2005

An Empirical Study Identifying the Factors that Impact eHealth Infrastructure and eHealth Use

Kallol Bagchi
University of Texas at El Paso, kbagchi@utep.edu

Godwin Udo
University of Texas at El Paso, gudo@utep.edu

Mark Kesh
University of Texas at El Paso, mkeh@utep.edu

Follow this and additional works at: http://aisel.aisnet.org/amcis2005

Recommended Citation
http://aisel.aisnet.org/amcis2005/295

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2005 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
ABSTRACT

This cross-national study examines factors that impact eHealth infrastructure and usage growth. The independent variables in the study are: confidence in health care system, perception of one’s state of health, per capita national expenditure in health, % of people belonging to voluntary organizations concerned with health, human development index, cost of health care per capita, and cost of health care per GDP (data gathered from World Value Survey (WVS), World Bank, United Nations, and World Health Organization (WHO) survey results). Data from a set of more than 40 nations involving nearly 8000 sites were gathered to allow analysis of factors encouraging the eHealth growth. The set of nations includes both developed and developing nations. Preliminary results suggest that confidence in health care system, voluntary health groups, GDP, human development, cost of health care per capita are significant in eHealth infrastructure -- explaining 32-54% of the variability of the eHealth infrastructure. Preliminary results also suggest that human development, infrastructure and perception of health are significant in eHealth usage -- explaining 45-52% of the variability of the eHealth usage.

Keywords

Internet, eHealth, infrastructure, cross-national study, world value survey, health information access, telemedicine.

INTRODUCTION

The Internet and Web technologies which were first introduced to the public in the 1990s have exploded all over the world (Berners-Lee, 1998; WWW, 1997). The advent of the Internet has greatly improved the creation and dissemination of information on public health not only to the residents of developed nations but also to those in developing countries. Information technology in general and the Internet/Web in particular, has drastically affected the ways and manners health information are collected, stored and distributed for use by both healthcare providers and consumers, which in turn affects the public policy and the health of the public (Fox and Rainee, 2000, Kaplan and Brennan, 2001; Kiley, 2000). This phenomenon is referred to as “eHealth.” In this article we focus on Internet/Web based eHealth systems. The health information management has rapidly turned into a high-value and growing market within the healthcare industry. According to Rodrigues and Risk (2003) the potential benefits of eHealth are very attractive to governments in many parts of the world and have led to specific actions by public health policy makers in the following areas:

- promotion of education, training, and national planning capacity in information systems and technology;
- convening groups for the implementation of standards;
- providing funding for research and development;
- ensuring the equitable distribution of resources, particularly to places and people considered by private enterprise to provide low opportunities for profit;
- protecting rights of privacy, intellectual property, and security; and
- overcoming the jurisdictional barriers to cooperation, particularly when there are conflicting regulations.
Bagchi et al.  
Factors Impacting eHealth

Empirical research on factors that influence eHealth growth is scanty. In this paper, we will try to find an answer to the research question: What drives the eHealth infrastructure of a nation and its usage? In other words, what are the factors affecting the development of eHealth infrastructure and its usage. Specifically, what roles do the following factors play: GDP, Internet use, voluntary health groups, health care expenditure by GDP, perception of health and other factors?

THE EHEALTH INFRASTRUCTURE AND USAGE MODEL

Figure 1 shows how various factors are impacting the eHealth infrastructure and eHealth usage. The dotted arrows denote the impact of control variables. The eHealth infrastructure of a nation is the set of eHealth related web sites in that nation. The eHealth usage can be defined as using/accessing these web sites for information. For example, recent studies on European nations (SIBIS, 2003; BISER, 2004) found that people of a nation mainly search the Internet for three specific types of health-related information: lifestyle information (e.g. diets, nutrition and fitness), information about particular illnesses, treatments or medication, and information about health services (e.g. what services are available and where and when they are available). The study variables which are shown in Figure 1 are described in the following section.

