.332, 0.312, 0.362, \text{ and } 0.574$, respectively). With regard to GDP and HE (which is a percentage of GDP), a rather reasonable explanation for this positive correlation can be provided. As the quality of healthcare services increases, more medical equipment and medical consumables are purchased, more patients are treated and, thus, more HCW are generated. Similar findings had been found by Windfeld and Brooks (2015) [7], who had also showed a positive correlation between HCWGR and the GDP and HE. In that case, the coefficients of determination were statistically significant and equal to 25.7% and 34.9%, respectively.
Figure 4. Linear regression models of HCWGR vs. (a) GDP per capita; (b) health expenditure per capita; (c) HDI; and (d) IHDI. The regression line was constructed from all data.

Figure 5. Linear regression models of HCWGR vs. (a) MPI; (b) life expectancy at birth; (c) mean years of schooling; and (d) HIV prevalence. The regression line was constructed from all data.
(MPI, tuberculosis deaths, malaria deaths, U-5 mortality rates) can be explained by the fact that weak poorer (i.e., increase of MPI), HCWGR decrease. This is also marginally visually evident from Figure 5a.

Prevalence, Deaths due to tuberculosis, Deaths due to malaria and U-5 mortality rate (R$^2$ 2017 not significant at p < 0.1). Nevertheless, the negative correlation with the other four indices (MPI, tuberculosis deaths, malaria deaths, U-5 mortality rates) can be explained by the fact that weak health services and mismanagement are observed in the poorer (developing) countries in which the healthcare facilities operate scarcely and inefficiently. The negative correlation between HCWGR and the above four indices is also concluded from the Pearson and Spearman correlations (Table 3), that have all negative signs (except the Pearson coefficients for MPI and Deaths due to malaria, which are not significant at p < 0.01).

The negative significant Spearman correlation coefficient for MPI indicates some type of negative non-linear correlation between MPI and HCWGR, which practically indicates that as a country gets poorer (i.e., increase of MPI), HCWGR decrease. This is also marginally visually evident from Figure 5a.
Table 3. Pearson and Spearman correlation coefficients between HCWGR and each of the 12 indices.

<table>
<thead>
<tr>
<th>Index</th>
<th>Correlated with HCWGR</th>
<th>Africa</th>
<th>America</th>
<th>Asia</th>
<th>Europe</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>ns, [ns] n = 8</td>
<td>0.985 [1.000] n = 6</td>
<td>ns, [0.688] n = 18</td>
<td>ns, [ns] n = 10</td>
<td>0.592 [0.699] n = 42</td>
<td></td>
</tr>
<tr>
<td>HE</td>
<td>ns, [ns] n = 8</td>
<td>0.939 [0.943] n = 6</td>
<td>ns, [0.725] n = 17</td>
<td>ns, [ns] n = 10</td>
<td>0.599 [0.687] n = 41</td>
<td></td>
</tr>
<tr>
<td>HDI</td>
<td>ns, [ns] n = 8</td>
<td>ns [1.000] n = 6</td>
<td>ns, [0.678] n = 18</td>
<td>ns, [ns] n = 10</td>
<td>0.612 [0.671] n = 42</td>
<td></td>
</tr>
<tr>
<td>IHDI</td>
<td>ns, [ns] n = 6</td>
<td>ns, [ns] n = 6</td>
<td>ns, [0.656] n = 16</td>
<td>ns, [ns] n = 10</td>
<td>0.576 [0.652] n = 38</td>
<td></td>
</tr>
<tr>
<td>MPI</td>
<td>ns, [ns] n = 6</td>
<td>ns, [ns] n = 6</td>
<td>ns, [ns] n = 12</td>
<td>n = 0</td>
<td>ns [-0.616] n = 21</td>
<td></td>
</tr>
<tr>
<td>LE</td>
<td>ns, [ns] n = 8</td>
<td>ns, [ns] n = 6</td>
<td>ns, [ns] n = 18</td>
<td>0.824 [0.840] n = 10</td>
<td>0.559 [0.687] n = 42</td>
<td></td>
</tr>
<tr>
<td>Mean years of schooling</td>
<td>ns, [ns] n = 8</td>
<td>0.959 [0.943] n = 6</td>
<td>ns, [ns] n = 18</td>
<td>ns, [ns] n = 10</td>
<td>0.601 [0.633] n = 42</td>
<td></td>
</tr>
<tr>
<td>HIV prevalence</td>
<td>ns, [ns] n = 8</td>
<td>ns, [ns] n = 3</td>
<td>ns, [ns] n = 10</td>
<td>ns, [ns] n = 3</td>
<td>ns, [ns] n = 24</td>
<td></td>
</tr>
<tr>
<td>Deaths tuberculosis</td>
<td>ns, [ns] n = 8</td>
<td>ns, [ns] n = 6</td>
<td>ns, [−0.635] n = 18</td>
<td>ns, [ns] n = 10</td>
<td>−0.466 [−0.596] n = 42</td>
<td></td>
</tr>
<tr>
<td>Deaths malaria</td>
<td>ns, [ns] n = 5</td>
<td>ns, [ns] n = 3</td>
<td>ns, [−0.728] n = 13</td>
<td>n = 0</td>
<td>ns [−0.656] n = 21</td>
<td></td>
</tr>
<tr>
<td>U-5 mortal. rate</td>
<td>ns, [ns] n = 8</td>
<td>ns, [ns] n = 6</td>
<td>ns, [−0.604] n = 18</td>
<td>ns, [ns] n = 10</td>
<td>−0.498 [−0.650] n = 42</td>
<td></td>
</tr>
<tr>
<td>CDE</td>
<td>ns, [ns] n = 8</td>
<td>0.987 [0.943] n = 6</td>
<td>ns, [0.718] n = 18</td>
<td>ns, [ns] n = 10</td>
<td>0.758 [0.727] n = 42</td>
<td></td>
</tr>
</tbody>
</table>

