Do Language Fluency and Other Socioeconomic Factors Influence the Use of PubMed and MedlinePlus?

L. Sheets[1, 2]; A. Gavino[2]; F. Callaghan[2]; P. Fontelo[2]

[1]University of Missouri Informatics Institute, Columbia, Missouri, USA; [2]National Library of Medicine, Bethesda, Maryland, USA

Keywords
Library information systems, access to information, diffusion of innovation, communication barriers, economic barriers

Summary
Background: Increased usage of MedlinePlus by Spanish-speakers was observed after introduction of MedlinePlus in Spanish. This probably reflects increased usage of MEDLINE and PubMed by those with greater fluency in the language in which it is presented; but this has never been demonstrated in English speakers. Evidence that lack of English fluency deters international healthcare personnel from using PubMed could support the use of multi-language search tools like BabelMeSH.
Objectives: This study aims to measure the effects of language fluency and other socioeconomic factors on PubMed MEDLINE and MedlinePlus access by international users.
Methods: We retrospectively reviewed server pageviews of PubMed and MedlinePlus from various periods of time, and analyzed them against country statistics on language fluency, GDP, literacy rate, Internet usage, medical schools, and physicians per capita, to determine whether they were associated.
Results: We found fluency in English to be positively associated with pageviews of PubMed and MedlinePlus in countries with high literacy rates. Spanish was generally found to be positively associated with pageviews of MedlinePlus en Español. The other parameters also showed varying degrees of association with pageviews.
Conclusions: After adjusting for the other factors investigated in this study, language fluency was a consistently significant predictor of the use of PubMed, MedlinePlus English and MedlinePlus en Español. This study may support the need for multi-language search tools and may increase access of health information resources from non-English speaking countries.

Correspondence to:
Lincoln R Sheets, MD
Missouri University Informatics Institute
EBW 241, Columbia MO 65211
Email: lincolnsheets@gmail.com

Appl Clin Inform 2013; 4: 170–184
DOI: 10.4338/ACI-2013-01-RA-0006
received: January 23,2013
accepted: March 30, 2013
published: April 10, 2013
Citation: Sheets L, Gavino A, Callaghan F, Fontelo P. Do language fluency and other socioeconomic factors influence the use of PubMed and MedlinePlus? Appl Clin Inf 2013; 4: 170–184
http://dx.doi.org/10.4338/ACI-2013-01-RA-0006

1. Background

In 2010, English ranked first among the top ten Internet languages with over 565 million or 26.8% of worldwide users [1]. It has been considered the *de facto* language of the Internet because websites are mostly in English [2]. Proficiency in English can be a major factor in the digital divide since it directly influences access to the Internet and information technology [2, 3]. On the contrary, a recent study showed that the adoption of English is not a significant contributor in the digital divide [4].

Lower utilization of health care services has been attributed to lack of fluency in English [5]. Studies in the US and elsewhere have demonstrated how lack of language fluency negatively affected patient education [6], compliance to medications [7], and overall clinical experience [8, 9]. This trend is not confined to English but was also observed in French-speaking populations [10]. To alleviate this problem, health facilities have responded by providing professional or *ad hoc* interpreters to assist physicians in communicating with their patients [11], or by training their doctors to learn a second language [12].

Another area in healthcare that can be affected by language skills is access to health information by providers, patients, and the general population. In a study by Leon *et al.*, an increase in usage of MedlinePlus by Spanish-speakers was observed after introduction of MedlinePlus *en Español* [13]. This probably reflects increased usage of health information resources by those with greater fluency in the language. However, there are limited studies demonstrating the relationship between language skills and health information resource usage. While many studies have addressed global economic disparities in access to evidence-based medical resources [14-16], we are not aware of previous efforts to systematically quantify the predictive power of linguistic and other socioeconomic factors.

The National Library of Medicine (NLM), the largest repository of biomedical data worldwide, provides valuable medical information to health professionals through PubMed MEDLINE and to public consumers through MedlinePlus [17]. While most of its resources are in English, NLM launched MedlinePlus *en Español* in 2002 to serve the increasing number of Hispanics in the United States [18, 19]. PubMed collects scholarly citations, abstracts, and articles; while MedlinePlus and MedlinePlus *en Español* provide reliable patient-oriented articles on medical topics.

2. Objectives

This study aims to determine the effect of language on the use of PubMed and MedlinePlus across the world. Discovering country-level correlations between linguistic characteristics and access to these important health information resources may also point out previously undocumented sources of health disparities between populations. Evidence that language barriers deter usage of online health information resources could support the use of multi-language search tools such as BabelMeSH [20]. Furthermore, we wanted to find out if other social determinants, such as the gross domestic product (GDP), Internet use, literacy, and the number of physicians and medical schools in the country, are associated with the utilization of these NLM health resources.