Internet Users Per Thousand

The eHealth infrastructure and usage in a nation directly depend on the Internet users per 1000. High Internet penetration or use of a nation may not necessarily translate to high Internet access of all people in that country; however low Internet users per 1000 of a nation probably mean that eHealth infrastructure and use can not be high in that nation. The Internet use, although growing rapidly, is still evolving. One of the highest users of the Internet is the U.S. The Internet users in year 2002 in U.S. were barely 55.14% of the total population and the household Internet penetration was 49.9% of the total households in 2001 (Dutta et al., 2004). The condition in developing nations is far worse. When we control for population, underdeveloped nations with large populations such as India, China and Brazil may actually show a negative relationship between Internet users per 1000 and health sites. These nations may have a large number of health-related web sites, which will pale into insignificance when population is considered as a control variable, especially since population may play a role in the size of number of web sites.

Eng et al., have observed that the Internet has opened the door for an average user to visit many places to gather information on health-related issues (Eng et al., 1998; Eng 2001). Thus, the number of Internet users per 1000 could be positively related to the number of health-related web sites. It has been observed that an increase in Internet-based health information access has resulted in recent times due partly to Internet-saavy consumers (Sieving, 1999).

In a study involving European nations/regions, it was found that the propensity to search for health-related information by Internet users in a region is strongly associated with the level of Internet penetration in that region (a Rsq. Value of 0.49). Additionally, a similar (though weaker) association was found across various European countries in an earlier SIBIS study (SIBIS 2003; BISER, 2004).

H1. The higher the number of Internet users per thousand of the publics of a nation, the higher the eHealth infrastructure and eHealth use of that nation

Economy or Personal Wealth

Previous researchers have noted that higher income is related to increased use of eHealth infrastructure (Diaz et al., 2002; Houston et al., 2002). Groot and Brink (2003) observed that health concerns tend to increase with economic development. For example, it has been observed that total pharmaceutical expenditure, as well as other health expenditure, is linked to the economic development level of a country, and tends to increase only when GDP increases. On the other hand, even in some rich nations such as the U.S. a large number of people are without Internet access, which translates to without direct e-health access. Because of better infrastructural facilities in nations with high GDP, many people can still use the Internet/Web from the public libraries. It has been observed that the poorest countries have the worst access to information and communication and thus to eHealth infrastructure and use (UNDP, 1999). Eysenbach et al.(2001) have observed that one fundamental problem of telemedicine and using the Internet for health education is that poor people who are at highest risk of preventable or treatable health problems have the greatest need for information and are the least likely to have access to such technologies.
The SIBIS report (2003) notes that almost one in five (19.8%) of the European Union population aged 15 years and over reported searching online for health-related information in the 12 months before the survey. However, it is a lot lower than the more than two in five (44.9%) of the US sample who reported this form of eHealth activity. A nation’s economy was found to be related to this kind of eHealth access, although not strongly (a Rsq value of .07) (SIBIS, 2003, BISER, 2004). We postulate that better economy of a nation will play a positive role in the eHealth infrastructure of a nation.

**H2. The stronger the economy of a nation as indicated by GDP per capita, the higher the eHealth infrastructure and eHealth usage of that nation.**

**Health Expenditure**

Health care expenditures are on the rise for many nations. For example, the US healthcare expenditures are expected to double from 1998 to 2007 (Health Industry Today, 1999). Internet provides a low cost universal access to data. As a result, eHealth infrastructures have mushroomed in different nations. Preventive health strategies and reduction of costs are two reasons that may partly be responsible for such growth in eHealth infrastructure (Building Bridges, 2002). Thus our next hypothesis follows:

**H3: The higher the health care expenditure of a nation, the better the eHealth infrastructure and the higher the eHealth usage.**

