SC|M
Studies in Communication | Media

FULL PAPER

Online health information seeking in Europe: Do digital divides persist?

Online-Suche nach Gesundheitsinformation in Europa: Fortbestehende Digital Divides?

Marko Bachl

Marko Bachl, Universität Hohenheim, Institut für Kommunikationswissenschaft, Fruwirthstraße 46, 70599 Stuttgart; Kontakt: marko.bachl(at)uni-hohenheim.de

FULL PAPER

Online health information seeking in Europe: Do digital divides persist?

Online-Suche nach Gesundheitsinformation in Europa: Fortbestehende Digital Divides?

Marko Bachl

Abstract: First- and second-level digital divides are of great concern in health communication research and practice. It is feared that the increased importance of online health information might also increase informational inequalities between those who can and those who for different reasons cannot benefit from online health communication. Using a large-scale representative survey from the 28 member states of the European Union (*N* = 26 566), we investigated macro-level divides between countries and micro-level divides based on sociodemographic and health-related individual characteristics. The Internet is established as a channel for health communication for substantial parts of the European population. Yet meaningful differences were still persistent even within the highly-developed region of the EU. Internet access divides were most pronounced in comparison, both at the macro and the micro level. Substantial differences were also found between the users of health-related Internet services regarding a wider range of online practices and online health literacy.

Keywords: Health communication, digital divides, online health information seeking, health-related Internet use, European comparison, representative general population survey.

Zusammenfassung: Digitale Spaltungen, also Unterschiede im Zugang zu und der Nutzung von digitalen Informations- und Kommunikationstechniken, haben für die Gesundheitskommunikation große wissenschaftliche und praktische Relevanz. Es besteht die Sorge, dass die Vorteile der fortschreitenden Etablierung von Online-Angeboten in der Gesundheitskommunikation nicht allen Gesellschaftsschichten gleichermaßen zugutekommen, sondern sich im Gegenteil bestehende Ungleichheiten im Zugang zu gesundheitsrelevanten Informationen verstärken. In diesem Artikel untersuchen wir mit einer repräsentativen Befragung in den 28 Mitgliedsstaaten der Europäischen Union (*N* = 26 566), wie sich der Zugang zu und die Nutzung von gesundheitsbezogenen Internet-Angeboten zwischen den Ländern (Makro-Ebene) und zwischen Individuen (Mikro-Ebene) unterscheiden. Wir zeigen, dass das Internet für viele Europäer ein relevanter Kanal der Gesundheitskommunikation ist. Allerdings bestehen innerhalb der hochentwickelten EU weiterhin bedeutende Ungleichheiten. Am deutlichsten ausgeprägt sind Unterschiede im Zugang zum Internet. Doch auch unter den Internetnutzern finden sich substantielle Unterschiede hinsichtlich verschiedener gesundheitsbezogener Online-Aktivitäten und der hierfür relevanten Fähigkeiten.

Schlagwörter: Gesundheitskommunikation, Digitale Spaltung, gesundheitsbezogene Internetnutzung, europäischer Vergleich, repräsentative Befragung.

1. Online health information seeking in Europe: Do digital divides persist?

The Internet is today an important source for information on health and illnesses. The consequences for health communication have been controversially discussed (Gibbons et al., 2011; Kreps & Neuhauser, 2010). On the one hand, the technological developments sparked great hopes. Citizens should be more informed, because all kinds of health-related information are just one mouse-click away. Health content can be tailored to meet the specific needs of diverse audiences. The medium allows interactive exchanges with healthcare professionals and between patients, relatives, and all interested citizens – to name only some of the desirable qualities and outcomes.

On the other hand, the rise of the Internet as an important channel for health communication brought along substantial challenges. Most importantly, *digital divides* between those who can and those who for different reasons cannot benefit from online health communication continue to be of great concern. It is feared that the increased importance of online health information might also increase informational inequalities, and that these inequalities might correspond with and thus exaggerate existing health disparities. The warning of Viswanath and Kreuter (2007) remains relevant: "[W]e contend that without careful and systematic research and policy, e-health may work primarily to the advantage of individuals and communities with greater resources and healthcare systems that serve them" (p. 131). In this article, we answer their call by updating the empirical evidence on differences in online health information seeking from a European perspective.