3. Methods

In this study, pageviews by country for PubMed and MedlinePlus were analyzed separately. Pageview counts were anonymized by client except for the root domain, to measure the usage of these resources by country. For example, all pageviews that were resolved to a root domain of .us were counted as being viewed from the United States, and those with a root domain of .jp as being viewed from Japan. PubMed pageview data were provided by the National Center for Biotechnology Information (NCBI) while MedlinePlus pageviews were provided by the National Library of Medicine. To make the values comparable across countries, pageviews per million capita were calculated by dividing each country's pageviews by their populations and multiplying by one million.

For a reliable and independent source of demographic data we chose the Central Intelligence Agency (CIA) World Factbook [21], whose online database is updated weekly. From this source we obtained each country's population, gross domestic product (GDP), literacy rate, and number of Internet users. The number of medical schools and physicians in each country was derived from World Health Organization data [22, 23], although these figures varied in age from 2000 to 2010, depending on each country's reporting frequency. Data on English-speakers and Spanish-speakers by country were obtained from various sources [24-27] of varying age and reliability.

Our primary goal was to investigate the association between the percentage of English/Spanish speakers in a country, and the number of PubMed or MedlinePlus pageviews per million population, after adjusting for other related predictors: GDP per capita (USD), literacy rate (%), Internet users (%), medical schools (per million capita), and number of physicians (per 1000 capita). We used linear regression to create three models:

1. predicting the number of PubMed pageviews per million population based on the percentage of English speakers in a country and the above predictors;
2. predicting the number of MedlinePlus English pageviews per million population based on the percentage of English speakers in a country and the above predictors; and
3. predicting the number of MedlinePlus *en Español* pageviews per million population based on the percentage of Spanish speakers in a country and the above predictors.

Linearity was assessed for all the models, and a log transformation for the number of pageviews was required for all the models. Note that a direct result of the log transformation of the outcome is that we are modeling the percent increase or decrease in the number of pageviews, rather than the number of pageviews *per se*. No further steps were required to satisfy the assumption of linearity, such as converting the continuous predictors to categorical predictors, so all the predictors in the model were continuous. For all three models, we used all the predictors listed above in order to assess how well each of these factors predicts the outcome. We also tested whether there were any significant interactions between the percentage of English speakers (Model 1 and 2) or Spanish speakers (Model 3) and the other predictors in the model, in order to investigate whether these predictors influence the effect of language on pageviews.

We used the R-squared statistic to measure the amount of variation in the number of pageviews accounted for by the models, and the F test to check whether the model significantly predicts the outcome. All other tests and confidence intervals are based on the t-distribution and all p values are two-sided. All the variables were continuous, and any interactions between literacy and the other variables were also included in the model as continuous predictors. These interactions were logarithmic so it was not possible to express linear relationships between the variables, but by specifying one of the variables at various representative points (e.g. 10%, 50%, and 90%) in the results tables, we were able to illustrate how the number of pageviews changed at different points in the logarithmic curve. This is because, when an interaction between two predictors is included in a model, it is not possible to say how an increase in one predictor affects the outcome as a single linear relationship. Therefore, while we couldn't say that the number of PubMed pageviews increased by some percentage for every 1% increase in English speakers overall, for example, we could illustrate that in countries with literacy rates around 90%, the number of pageviews would increase by 11% for every 1% increase in English.

All statistical tests were performed at the 5% level of significance. All the interactions were significant overall but some confidence intervals may not be significant when calculated for specific representative points, because they are based on a smaller sample size (such as the number of values we had around the 10% point, for example). We used the R software [28] to perform all statistical analyses.

4. Results

We found population, GDP per capita, literacy rate, rate of Internet use, medical schools per capita, physicians per capita, PubMed pageviews, and MedlinePlus English pageviews for 97 countries. This excludes 53 countries for which some of these data were unknown, 89 countries for which no reliable counts of English speakers could be found, and four root domains for entities not considered countries by the CIA World Factbook: the European Union, French Guiana, Guadeloupe, and Reunion. In the case of MedlinePlus *en Español* pageviews, only 63 countries were included in the final tabulation due to missing or unknown data. ►Table 1, ►Table 2, ►Table 3 show the data for PubMed, MedlinePlus English, and MedlinePlus *en Español*, respectively, with each table abridged to the ten largest and ten smallest countries by population. ►Table 4 shows the total number of pageviews served by the three online resources during their respective sample periods, as well as the exclusions made due to unrecognized or unknown root domains.