**Perception of Health**

Nations where perceptions of good health exist, many web sites also exist that give advices on health nutrition and fitness related topics. As an illustration, many links to such web sites exist and can be found in the web site in U.S. ([http://www.pecentral.org/websites/healthsites.html](http://www.pecentral.org/websites/healthsites.html)). It can also be counter argued that if a person has good health, he/she will be less inclined to use eHealth and that sick people will be using eHealth more. Preliminary evidence suggests to the contrary. Houston et al., (2002) reported that their research indicated that majority of eHealth users who participated in their
research project were in good health; however, those individuals with apparent illnesses were more-frequent users of the health information. Many people use eHealth for maintaining a high lifestyle consistent with high standards of living in industrialized nations. Over one third of Internet users in the EU (36.4%) reported online searching for health-related information during the 12 months reference period. Almost three in five (58.3%) respondents of the US sample reported this form of eHealth activity. Within Europe, the prevalence of reported online health information seeking amongst Internet users varied considerably across the countries, with highest rates in Ireland (48.1%) and lowest rates in Greece (21.6%) (SIBIS, 2003).

Although its importance varies across countries, online search for health information is clearly becoming a significant element of the health-related activities of the population. According to the BISER survey (2004), both searching for practical information about service availability, and searching for health-related lifestyle information is more related to specific contextual or cultural factors in the regions/countries. A variable is taken from the World Value Survey (WVS) to measure the perception of health (Inglehart et al., 2004). The question asked was: All in all, how would you describe your state of health these days? Would you say it is: (The measure is % Very good/Good)?

**H4:** The better the perception of good health of a nation, the higher the eHealth infrastructure and lower the eHealth usage.

### Confidence in Health Care Systems

Americans’ satisfaction with the quality of medical care they personally have received remained stable. However, dissatisfaction with the health care system as a whole and concerns are growing among Americans in their ability to afford quality health care in the future, according to the 2004 Health Confidence Survey (HCS, 2002). Confidence in health care in nations such as Canada is also eroding. (HealthCanada, 2004). As the confidence in health care sags (especially in local systems), it is expected that publics will seek health information on Web sites, educate themselves with alternatives and various other health information. As an example, in Italy, people from southern regions travel considerable distances for care because of the dissatisfaction with the quality of care obtainable in the home or nearby regions and one of their sources to obtain this information on better health care is Internet based eHealth systems (BISER, 2004). This indicator may not impact eHealth infrastructure.

A variable is taken from the WVS to measure public confidence in health care. The question asked was: Please look at this card and tell me, for each item listed, how much confidence you have in them, is it a great deal, quite a lot, not very much or none at all? Health care system (The measure is: a great deal/quite a lot (%)).

**H5:** The lower the confidence in health care systems of a nation, the better the eHealth usage.

### Voluntary Organizations on Health

Voluntary organizations on health represent “alternatives, options, experimentation, supplementation, and leadership in the provision of various services and support mechanisms” (O’Connell, 1996). Voluntary organizations can play a major role in rural areas or in tribal or minority health where many facilities for health-related activities do not exist. Many voluntary organizations create and maintain web sites for educating the public at large. On the other hand, it can be argued that voluntary organizations exist primarily to complement absence of government activities in promoting health and so infrastructure and usage of eHealth is low in those nations. On balance we think the role of such voluntary organizations is supplementary to governmental activities and thus their impacts are positive on eHealth infrastructure and usage. The variable is taken from the WVS. The question asked was: Please look at the following list of voluntary organizations and activities and say which if any, do you belong to: voluntary organizations concerned with health. The measure is: % belongs to.

**H6:** The higher the number of voluntary health organizations of a nation, the better the eHealth infrastructure and higher the eHealth usage.

### The Human Development Index
The human development index (HDI) denotes the country’s achievement in human development (which includes economy and education) and its index include a factor related to health: life expectancy. The higher the life expectancy, the better the HDI value (UNDP, 2001). It is expected that eHealth infrastructure and use will be more prevalent in nations with high HDI values as better longevity means better health care systems which in present day translates to eHealth infrastructure and access (Sen, 1999).

**H7: The higher the human development of a nation, the better the eHealth infrastructure and higher the eHealth usage.**

### eHealth Infrastructure and Usage

Finally the usage of e-Health will be dependent on e-Health infrastructure. The better the eHealth infrastructure, the more is its use. It is obvious that without a proper development of infrastructure, the eHealth use can not flourish.