2. Analytic framework and literature overview

Digital divides are investigated at different levels of analysis and regarding multiple dimensions (Ball-Rokeach & Jung, 2008; Norris, 2001). Analyses at the individual level look at differences between individuals based on, for example, gender, socioeconomic status, ethnicity, or cognitive abilities. Aggregate level analyses compare communities, countries, or continents. Various dimensions of digital divides are discussed in the literature. One prominent distinction is between first-level access divides and second-level skills and uses divides (Hargittai & Hsieh, 2013; van Deursen & van Dijk, 2014). First-level divides are concerned with the most basic question of who has at all access to information and communication technologies (ICTs). Without access to ICT, it is of course impossible to benefit directly from its developments. At the macro level, members of groups with lower access rates may be disadvantaged, because indirect access (for example asking your neighbors if they could get some information from the Internet) is harder to come by. Moreover, less ICT services may be supplied to the users in such communities, because the overall demand may be too low to make services rewarding. Second-level divides describe differences in how those with access use ICTs. Special attention is also paid to the skills which are necessary to benefit from ICT use (Hargittai, 2002; van Deursen & van Dijk, 2011). This strand of research became more important with the recent developments of high ICT access rates in many modern societies. It is assumed that soon most people will have access to ICT, but other factors beyond access, such as cognitive abilities and cultural preferences, will influence who will be able to benefit most from using ICT. Digital divides research is generally interested in the use of all kinds of ICTs. In this article, however, we focus on the Internet as the most important ICT in health communication for the general population.

Both first- and second-level digital divides are of relevance for health communication. Internet access is the necessary precondition for any health-related Internet use (access divides). Health-related online activities are quite specialized uses of the Internet that may vary within and between populations (uses divides). They require relevant skills, which may also be unequally distributed (skills divides). Similarly, both macro-level and micro-level analyses are important. Macro-level results may guide (supra-) national policy initiatives like – in our case – the "eHealth Action Plan 2012-2020" of the European Commission (2012). Evidence on individual-level differences informs about which characteristics are starting points for interventions to close the respective divides. It also provides guidance on who will be reached by which online communication efforts and who should better be addressed offline.

2.1 Macro-level digital divides in Europe

Relevant in the scope of this article are macro-level differences between the member countries of the European Union (EU). The EU is overall an economically and technologically well-developed region, but it still consists of a diverse set of countries. First-level access divides between the member states of the EU have been closing (Kyriakidou, Michalakelis, & Sphicopoulos, 2011). Yet differences are still persistent, mainly along the lines of the European integration process and economic wealth. Newer member countries in Eastern Europe and economically less prospering countries in the south showed lower levels of Internet adoption. In contrast, Internet access in some northern and western European states has reached almost complete saturation (Cruz-Jesus, Oliveira, & Bacao, 2012). Similar differences were also identified for Internet adoption by the public and business sectors (Seri, Bianchi, & Matteucci, 2014) and in more specific Internet skills and usage patterns (van Dijk, 2009).

There is very limited evidence on macro-level differences in health-related Internet use in Europe. One available study, a comparative survey from seven European countries (Denmark, Germany, Greece, Latvia, Norway, Poland and Portugal) in 2005 and 2007, found country-specific differences in several online health activities. Health-related Internet use was more common in the northern and western countries compared to those in the south and east. Differences were also identified regarding more specific health-related online activities (Andreassen et al., 2007; Kummervold et al., 2008; Santana et al., 2011).

2.2 Micro-level digital divides in health-related Internet use

Many studies have addressed individual-level differences in health-related Internet use. First-level access divides were found along the lines of sociodemographic and economic factors, such as age, sex/gender, education, race/ethnicity, urbanity of place of residence, and income. Although these gaps seemed to have reduced over time, they were still persistent in more recent studies (Fox & Duggan, 2013; Prestin, Vieux, & Chou, 2015). The results on the relationship of health-related variables and Internet access were mixed. While earlier studies reported disparities, such that those with worse health or chronic illnesses were less likely to use the Internet (Kontos, Emmons, Puleo, & Viswanath, 2010; Renahy, Parizot, & Chauvin, 2008), a more recent study did not find effects of health- and health-care-related factors (Prestin et al., 2015).