4.1 Results for PubMed

►Table 5 shows the results (coefficients and *p* values) of the linear regression analysis for PubMed use by country. Overall, the factors jointly and significantly ($p<0.001$) predicted approximately 84% of the variation in the number of PubMed pageviews ($R^2 = 0.838$). ►Table 5 also shows the percentage change in the number of PubMed pageviews associated with a change in one of the predictors, and the corresponding 95% confidence interval (CI). For example, a $1,000 USD increase in the GDP per capita of a country corresponds to a 4.1% increase in the number of PubMed pageviews, on average. Furthermore, we found that there is a significant interaction between the percentage of English speakers and literacy rate in a country, which means that the effect of English speakers on the number of PubMed pageviews depends on the literacy rate and, conversely, the effect of literacy on the number of PubMed pageviews depends on the percentage of English speakers. In order to describe the percentage change in the number of PubMed pageviews associated with an increase in the percentage of English speakers, we also needed to specify the literacy rate. Therefore, in order to help with interpretation of the model, we calculated the effects of literacy rates of 10%, 50% and 90%, to represent low, middle and high literacy countries, respectively. For example, a 1% increase in the number of English speakers corresponds to an 11% increase in PubMed pageviews for a country with 90% literacy, whereas it corresponds to a 36.8% decrease in the number of PubMed pageviews if literacy is only 10%. Hence, our model predicts that an increase in the number of English speakers will be associated with an increase the number of PubMed pageviews, but only if the literacy rate is high. All factors except Internet users and medical schools were found to be significant predictors of the number of PubMed pageviews.

4.2 Results for MedlinePlus English

►Table 6 shows the results (coefficients and *p* values) of the linear regression analysis for MedlinePlus English use by country. Overall, the factors jointly and significantly ($p<0.001$) accounted for approximately 69% of the variation in the number of MedlinePlus English pageviews ($R^2 = 0.687$). Again, we found a significant interaction between the percentage of English speakers and literacy rate, which means that the effect of English speakers on MedlinePlus English pageviews depends on the literacy rate of that country. For example, a 1% increase in the number of English speakers corresponds to a 16.1% increase in MedlinePlus English pageviews if the country has 90% literacy, whereas it corresponds to a 26.9% decrease in the number of MedlinePlus English pageviews if literacy rate is 10%. Hence, an increase in the number of English speakers increases the number of MedlinePlus English pageviews only if the literacy rate is high. All factors except GDP and physicians were found to be significant predictors of MedlinePlus English pageviews.

4.3 Results for MedlinePlus en Español

►Table 7 shows the results (coefficients and *p* values) of the linear regression analysis for MedlinePlus *en Español* use by country. Overall, the factors jointly and significantly ($p<0.001$) predicted approximately 88% of the variation in the number of MedlinePlus *en Español* pageviews ($R^2 = 0.883$). We found that the percentage of Spanish speakers had significant interactions with the GDP per capita, and percentage of Internet users. In this case, when we estimate the percentage change in MedlinePlus *en Español* pageviews given a percentage increase in Spanish speakers, we must also specify the GDP and the level of Internet use since the amount of change in the number of pageviews associated with an increase in the number of Spanish speakers will be different depending on the levels of GDP and Internet use. We used $4,600 USD (the 10^{th} percentile of GDP per capita) and $39,500 USD (the 90^{th} percentile) to represent low and high GDP countries, respectively, and, 10%

and 90% Internet users to represent low and high Internet usage, respectively. For example, a 1% increase in the number of Spanish speakers corresponds to a 90.4% increase in the number of MedlinePlus *en Español* pageviews for a country with $4,600 USD GDP per capita and 10% Internet use ($p < .001$), whereas it corresponds to a 132% increase in MedlinePlus *en Español* pageviews if the country has high GDP and high Internet use ($p<0.001$). Hence, an increase in the number of Spanish speakers increases MedlinePlus *en Español* pageviews regardless of the level of GDP or Internet use, but the amount of increase is even greater if GDP per capita and the percentage of Internet users are high. Only the percentage of Spanish speakers, GDP per capita, percent Internet users and their corresponding interactions were found to be significant predictors of MedlinePlus *en Español* pageviews.

5. Discussion

Although the effects of language barriers in health communication have been demonstrated [5-11], their effects on health information access in countries whose primary language is less pervasive has not been reported. In this study, we reviewed pageview data and country statistics to determine if proficiency in a language was associated with higher access to NLM's databases. Using the rate of dominant-language speakers to represent language proficiency by country, we found that use of PubMed and MedlinePlus English was predicted by fluency in English, and that use of MedlinePlus *en Español* was significantly predicted by fluency in Spanish. For English, the relationship between the number of pageviews of PubMed and MedlinePlus English was further affected by the level of literacy in the country, such that high literacy was needed for an increase in English fluency to effect an increase in the number of pageviews. Conversely, for countries with low literacy rates, an increase in the percentage of English speakers did not lead to an increase in pageviews. For Spanish speaking countries, an increase in the percentage of Spanish speakers was associated with an increase in the number of pageviews in MedlinePlus *en Español*. This relationship was further moderated by GDP per capita and Internet use. This data seems to validate the study on the effects of language fluency on the use of MedlinePlus *en Español* where the introduction of MedlinePlus in Spanish significantly increased its usage in Spanish-speaking countries [13].