**H8. The e-health usage is higher for nations with better e-health infrastructures.**

### DATA AND RESULTS

Table 1 details the variables used in this study and their sources. The eHealth infrastructure measure was obtained from the search “~health” or “~health care” of websites that exist on the Web on any given nation. The search can capture all identical or similar words starting with “health” or “health care”. The search engine used was Google (Google, 2005). We selected web groups containing websites only belonging to a nation for this study. We used another search machine (A9) for comparing the results with Google. Since the correlation results for a few nations were very high from two search engines, we decided to keep the results from the Google search engine. We also checked search results from two different periods of time (weeks apart) using the Google search engine and statistical differences between these results were insignificant.

<table>
<thead>
<tr>
<th>Variable/Year</th>
<th>What the Variable Denotes</th>
<th>Source of Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDI</td>
<td>Human Development Index, 1999</td>
<td>United Nations</td>
</tr>
<tr>
<td>INTNT</td>
<td>Internet Users per 1000 in 2001</td>
<td>World Bank, 2005</td>
</tr>
<tr>
<td>GDP</td>
<td>GDP per capita in 2002</td>
<td>World Bank, 2005</td>
</tr>
<tr>
<td>USAGE</td>
<td>Ranking average of first 200 Health web sites usages from Alexa, 2005</td>
<td>Alexa Ranking</td>
</tr>
<tr>
<td>INFRASTR</td>
<td>The number of Health Web sites of a nation, 2005</td>
<td>Google Search</td>
</tr>
<tr>
<td>HHCOSTGDP</td>
<td>Health Care cost Per GDP $</td>
<td>WHO, 2000</td>
</tr>
<tr>
<td>VOLUNT</td>
<td>Belongs to Voluntary groups concerned with Health and do unpaid voluntary work</td>
<td>World Value Survey, 2004</td>
</tr>
<tr>
<td>GRHLT</td>
<td>Belongs to Voluntary groups concerned with Health</td>
<td>World Value Survey, 2004</td>
</tr>
<tr>
<td>HPERCEP</td>
<td>Average of respondents who considered their state of health (% Very good/good) in 1990 and 2000</td>
<td>World Value Survey, 2004</td>
</tr>
<tr>
<td>HCONF</td>
<td>Confidence in Health Care (% responding: A great deal/quite a lot)</td>
<td>World Value Survey, 2004</td>
</tr>
</tbody>
</table>

Table 1. The Variables, Their Meanings and their sources

Alexa (2005) is a web service that gives free use to the information gathered by its Web Crawl feature. The service includes more than 100 terabytes of data from over 4 billion web pages, as of October, 2004 (Internetweek, 2004). For measuring e-Health usage, we took the first 200 web sites of each country as ranked by Alexa web services and calculated the average ranking of health web sites from each nation. To keep things manageable, web sites managed from within a nation were only considered. A ranking with low values is ordered high and vice versa. This kind of Web ranking based research has been
done before (Abernethy and Reichgelt, 2003). The authors used the Alexa ranking scheme to determine a measure of "passive" Web participation of a nation. In our study, we tried to consider all major concerns raised by these authors in such ranking schemes. Additionally, we increased our sample size for each nation to 200, thereby minimizing the errors that may arise from such rankings. Examples of information that can be accessed are site popularity, related sites, detailed usage/traffic statistics and site contact information. We used site popularity as the measure for web site ranking. For forty nations, we used approximately 8000 web sites to calculate the average of rankings of each of the forty nations.

The human development index (HDI) value of year 1999 was obtained from the United Nations Development Programme report (UNDP, 2001). The three primary components in this index are GDP per capita, knowledge, and longevity (i.e., life expectancy). We also used the WVS report of 2004 to select most of the health-related variables in the study. The Internet users per 1000 measurement (INTNT) was obtained from the World Bank database as well as health costs per GDP (HCOSTGDP). The data for this study came from four separate sources: WVS, World Bank, United Nations, and WHO but the methods of data collection and populations are very similar across all sources. All sources used tested survey instruments to sample the perception of participants. Therefore, we do not expect the issue of multiple sources of data to be of a great concern.