First value in each column is the Pearson linear correlation coefficient; values in brackets are the Spearman rank order correlation coefficients; the last value indicates the sample size from which the correlation was based on; only the coefficients that are statistically significant at $p < 0.01$ are presented.
With regard to the correlations per continent, Table 3 reveals that there is no correlation in Africa between HCWGR and all indices, since the Pearson and Spearman coefficients are not statistically significant in any of the cases. Moreover, the Pearson and Spearman coefficients between HCWGR and all indices are also not significant in Europe, except in the case of life expectancy at birth (LE), in which there is positive correlation between HCWGR and LE (Pearson coeff.: 0.824, Spearman coeff.: 0.84).

In the case of the Americas, there is positive correlation between HCWGR and GDP, HE, mean schooling years, and CDE. However, there is no correlation between HCWGR and the other eight indices (HDI, IHDI, MPI, LE, HIV prevalence, deaths due to tuberculosis and malaria, U-5 mortality rate), since the Pearson and Spearman coefficients are not statistically significant.

Finally, the Pearson coefficients from the correlations between HCWGR and all indices in Asia are not significant either. Nevertheless, the Spearman coefficients from the correlations between HCWGR and eight indices (GDP, HE, HDI, IHDI, deaths due to tuberculosis and malaria, U-5 mortality rate, and CDE) in Asia reveal some correlation (positive correlation: GDP, HE, HDI, IHDI, CDE—negative correlation: deaths due to tuberculosis and malaria, U-5 mortality rate). Instead, the Spearman coefficients from the correlations between HCWGR and MPI, LE, mean schooling years and HIV prevalence are not significant.

It should be mentioned that Asia presents the highest HCWGR fluctuations whereas, in Europe, the HCWGR have a lower variability. Table 3 includes all of the Pearson linear and Spearman rank coefficients (per continent and overall) that are statistically significant at \( p < 0.01 \).

### 3.3. Multiple Linear Regression Modeling

An attempt was made to describe the HCWGR through multiple regression modeling. The full empirical linear model that was developed was in the form of:

$$\text{HCWGR} = \text{constant} + a \times X_1 + b \times X_2 + c \times X_3 + \ldots + n \times X_n$$  \hspace{1cm} (1)

where:
- **HCWGR**: the Health-Care Waste Generation Rate in kg/bed/day
- **Constant**: a constant in kg/bed/day
- **a, b, c, \ldots, n**: coefficients
- **X1, X2, \ldots, Xn**: independent variables (predictors). Eight of the twelve parameters (health expenditure, HDI, CDE, LE, schooling years, tuberculosis induced deaths, malaria induced deaths, under-five mortality rate) were used during modeling.