Other socioeconomic factors, such as GDP, literacy, Internet usage, and number of medical schools and physicians, had varying effects on the pageviews for the websites. Aside from English fluency, for PubMed, GDP, literacy and number of physicians were significantly associated with increased pageviews. GDP is a universally recognized measure of a country's economy and living standards [29]. It is likely that high GDP countries have greater resources to acquire health information technologies needed to support their medical research infrastructure and health information retrieval. GDP has also been shown to be a good predictor of Internet access [30]. If we project this trend, we would expect that Internet user rate would have the same significant effect as GDP on PubMed pageviews. However, this was not observed in our study. A possible explanation for this is that majority of Internet users are not in the scientific community which make them less likely to access PubMed; or maybe, because Internet use and GDP are correlated, having GDP in the model left no "independent" predictive power for Internet use. Conversely, since PubMed is primarily accessed by health professionals and scientists, the number of physicians correlated significantly more with PubMed pageviews by country.

In contrast to PubMed, MedlinePlus English pageviews were significantly influenced by all factors except GDP and number of physicians. The dissociation of GDP with the use of health-consumer-oriented resources like MedlinePlus may support the observation of Lorence and Park that consumer health information-seeking patterns are comparable across populations regardless of financial status, provided technologies are accessible [31]. Moreover, Pew Internet found that 80% of Internet users access health information websites, making it the third most used service after email and web search [32]. For MedlinePlus *en Español*, only Internet users, GDP and percentage of Spanish speakers were significant predictors of the number of pageviews.

An observation we made while processing the pageview data for MedlinePlus *en Español* is that Uruguay ranked first among Spanish-speaking countries with over 55 million pageviews or 24% of the total pageviews in the past four years. This came as a surprise since Uruguay's population is only

3.3 million, while Mexico, which has a population of over 113 million people, ranked only second with 42.65 million pageviews. Part of this difference may be predicted by Uruguay's higher Internet usage (38%) compared to Mexico's (21%); however, there may be unmeasured cultural or other differences reflected in this finding. In general our study shares all those limitations associated with ecological studies. including potential confounding by unmeasured variables, unmeasured dependencies between the variables we did measure, and the aggregation or averaging of individual variations in variables measured at the country level. We attempted to minimize these limitations by analyzing several socioeconomic determinants without presuppositions about which were dominant or interdependent; for example, treating literacy rate, Internet usage, and medical schools per capita as independent variables. While our analysis found no statistical collinearity, we could not rule out other sorts of complex dependencies between the variables we measured.

The study may also be limited by incomplete or unverified sources of several key data that are mostly derived from websites. Foremost among these was aggregation of English-speaker and Spanish-speaker counts by country from several sources that did not use identical definitions and were not all verifiable. Even with the use of these questionable data sources, incomplete data prompted us to exclude more than half the world's countries for the English analysis and consider only a third for the Spanish analysis. An additional aggregation error may have been introduced by using population estimates for the year 2010 but older data for other variables, when no current data were available. In the worst cases, some surveys of physicians by country were up to 10 years old at the time of this study. While website statistics for PubMed were recent, they required three caveats: they were based on Internet domain name registration, which might not always be an accurate indication of geographic location; they were measured in number of pages viewed, so the number of queries they represent was probably less; and they were collected during a single week in July 2010 which may not have been generally representative of year-round use. In contrast, MedlinePlus statistics were more comprehensive with four years' worth of data. Finally, we only included pageviews accessed from computers. In recent years, pageviews from mobile devices were tabulated in a separate database.

Future research in this area could include enhanced data collection to correct some of the problems experienced in this study; for example, use of a more uniform and verifiable database on language speakers by country if one could be ascertained, a larger sample of countries by restoring other missing data points, a longer time period for the PubMed data sample, and a count of the number of PubMed and MedlinePlus queries rather than pageviews. The effect of BabelMeSH could also be investigated by analyzing PubMed use rates in countries with widespread use of the BabelMeSH-supported languages compared to other countries.

6. Conclusions

Among the factors investigated in this study, language fluency, expressed as the percentage of English speakers and Spanish speakers by country, was a consistently significant predictor of the use of online health resources, such as PubMed and MedlinePlus. This finding may support a need for multi-language tools that may increase access to health information resources from non-English-speaking countries.

Clinical Relevance Statement

Clinicians and healthcare systems that serve disadvantaged or linguistic-minority populations may need to communicate with patients differently or more extensively to compensate for decreased utilization of online evidence-based medical resources. This disparity in health literacy might also be reduced by multi-language tools that increase access to online health resources.

Conflict of Interest

The authors declare that they have no conflicts of interest in the research.

Human Subjects Protections

Human and animal subjects were not included in the project.