Among Internet users, women and individuals with higher education and more income were more likely to use the Internet for health-related purposes. Being directly or indirectly affected by health issues also increased the likelihood of health-related Internet use. Similar differences were identified in studies on more specific health-related online activities (see Cline & Haynes, 2001; Higgins, Sixsmith, Barry, & Domegan, 2011; Zschorlich et al., 2015, for reviews). The importance of the skills that are necessary for fully benefiting from online health activities has also been recognized. The "ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem" (Norman & Skinner, 2006b) should be taken into account additionally to the more general Internet skills and health literacy (Bickmore & Paasche-Orlow, 2012).

2.3 Study objectives

There is plenty of research on online health information seeking. Yet, most published studies analyzed the situation in the USA. In a recent systematic review of 74 articles published 2008 to 2013, Zschorlich et al. (2015) identified 40 studies on the USA, but only ten on the UK, eight on Australia, five on Canada, five on other single European countries, and six comparative studies. Empirical evidence from other industrialized countries besides the USA remained sparse. This particularly applies to comparative studies from Europe. The most recently published academic study that we know of reported on data from a 2007 survey that covered seven countries (Kummervold et al., 2008; Santana et al., 2011). But as Internet adoption rates kept growing, usage patterns developed, and new online health content and services were introduced, it is important to keep track of changes in differential access to and use of the Internet as health-related information source and communication channel.

In this article, we aim to update the empirical evidence on health-related Internet use in Europe. The analyses first cover the more general questions of who used the Internet at all for health-related purposes. We then turn to more specific questions regarding domains, information types, and sources of health-related Internet use and online health literacy. We compare all outcomes between countries to uncover macro-level divides. Micro-level divides are identified by probing the effects of individual-level sociodemographic and health-related characteristics on the outcomes.

3. Methods

3.1 Data source

We analyze data from the Flash Eurobarometer 404 (European Commission, 2015). The survey was conducted as part of the "eHealth Action Plan 2012–2020" of the European Commission (2012). The data set, original questionnaires, and further documentation are publicly available at the GESIS data archive, and the replication code for all statistical analyses is published in the supplemental material to this article. TNS Political & Social conducted the field work from September 18th to 20th, 2014. The samples for each country were generated by regionally stratified random digit dialing on both fixed landlines and mobile phones. 26 566 residents of the 28 EU member countries were surveyed with computer-assisted telephone interviews ($n \approx 1000$ per country, except for Cyprus, Luxembourg, and Malta, where $n \approx 500$). The data are representative for the populations of EU residents aged 15 and older in each country. Extensive descriptions of all variables are documented in a report of the European Commission (2014).

3.2 Measures

Internet access and general health-related Internet use. Respondents were asked how frequently they used the Internet for private purposes. Respondents who answered "never" or reported not to have Internet access were classified as non-users and all other respondents as Internet users. General health-related Internet use was measured among Internet users with the question "Within the last 12 months, have you used the Internet to search for health-related information? This could include information on an injury, a disease, illness, nutrition, improving health, etc.". Respondents who answered "No, never" were classified as non-users and all other respondents as users, regardless of usage frequency.

Domain-specific uses of health-related online information. Users of health-related online information were then asked more specifically about the domains of their activities. Response categories were "General information on health-related topics or ways to improve your health", "Information on a specific injury, disease, illness or condition" (hereafter: *information on health problems*), "Information to get a second opinion after having visited your doctor" (hereafter: *second opinions*), and "Specific information on a medical treatment or procedure" (hereafter: *information on treatments*). Respondents could select up to two out of the four response categories. The first domain was followed by a set of diverse topics on the specific contents of the searches (e.g., information on lifestyle choices, information on pregnancy, childbirth and early infancy, information on vaccinations). The other domains were followed by lists which were similar to each other but differed from the first domain (see below). We therefore excluded the first response category, because it was too topically diverse to allow for a comparison with the follow-up questions of other domains. Separate analyses for this domain would have exceeded the scope of the present article. For a comparison of general health-related information seeking and specific information on health problems, see Bachl (2016).

Types of health-related online information and information sources. Respondents were asked further questions for each selected domain. They were read a list on which they indicated what they were looking for more specifically. The lists included several more detailed areas of factual information (e.g., “Information on the symptoms of a specific disease”). We summarized the information items, such that a respondent who indicated at least one of the items was classified as searching for *factual information*. Other types of information were *reports of other patients* (“Testimonials or experiences from other patients”) and *seeking emotional support* (“Emotional support in dealing with a health issue”). Respondents were also asked which sources they used in each domain. Response categories included general sources (e.g., “Internet search engines”) and health-specific sources (e.g., “Websites from official health organizations”). Multiple responses were permitted. The domain-specific answers were summarized so that the final variables indicated the types of information and sources used by a respondent, regardless of the specific domain.