GDP was not included in the model, due to its direct correlation with HE. In addition, the HIV prevalence was also removed, since Table 3 showed that this factor alone does not affect HCWGR at all. IHDI was also removed from modeling due to its apparent correlation with HDI. Finally, MPI was also not included in the model, since, according to Table 3, it does not correlate linearly with HCWGR and there are several lacking data (sample size \( n \) is 21 for MPI). A best reduced model was fit by starting developing the full model (with all eight predictors and a constant) and then by sequentially removing terms. The goal of this process was to remove all non-statistically significant terms and to reach a best reduced model (BRM) that still describes adequately the data with the fewest, yet, significant parameters [41] and with the highest \( R^2 \). The data from all 42 countries combined were used in the modeling. By following the above process, the resulting best reduced model is shown in Equation (2). Equations (3) and (4) are additional models with slightly lower \( R^2 \) values than that of Equation (2), but with different parameters. Equations (2)–(4) could be interchangeably used depending on the availability of the shown indices (parameters) in each country.

\[
\begin{align*}
\text{HCWGR} & = 0.014 \ [0.0042] \ \text{LE} + 0.31 \ [0.047] \ \text{CDE} & \text{Adjusted } R^2 & = 84.73\% \\
\text{HCWGR} & = 1.5 \ [0.47] \ \text{HDI} + 0.29 \ [0.053] \ \text{CDE} & \text{Adjusted } R^2 & = 84.52\% 
\end{align*}
\]
HCWGR = 0.13 [0.048] SCH_Y + 0.278 [0.066] CDE

Adjusted $R^2 = 83.64\%$ 

where:

LE: Life expectancy (in years),
HDI: Human Development Index, as defined earlier,
SCH_Y: Mean years of schooling (years), and
CDE: CO\(_2\) emissions in tonnes per capita per year.

Values in brackets are the standard errors of the corresponding coefficients. All coefficients were statistically significant at $p \leq 0.003$, and at $p \leq 0.01$, for the SCH_Y coefficient of Equation (4).

The above BRMs contain no constant, which was proven to be statistically non-significant in all cases. Equation (2), which is the best of the three models since it obtained the highest $R^2$, reveals that the two factors that mostly affect HCWGR are the life expectancy (LE) and the total CO\(_2\) emissions (CDE) of a country. Both of those factors positively affect HCWGR, since, as they increase, HCWGR increases, too. The variability explained by this model is around 85%. Equation (2) reveals that as life expectancy increases, HCW increases, too, due to a likely improvement of the health services needed to achieve the high LE. Similarly, CDE, which are directly associated to health care management systems [42], appears as an additional good predictor of HCWGR. It is interesting to note that CDE was statistically significant (with a $p$ always less than 0.0001) in all three equations with a coefficient equal to around 0.3 in all cases. Therefore, CDE is the most statistically significant term among all model terms. According to Equations (3) and (4), HDI and schooling years positively affect HCWGR, as had been revealed from the correlation coefficients of Table 3, as well. That is, the more developed a country is, the better its medical system becomes, which eventually leads to the generation of higher medical wastes compared to less developed countries.

4. Discussion and Conclusions

The main conclusions from the present work are the following:

- The practical application of the work is that certain socio-economic and environmental indices per country can be used to mathematically predict HCWGR so as to avoid direct and costly HCW weight measurements.
- A positive correlation between HCWGR and seven of the twelve indices (GDP, HE, HDI, IHDI, life expectancy, mean years of schooling, and CDE) was observed.
- A negative correlation between HCWGR generation rate and four of the twelve indices (MPI, HIV prevalence, deaths due to tuberculosis, deaths due to malaria, and U-5 mortality rate) was observed.
- Using the Pearson and Spearman correlation coefficients, it was found that the HIV prevalence was not a statistically significant predictor of the HCWGR.
- Based on multiple linear regression modeling, the resulting best reduced model indicated that life expectancy and carbon dioxide emissions positively affect healthcare waste generation and can be used as predictors to adequately describe HCWGR (see Equation (2)). The resulting empirical multiple regression model explained 85% of the variability of the response. Two additional models, Equations (3) and (4), showed that HDI and mean years of schooling can be also used as HCWGR predictors.
- The annual CO\(_2\) emissions was the index that affected the HCWGR the most.
- More factors should be investigated in future work to try to augment and validate the proposed regression model and to incorporate principal component analysis to separate and group the significant predictors. In addition, efforts should be made to distinguish the hazardous fraction from the total HCW and to develop similar modeling with the former fraction as well.
Author Contributions: Minas Minoglou and Spyridoula Gerassimidou collected the original data, performed some of the statistical analysis and wrote parts of the paper. Dimitrios Komilis conceived the idea, provided guidance, finalized the statistical analysis, and had the overall supervision during the writing of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References


© 2017 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).