Acknowledgements
The authors would like to thank Dennis Benson, PhD, of the National Center for Biotechnology Information, for several useful discussions. This research was supported by the Intramural Research Program of the National Institutes of Health, National Library of Medicine, and Lister Hill National Center for Biomedical Communications.

Table 1 Independent variables and PubMed pageviews (Ten largest and ten smallest countries by population, of 97 countries tabulated)

Country	Population	English speakers	GDP per capita US$	Literacy rate	Internet usage	Medical schools per million	Physicians per 1000	PubMed page-views (13–17 Sep 2010)	PubMed page-views per million
China	1,330,141,295	1%	6,600	92%	22%	0.09	1.4	1,706,748	1,283
India	1,173,108,018	23%	3,100	61%	7%	0.16	0.55	968,914	826
United States	310,232,863	96%	46,000	99%	74%	0.48	2.56	13,666,314	44,052
Pakistan	184,404,791	11%	2,500	50%	10%	0.29	0.1	47,145	256
Bangladesh	156,118,464	2%	1,500	48%	0%	0.19	0.27	9,626	62
Nigeria	152,217,341	53%	2,300	68%	7%	0.14	0.36	24,730	162
Russia	139,390,205	5%	15,100	99%	32%	0.45	4.41	172,698	1239
Japan	126,804,433	12%	32,700	99%	72%	0.63	2.13	2,510,063	19,795
Mexico	112,468,855	5%	13,200	86%	21%	0.52	2.7	282,724	2,514
Philippines	99,900,177	55%	3,300	93%	6%	0.46	0.9	137,926	1,381
Grenada	107,818	91%	10,300	96%	22%	9.27	0.74	1,547	14,348
Micronesia	107,154	58%	2,200	89%	15%	0	0.56	99	924
St. Vincent	104,217	95%	10,200	96%	63%	19.19	0.85	178	1,708
Seychelles	88,340	38%	20,800	92%	36%	22.64	1.37	203	2,298
Antigua and Barbuda	86,754	80%	17,800	86%	75%	23.05	0.14	311	3,585
Andorra	84,525	22%	44,900	100%	70%	0	2.95	130	1,538
Dominica	72,813	94%	10,200	94%	38%	27.47	0.52	586	8,048
St. Kitts	49,898	78%	14,700	98%	32%	100.2	0.92	753	15,091
Cook Is.	11,488	20%	9,100	95%	44%	0	1.74	8	696
Niue	1,354	6%	5,800	95%	74%	0	2.95	0	0

Table 2 Independent variables and MedlinePlus English pageviews (Ten largest and ten smallest countries by population, of 97 countries tabulated)

Country	Population	English speakers	GDP per capita US$	Literacy rate	Internet usage	Medical schools per million	Physicians per 1000	Medline Plus English page-views (01 Jan. 2007 – 27 Dec. 2010)	Medline Plus English page-views per 10,000
China	1,330,141,295	1%	6,600	92%	22%	0.09	1.4	1,393,354	10
India	117,310,8018	23%	3,100	61%	7%	0.16	0.55	6,206,718	53
United States	310,232,863	96%	46,000	99%	74%	0.48	2.56	313,085,661	10,092
Pakistan	184,404,791	11%	2,500	50%	10%	0.29	0.1	599,945	33
Bangladesh	156,118,464	2%	1,500	48%	0%	0.19	0.27	137,517	9
Nigeria	152,217,341	53%	2,300	68%	7%	0.14	0.36	175,088	12
Russia	139,390,205	5%	15,100	99%	32%	0.45	4.41	374,610	27
Japan	126,804,433	12%	32,700	99%	72%	0.63	2.13	1,347,614	106
Mexico	112,468,855	5%	13,200	86%	21%	0.52	2.7	6,033,607	536
Philippines	99,900,177	55%	3,300	93%	6%	0.46	0.9	4,238,785	424
Grenada	107,818	91%	10,300	96%	22%	9.27	0.74	66,259	6,145
Micronesia	107,154	58%	2,200	89%	15%	0	0.56	3,613	337
St. Vincent	104,217	95%	10,200	96%	63%	19.19	0.85	15,218	1,460
Seychelles	88,340	38%	20,800	92%	36%	22.64	1.37	6,629	750
Antigua and Barbuda	86,754	80%	17,800	86%	75%	23.05	0.14	61,036	7,036
Andorra	84,525	22%	44,900	100%	70%	0	2.95	7,276	861
Dominica	72,813	94%	10,200	94%	38%	27.47	0.52	91,359	12,547
St. Kitts	49,898	78%	14,700	98%	32%	100.2	0.92	27,821	5,576
Cook Is.	11,488	20%	9,100	95%	44%	0	1.74	1,397	1,216
Niue	1,354	6%	5,800	95%	74%	0	2.95	165	1,219

Table 3 Independent variables and MedlinePlus en Español pageviews (Ten largest and ten smallest countries by population, of 63 countries tabulated)