The WVS variables are detailed in Table 1. Two indicators, VOLUNT and GRHLT are similar but not exactly identical variables, measuring the voluntary health activities. VOLUNT measures active voluntary work whereas GRHLT measures passive membership in voluntary societies. Other WVS variables are self-explanatory.

### Table 2. The PearsonCorrelation Table

Most of the correlations of usage and infrastructure with eHealth variables were significant and in predicted directions and thus most of the hypotheses were supported (Table 2). Human Development Index was not correlated with eHealth infrastructure and confidence in health care systems (HCCONF) was not correlated with eHealth infrastructure or usage. Note that usage correlation results are mostly negative, as usage was measured reversely. Some indicators were not significant and so were eliminated in an earlier regression, to keep the regression parsimonious. Refer to Table 3. The regression results were checked for problems such as multi collinearity, autocorrelation etc (Netter et al., 1996). The multi collinearity was estimated by VIF values. If VIF values of variables are less than 10, that can be regarded as absence of serious multi collinearity in regression. Since some of the selected indicators were strongly correlated with each other, including these in a single...
Regression model would give rise to serious multi collinearity problem in results. So several separate regression models were constructed with these indicators and these were free from multi collinearity. Autocorrelation was measured by Durbin-Watson statistics (DW Stat). The power of regression got reduced in models 1, 2 and 4, due to small sample size emanating from missing data. We used step regression after introducing the controls (Internet Users, Population, GDP).

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTNT (H1)</td>
<td>-0.21</td>
<td>-0.38</td>
<td>-0.19</td>
<td>-0.53</td>
<td>0.07</td>
</tr>
<tr>
<td>POP (Control)</td>
<td>0.45***</td>
<td>.035**</td>
<td>0.20</td>
<td>-0.44**</td>
<td>-0.28*</td>
</tr>
<tr>
<td>GDP (H2)</td>
<td>0.54*</td>
<td>0.67***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCOSTGDP (H3)</td>
<td>0.70**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPERCEP (H4)</td>
<td>0.54*</td>
<td>0.67***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCONF (H5)</td>
<td>-0.31*</td>
<td>-0.33*</td>
<td>0.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOLUNT (H6)</td>
<td>0.44**</td>
<td>0.49***</td>
<td>0.50**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDI (H7)</td>
<td>0.46*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INFRASTR (H8)</td>
<td>-0.17</td>
<td>-0.44***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>25</td>
<td>24</td>
<td>35</td>
<td>25</td>
<td>39</td>
</tr>
<tr>
<td>Adj Rsq</td>
<td>1.11-4.22</td>
<td>1.2-4.9</td>
<td>2.0-4.0</td>
<td>1.6-9.3</td>
<td>1.4-3.1</td>
</tr>
<tr>
<td>VIF</td>
<td>2.02</td>
<td>1.98</td>
<td>1.71</td>
<td>1.71</td>
<td>1.88</td>
</tr>
<tr>
<td>Dep. Variable</td>
<td>INFRASTR</td>
<td>INFRASTR</td>
<td>INFRASTR</td>
<td>USAGE</td>
<td>USAGE</td>
</tr>
</tbody>
</table>