Online health literacy. Online health literacy was assessed from all users of health-related online information with five items similar to the eHEALS scale (Norman & Skinner, 2006a). Example items were “You know where to find reliable health-related information on the Internet” or “You understand the terminology used on the Internet for health-related topics”. Answers were provided on 4-point scales. The items were summarized into a mean index with sufficient internal consistency (Cronbach’s $\alpha = .74$) and a range from 0 (low literacy) to 3 (high literacy).

Sociodemographic and health-related predictors. The following sociodemographic characteristics were included to determine individual-level digital divides: age (measured in years and rescaled to 10 years for ease of interpretation); sex (0 = male, 1 = female); urbanity of place of residence (0 = “Rural area or village”, 1 = “Small or medium-sized town”, 2 = “Large town/city”); education (age at the end of full-time education in years).

Health status was measured by the question “How would you rate your level of health in general?”. Response categories were “Very bad”, “Fairly bad”, “Fairly good”, and “Very good”. They were assigned the values 0 to 3, so that higher values indicated better health. Health knowledge was measured by the question “How would you assess your general knowledge of health-related topics?” with the same response categories. Likewise, the values 0 to 3 were assigned so that higher values indicated better health knowledge.

Surrogate information seeking. Surrogate information seeking is an important control variable, because surrogate seekers differ in their search behavior (Sadasivam et al., 2013). Respondents were asked for whom they searched health-related information in each domain. Response categories included “Yourself” and several other persons. Respondents who did not search information for themselves were classified as surrogate seekers.

3.3 Statistical analysis

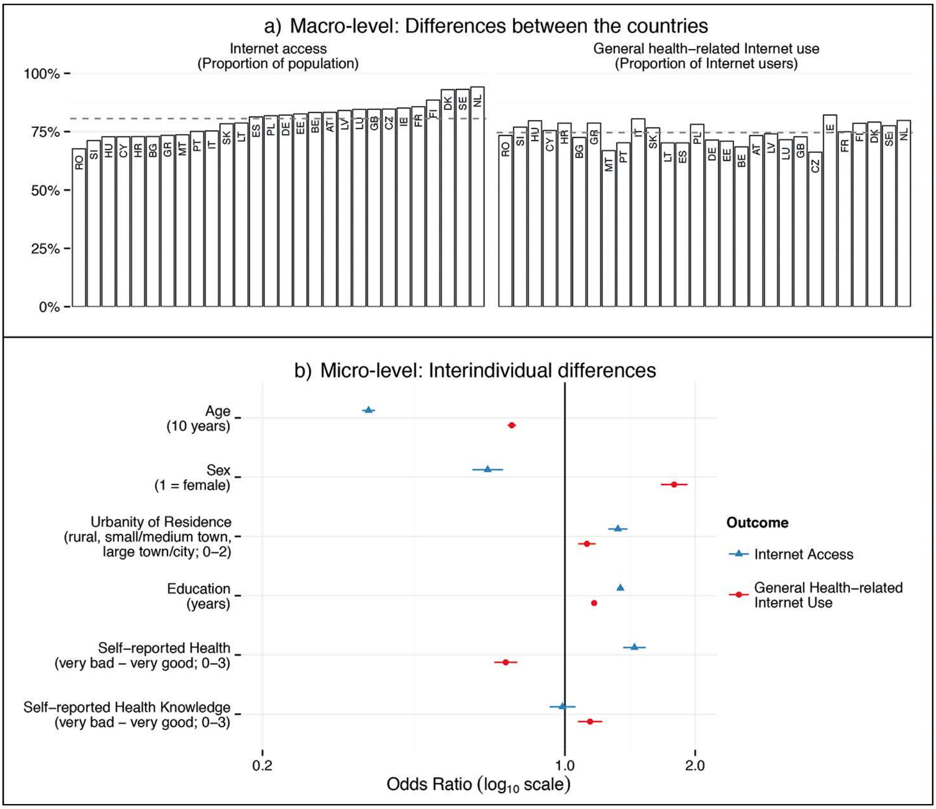
The comparisons between countries were adjusted for sampling bias by using the within-country weights provided with the data set. The weighting procedure corrected for unit non-response regarding sex, age, region, and size of locality. The *R* (R Core Team, 2015) package *survey* (Lumley, 2004) was used for the calculation of the adjusted statistics.