Country	Population	Spanish speakers	GDP per capita US$	Literacy rate	Internet usage	Medical schools per million	Physicians per 1000	Medline Plus en Español pageviews (01 Jan. 2007 – 30 Nov. 2010)	Medline Plus en Español page-views per 10,000
China	1,330,141,295	0	6,600	92%	22%	0.09	1.4	38,128	0
United States	31,0232,863	15.8	46,000	99%	74%	0.48	2.56	18,520,176	597
Russia	139,390,205	0.01	15,100	99%	32%	0.45	4.41	15,503	1
Japan	126,804,433	0.1	32,700	99%	72%	0.63	2.13	97,221	8
Mexico	112,468,855	98.5	13,200	86%	21%	0.52	2.7	42,936,495	3,818
Philippines	99,900,177	3.1	3,300	93%	6%	0.46	0.9	12,152	1
Germany	82,282,988	3.2	34,100	99%	75%	0.51	3.5	462,927	56
Turkey	77,804,122	0.01	11,400	87%	31%	0.57	1.49	17,214	2
France	64,768,389	9.6	32,600	99%	66%	0.69	3.52	415,605	64
United Kingdom	62,348,447	6.4	34,800	99%	78%	0.61	2.02	409,756	66
Grenada	107,818	91%	10,300	96%	22%	9.27	0.74	66,259	6,145
Micronesia	107,154	58%	2,200	89%	15%	0	0.56	3,613	337
St. Vincent	104,217	95%	10,200	96%	63%	19.19	0.85	15,218	1,460
Seychelles	88,340	38%	20,800	92%	36%	22.64	1.37	6,629	750
Antigua and Barbuda	86,754	80%	17,800	86%	75%	23.05	0.14	61,036	7,036
Andorra	84,525	68.7	44,900	100%	70%	0	2.95	55,361	6,550
Dominica	72,813	94%	10,200	94%	38%	27.47	0.52	91,359	12,547
St. Kitts	49,898	78%	14,700	98%	32%	100.2	0.92	27,821	5,576
Cook Is.	11,488	20%	9,100	95%	44%	0	1.74	1,397	1,216
Niue	1,354	6%	5,800	95%	74%	0	2.95	165	1,219

Table 4 Total Pageviews by Health Information Resource

Source	Period	Total page-views	Number of root domains	Excluded
PubMed	13–17 Sept. 2010	39,958,804	229	21,529 (0.1%)
MedlinePlus English	01 Jan. 2007 – 27 Dec. 2010	468,382,743	240	91,503 (0.02%)
MedlinePlus en Español	01 Jan. 2007 – 30 Nov. 2010	230,975,440	50	381,276 (0.16%)

Table 5 Results from linear regression model predicting number of pageviews on PubMed; *p value significant at the 0.05 level; †p value for the overall contribution of this factor to the model including interaction effect

Predictor	Coefficients	p value	Increase in predictor value	Value of other interaction predictors	% change in number of PubMed views per increase in predictor	95% CI for prediction of % change in number of PubMed pageviews	p value
Intercept	4.66959	<0.001*					
English-speakers (%)	-0.05297	0.0001*†	10%	Literacy 10% Literacy 50% Literacy 90%	-36.8 -16.3 11.0	[-51.9, –17.1] [-27.1, –3.7] [3.8, 18.7]	0.0012* 0.0132* 0.0026*
GDP per capita (USD)	0.00004	0.0001*	$1,000		4.1	[2.1, 6.2]	0.0001*
Literacy (%)	0.01069	<0.001*†	10%	English 10% English 50% English 90%	19.4 58.3 109.8	[0.7, 41.6] [33.8, 87.3] [61.9, 172.0]	0.0417* <0.001* <0.001*
Internet users (%)	0.00864	0.1871	10%		9.0	[-4.2, 24.1]	0.1871
Medical schools (per million)	0.01286	0.1159	1 school per million		1.3	[-0.3, 2.9]	0.1159
Physicians (per 1000)	0.38755	0.0002*	1 physician per 1000		47.3	[21.1, 79.3]	0.0002*
Interaction between English-speakers and literacy	0.00070	0.0001*	-		-	-	

Table 6 Results from the linear regression model predicting number of pageviews on MedlinePlus English; * p value significant at the 0.05 level; † p value for the overall contribution of the factor to the model including interaction effect