Note: ***: p<.000, **: p<.01, *:p<.1

Table 3. Results of Infrastructure and Use Regressions

For the infrastructure regression (model 1), after controlling for population, Internet users and GDP per capita, we found that confidence in healthcare (HCONF) and voluntary work (VOLUNT) were significant in infrastructure regression (N=25). The adj. Rsq value was 0.51. For the infrastructure regression (model 2), after controlling for population, Internet users, we found that health care cost (HCOSTGDP), confidence in healthcare (HCONF) and voluntary work (VOLUNT) were significant in infrastructure regression (N=24). The adj. Rsq value was 0.54. In both models 1 and 2, the sign of HCONF was negative, suggesting that less confidence in healthcare translated to design of more health-related web sites. For the next infrastructure regression (model 3), after controlling for population, we found that perception of health (HPERCEP) was insignificant whereas confidence in health care systems (HCONF) and human development index (HDI) were significant in the infrastructure regression (N=35). The adj. Rsq value was 0.32. Model 1 and model 2 are free from autocorrelation. Model 3 shows that the results may not be free from autocorrelation at 5% significance level (could be due to an omitted variable) but may be free from autocorrelation at 1% significance level (the null hypothesis of no autocorrelation is not rejected).

We next report the regression on eHealth usage (the last two columns of Table 3). In model 4, after controlling for population, Internet use and GDP, we found that perception of health was significant whereas confidence in healthcare and infrastructure were insignificant in eHealth usage. The adj. Rsq value was .45 (N=25). Next, we controlled for population, Internet usage and found that in model 5, the perception of health was insignificant and HDI as well as infrastructure were significant in eHealth usage. The adj. Rsq value was .52 (N=39). We note that in the usage regressions, the signs of HDI, health perceptions and infrastructure were reversed (as shown in Table 3), as usage was measured reversely. The models 4 (at 1% significance level) and model 5 (at 5% significance level) are free from multicollinearity and autocorrelation.

All in all, we found that even after controlling for population, economy and Internet variables, several health-related variables emerged as significant in both infrastructure and usage regressions. Population was always significant, GDP sometimes in the models considered. Thus hypotheses H2-H7 were partly supported in the regressions. Voluntary group work (VOLUNT) emerged as strongly relevant in infrastructure but not in usage regressions. Our control variable population (POP) also emerged as strongly significant in most regressions. Internet use was not at all significant. More research is needed to find out why it was so. At least one economic variable (GDP, HDI, HCOSTGDP) also emerged as significant in most regressions.
Perception of one’s health emerged as significant in usage regression only whereas confidence in healthcare was significant only in infrastructure regressions.

CONCLUSIONS

This preliminary study was a first attempt to find out what factors affect the development of eHealth infrastructure and usage. The study explored the role of several indicators in explaining the development of eHealth infrastructure and its usage. It was found that after controlling for economic variables, population and Internet users, eHealth infrastructure and usage were influenced by health-related variables such as health expenditure per GDP, voluntary health groups, perception of health, and confidence in health care systems. The implications of the study include: (a) If public policy makers decide to spend on health care, the eHealth infrastructure will become well developed in the long run and the populations will use the eHealth systems more; (b) The more the public perceive their health to be very good, the more they are likely to use eHealth systems; (c) Since the results indicate that the populations have less confidence in the health care systems of nations with higher eHealth infrastructures, it implies that the poor nations with sparse infrastructure will benefit more by learning from the pitfalls of the rich nations; (d) The higher the human development index of a nation, the more the populations tend to use eHealth implying that poorer nations (low HDI) are less able to use eHealth; and (e) voluntary health-related groups play a major role in developing eHealth infrastructure. If some of these implications hold, it could then be concluded that investment in eHealth technologies by developed nations will yield a huge payoff for not only their populations but even more so for the populations of less developed nations.

One of the limitations of the study is the fact that the data are two to three years old because the sources of the study data (World Bank, WHO, WVS, and UN) are that much behind in their data collection because of the complexity involved in the process. The findings should be interpreted with caution until more research is done to clearly identify the factors that affect the eHealth infrastructure and use. In the future, the authors also intend to conduct studies to determine the impacts of eHealth on the health of the population.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the helpful comments of the track chairs and reviewers which improved the quality of the manuscript. The authors also acknowledge the help provided by Mr. Jagadish Dandu in running the experiments.

REFERENCES

5. Brennan, P. F. (1999) Health Informatics and Community Health: Support for Patients as Collaborators in Care, Methods Information Medicine, 38, 4-5, 274-278.