All multivariate analyses were carried out as multilevel models with the countries as Level 2 units. Multilevel models account for country-level effects that go beyond the effects of the predictor variables that were measured at the individual level. We used the *R* package *lme4* (Bates, Mächler, Bolker, & Walker, 2015) for multilevel modeling. The results are presented as coefficient plots with the uncertainty around the point estimates being represented by intervals of two standard errors. *z* and *t* values are reported to compare the relative importance of the predictors within a model. Odds ratios (*OR*) and linear coefficients (*b*) can be compared across the models. *P* values or confidence intervals at a defined α level are not reported because Null Hypothesis Testing was not an aim of the analyses.

4. Results

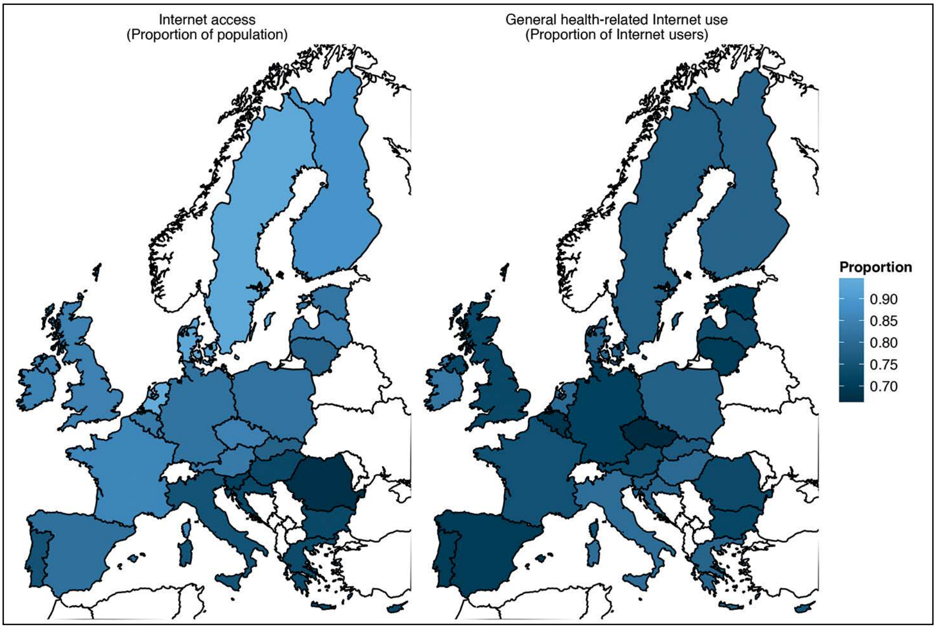
4.1 Internet access and general health-related Internet use

Figure 1: Differences in Internet access and general health-related Internet use



Notes: The analyses are based on *N* = 24 722 respondents and *n* = 18 049 Internet users. Upper panel: Horizontal lines show averages across the 28 countries, where every country has the same weight regardless of population size. Countries are denoted with ISO 3166-1 country codes. Lower panel: Coefficients from multilevel logit models with fixed slopes and intercepts varying between countries. Data points show the fixed effect estimates (odds ratios). Horizontal lines represent two standard errors. Different colors and shapes represent the outcome variables.

Figure 2: Differences in Internet Access and General Health-related Internet Use Between Countries: Geographical Distribution



Notes: Lighter blue indicates higher proportions. Country shapes from *cshapes* (Weidmann & Gleditsch, 2013).

Figure 1a and Figure 2 illustrate the macro-level divides in Internet access and general health-related Internet use. Overall, *Internet adoption rates* were high in the EU with an across-country average (i.e., the mean proportion across the 28 countries, where every country has the same weight regardless of population size) of 81 percent. Yet the proportions of Internet users still varied substantially: The Netherlands, Sweden, and Denmark approached almost total saturation. In contrast, only two thirds of Romania's population and less than three quarters of the populations of five south-eastern countries used the Internet. Figure 2 gives an impression of the geographical distribution. Internet access rates were highest in the northern and north-western countries and lowest in the newer, economically weaker EU member states in the south-east.