Predictor	Coefficients	p value	Increase in predictor value	Value of other interaction predictors	% change in number of MedlinePlus English pageviews per increase in predictor	95% CI for prediction of % change in number of MedlinePlus English pageviews	p value
Intercept	5.85561	<0.001*	-		-		
English-speakers (%)	-0.03713	0.0026*†	10%	Literacy 10% Literacy 50% Literacy 90%	-26.9 -7.9 16.1	[-50.6, 8.1] [-24.6, 12.5] [5.9, 27.3]	0.1146 0.4165 0.0018*
GDP per capita (USD)	0.00002	0.2338	$1,000		1.6	[-1.0, 4.4]	0.2338
Literacy (%)	0.03241	<0.001*†	10%	English 10% English 50% English 90%	46.5 84.6 132.7	[13.7, 88.8] [46.0, 133.5] [62.9, 232.5]	0.0036* <0.001* <0.001*
Internet users (%)	0.02005	0.0265*	10%		22.2	[2.4, 45.8]	0.0265*
Medical schools (per million)	0.03280	0.0039*	1 school per million		3.3	[1.1, 5.6]	0.0039*
Physicians (per 1000)	-0.15153	0.2658	1 physician per 1000		-14.1	[-34.3, 12.4]	0.2658
Interaction between English-speakers and literacy	0.00058	0.0251*	-		-	-	

Table 7 Results of the linear regression model to predict the number of pageviews of MedlinePlus en Español; * p value significant at the 0.05 level; † p value for the overall contribution of this factor to the model including interaction effect; ‡ 10th percentile of GDP; § 90th percentile of GDP

Predictor	Coefficients	p value	Increase in Predictor value	Value of other interaction predictors	% change in number of MedlinePlus en Español pageviews per increase in predictor	95% CI for prediction of % change in number of MedlinePlus en Español pageviews	p value
Intercept	3.08366	0.0079*	-		-		
Spanish -speakers (%)	0.06225	<0.001* †	10%	GDP$4,600‡ Internet 10% Internet 90% GDP$39,500§ Internet 10 Internet 90%	 90.4 264.0 67.81 32.0	 [72.2, 110.5] [159.9, 409.9] [42.7, 97.2] [82.5, 194.8]	 <0.001* <0.001* <0.001* <0.001*
GDP per capita (USD)	0.00005	0.0096*†	$1,000	Spanish 10% Spanish 90%	4.3 -5.9	[7.6, 10.9] [-10.7,-0.9]	0.0092* 0.0226*
Literacy (%)	0.02243	0.1286	10%		25.1	[-6.5, 67.5]	.1286
Internet users (%)	0.01529	0.0003*†	10%	Spanish 10% Spanish 90%	26.3 141.5	[2.6, 55.6] [70.3, 242.6]	0.0287* <0.001*
Medical schools (per million)	0.09860	0.1535	1 school per million		25.1	[9.2, 43.5]	0.0017*
Physicians (per 1000)	-0.01780	0.8896	1 physician per 1000		-1.8	[-24.0, 26.9]	0.8896
Interaction between Spanish-speakers and GDP	-0.000001	0.0006*	-		-	-	
Interaction between Spanish-speakers and Internet use	0.00081	0.0008*					

References

1. Top Ten Internet Languages - World Internet Statistics [Internet]. [Cited 2011 Aug 30]. Available from http://www.internetworldstats.com/stats7.htm
2. Wolk RM. The effects of English language dominance of the internet and the digital divide. 2004 International Symposium on Technology and Society, ISTAS '04 Globalizing Technological Education [Internet]. 2004 [cited 2011 Aug 31]. p. 174–80. Available from http://www.scopus.com/inward/record.url?eid=2-s2.0-7744225741&partnerID=40&md5=b5d0f889eed28eee0831c9b12f46b9a2
3. Ono H, Zavodny M. Immigrants, English Ability and the Digital Divide. IZA Discussion Paper Series [Internet]. 2007 Oct [cited 2011 Aug 30];(3124). Available from http://ftp.iza.org/dp3124.pdf
4. Tang H-WV, Yin M-S, Sheu R-S. The relationship between English language adoption and global digital inequality: A cross-country analysis of ICT readiness and use. Int J Distance Educ Tech 2011; 9(2): 13–22.
5. Fiscella K, Franks P, Doescher MP, Saver BG. Disparities in health care by race, ethnicity, and language among the insured: findings from a national sample. Med Care 2002; 40(1): 52–59.
6. David RA, Rhee M. The impact of language as a barrier to effective health care in an underserved urban Hispanic community. Mt Sinai J Med 1998; 65(5-6): 393–397.
7. Apter AJ, Reisine ST, Affleck G, Barrows E, ZuWallack RL. Adherence with twice-daily dosing of inhaled steroids. Socioeconomic and health-belief differences. Am J Respir Crit Care Med 1998; 157(6 Pt 1): 1810–1817.
8. Carrasquillo O, Orav EJ, Brennan TA, Burstin HR. Impact of language barriers on patient satisfaction in an emergency department. J Gen Intern Med 1999; 14(2): 82–87.
9. Gulati S, Watt L, Shaw N, Sung L, Poureslami IM, Klaassen R, et al. Communication and language challenges experienced by Chinese and South Asian immigrant parents of children with cancer in Canada: Implications for health services delivery. Pediatr Blood Cancer [Internet]. 2011 Mar 2 [cited 2011 Aug 30]. Available from http://www.ncbi.nlm.nih.gov/pubmed/21370432
10. Bischoff A, Tonnerre C, Loutan L, Stalder H. Language difficulties in an outpatient clinic in Switzerland. Soz Praventivmed 1999; 44(6): 283–287.
11. Baker DW, Hayes R, Fortier JP. Interpreter use and satisfaction with interpersonal aspects of care for Spanish-speaking patients. Med Care 1998; 36(10): 1461–1470.
12. Binder L, Nelson B, Smith D, Glass B, Haynes J, Wainscott M. Development, implementation, and evaluation of a medical Spanish curriculum for an emergency medicine residency program. J Emerg Med 1988; 6(5): 439–441.
13. Leon SA, Fontelo P. MedlinePlus en Español and Spanish-speakers. AMIA Annu Symp Proc 2007; 1028.
14. Edejer TT. Disseminating health information in developing countries: the role of the internet. BMJ 2000; 321(7264): 797–800.
15. Smith H, Bukirwa H, Mukasa O, Snell P, Adeh-Nsoh S, Mbuyita S, et al. Access to electronic health knowledge in five countries in Africa: a descriptive study. B MC Health Serv Res 2007; 7: 72.
16. Doughty K, Rothman L, Johnston L, Le K, Wu J, Howard A. Low-Income Countries' Orthopaedic Information Needs: Challenges and Opportunities. Clin Orthop Relat Res 2010; 486(10): 2598-2603.
17. Fact Sheet The National Library of Medicine [Internet]. 1975 [cited 2011 Sep 1]. Available from http://www.nlm.nih.gov/pubs/factsheets/nlm.html
18. Fact Sheet MedlinePlus® [Internet]. 2001 [cited 2011 Sep 8]. Available from http://www.nlm.nih.gov/pubs/factsheets/medlineplus.html
19. MedlinePlus Turns 10 Years Old [Internet]. 2008 [cited 2011 Sep 8]. Available from http://www.nlm.nih.gov/news/mplus_turns_ten.html
20. National Library of Medicine. <BabelMeSH>< MEDLINE/PubMed Multilanguage Search [Internet]. [Cited 2011 Aug 31]. Available from http://babelmesh.nlm.nih.gov/
21. CIA - The World Factbook [Internet]. [Cited 2010 Nov 13]. Available from https://www.cia.gov/library/publications/the-world-factbook/index.html
22. Copenhagen University. Global Directories of Education Institutions for Health Professions [Internet]. [Cited 2010 Nov 13]. Available from http://avicenna.ku.dk/
23. World Health Organization. Global Health Atlas [Internet]. [Cited 2010 Nov 13]. Available from http://apps.who.int/globalatlas/dataQuery/
24. Wikipedia. List of countries by English-speaking population [Internet]. Wikipedia - The Free Encyclopedia. 2010 [cited 2010 Nov 13]. Available from http://en.wikipedia.org/wiki/List_of_countries_by_English-speaking_population
25. Ethnologue report for language code: English [Internet]. Ethnologue Languages of the World. 2010 [cited 2010 Nov 13]. Available from http://www.ethnologue.com/show_language.asp?code=eng

26. Wikipedia. Spanish language - Wikipedia, the free encyclopedia [Internet]. [Cited 2011 Aug 31]. Available from http://en.wikipedia.org/wiki/Spanish_language
27. Spanish Speaking Population Statistics - Worldwide Population [Internet]. [Cited 2011 Aug 31]. Available from http://www.spanishseo.org/resources/worldwide-spanish-speaking-population
28. R Development Core Team. R: A language and environment for statistical computing [Internet]. R Foundation for Statistical Computing; 2010 [cited 2011 Sep 1]. Available from http://www.r-project.org/
29. Goossens Y, Makipaa A, Schepelmann P, van de Sand I, Kuhndtand M, Herrndorf M. Alternative progress indicators to Gross Domestic Product (GDP) as a means towards sustainable development [Internet]. European Parliament Policy Department: Economic and Scientific Policy; 2007 Oct [cited 2011 Aug 31]. Available from http://www.beyond-gdp.eu/download/bgdp-bp-goossens.pdf
30. Gil-de-Zuniga H. Reshaping Digital Inequality in the European Union: How Psychological Barriers Affect Internet Adoption Rates. Webology [Internet]. 2006 Dec [cited 2011 Sep 1];3(4). Available from http://www.webology.org/2006/v3n4/a32.html
31. Lorence D, Park H. Web-based consumer health information: public access, digital division, and remainders. MedGenMed 2006; 8(2): 4.
32. Fox S. 80% of internet users look for health information online [Internet]. Pew Research Center's Internet & American Life Project; 2011 Feb [cited 2011 Aug 31]. Available from http://pewinternet.org/Reports/2011/HealthTopics.aspx