Using the Internet to search for health-related information was common among Internet users: Across countries, an average of three quarters of Internet users did so. There was less variation between countries compared to Internet access, the extremes being Ireland and Italy at the top with over 80 percent and the Czech Republic and Malta with about two thirds at the bottom. Remarkably, no geographic pattern (see Figure 2) and no systematic relationship between Internet access rates and proportions of online health information users (see Figure 1a) were found.

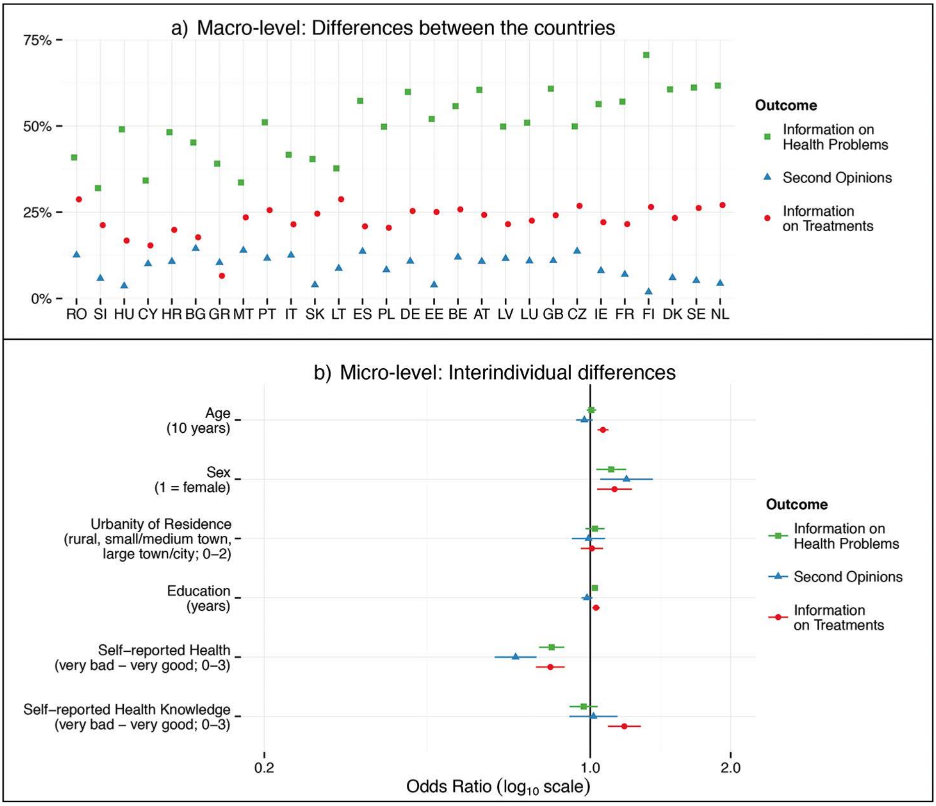
At the micro level, *Internet access divides* were identified for all commonly established sociodemographic characteristics (see Figure 1b). Age was the most important predictor (z = -61.37; OR = 0.35), with every 10 years reducing the odds to use the Internet by one third. Every year of education raised the odds 1.34-fold (z = 37.35;

OR = 1.34). Roughly equally important factors were sex (female; *z* = -10.12; *OR* = 0.66) and urbanity of place of residence (*z* = 11.07; *OR* = 1.33). Additionally, people with better health were more likely to use the Internet (*z* = 12.32; *OR* = 1.45).

Age (*z* = -24.83; *OR* = 0.75) and education (*z* = 21.17; *OR* = 1.17) were also the most important predictors discriminating among Internet users between *users and non-users of online health information*. Being 10 years older decreased the odds of health-related Internet use by three quarters. Every year of education raised the odds 1.17-fold. Women had 1.8-fold odds of searching for health-related information compared to men (*z* = 16.41; *OR* = 1.79). Residents of more urban areas were more likely to use the Internet for health purposes (*z* = 5.07; *OR* = 1.12). Both health-related characteristics predicted the use of online health information. Worse health status (*z* = -10.32; *OR* = 0.73) and better health knowledge (*z* = 4.12; *OR* = 1.14) raised the odds to do so.

4.2 Domain-specific uses of health-related online information

Figure 3: Domain-specific uses of health-related online information



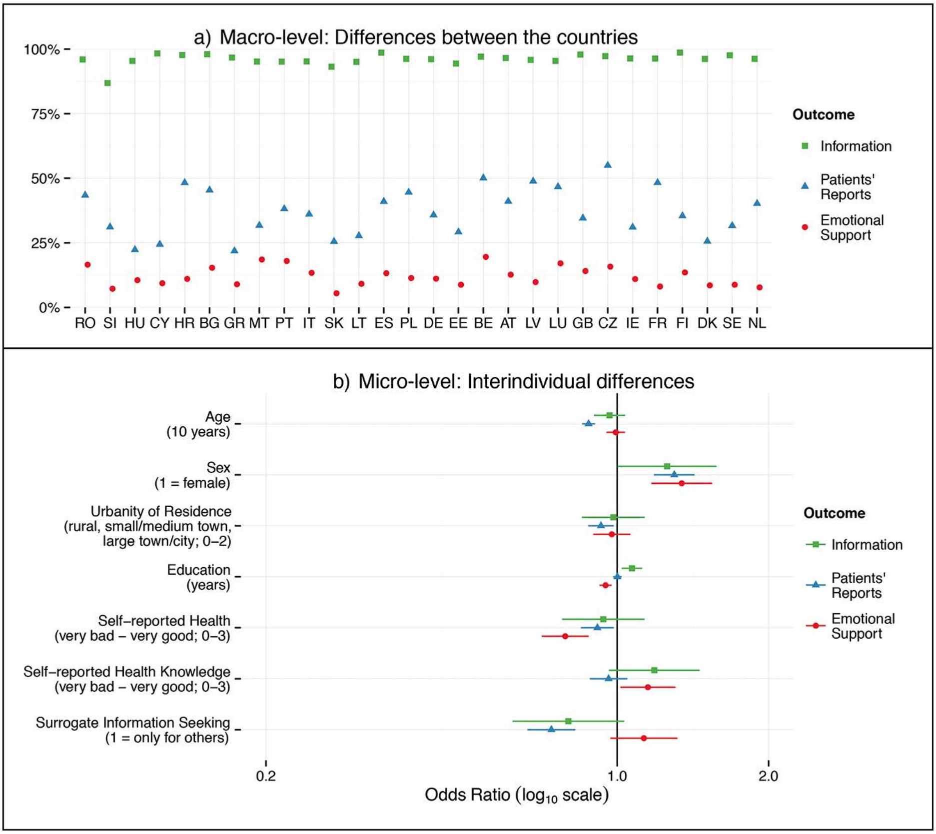
Notes: Different colors and shapes represent the outcome variables. The analyses are based on *n* = 13 134 respondents who used the Internet for health-related purposes. Upper panel: Proportions of all users of general health-related online information. The order of the countries on the x-axis follows the Internet adoptions rates in Figure 1. Sums exceed 100% because of multiple responses. Lower panel: Coefficients from multilevel logit models with fixed slopes and intercepts varying between countries. Data points show the fixed effect estimates (odds ratios). Horizontal lines represent two standard errors.

The domain-specific uses showed similar patterns in all EU member states (see Figure 3a). Searching for *information on health problems* was the most common task in all countries with an across-country average of 50 percent. The shares were somewhat higher in countries that had higher Internet adoption rates. Searching for *information on specific treatments* was the second most important domain in almost all countries (across-country average 23%). Only relatively few users of online health information searched for *second opinions* after they visited their doctor (across-country average 9%).

There were relatively minor individual-level differences in the domain-specific uses compared to the prior analyses (see Figure 3b). Unsurprisingly, health status was the most important predictor in all three models: Individuals with worse health were more likely to be active in all three domains (health problems: $z = -6.14$; $OR = 0.83$; second opinions: $z = -7.12$; $OR = 0.69$; treatments: $z = -5.58$; $OR = 0.82$). Women were more likely than men to search for information in all domains (health problems: $z = 2.81$; $OR = 1.11$; second opinions: $z = 2.75$; $OR = 1.20$; treatments: $z = 2.77$; $OR = 1.13$). Age ($z = 4.61$; $OR = 1.06$), education ($z = 3.01$; $OR = 1.03$), and self-reported health knowledge ($z = 4.12$; $OR = 1.18$) predicted the search for *information on treatments*.

4.3 Types of health-related online information and information sources

Figure 4: Use of types of health-related online information

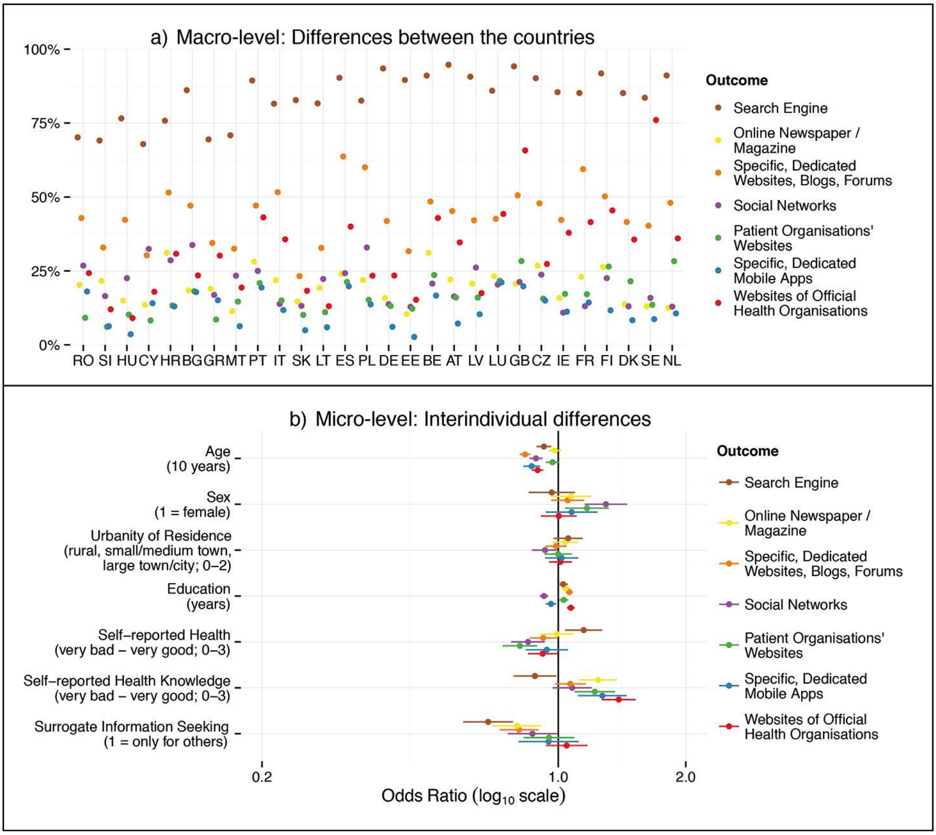


Notes: Different colors and shapes represent the outcome variables. The analyses are based on n = 9 269 respondents who used the Internet for health-related purposes in at least one of the three domains under investigation. Upper panel: Proportions of all users of health-related online information in the three domains information on health problems, second opinions, and information on treatments. The order of the countries on the x-axis follows the Internet adoptions rates in Figure 1. Sums exceed 100% because of multiple responses. Lower panel: Coefficients from multilevel logit models with fixed slopes and intercepts varying between countries. Data points show the fixed effect estimates (odds ratios). Horizontal lines represent two standard errors.

Almost all users of health information in the domains under investigation searched for more specific *factual information* (see Figure 4a). Searching for *reports from other patients* was also quite common. Between one fifth to half of the relevant subpopulations in each country used such reports. We could not, however, identify a meaningful pattern in the variation between the EU member states. Seeking *emotional support* was the least popular activity. The proportions varied between 5 percent in Slovakia and 20 percent in Belgium. Again, no pattern was identified in the between-country variation.

The individual-level differences are depicted in Figure 4b. Searching for *factual information* was so widespread that there were basically no interindividual differences. Only education (z = 2.87; OR = 1.07) and sex (z = 2.02; OR = 1.26) had small effects. Users of *patients' reports* were younger (z = -8.92; OR = 0.88), female (z = 5.61; OR = 1.30), and lived in more rural areas (z = -2.55; OR = 0.93). Individuals with worse health were more likely to seek peer testimonials (z = -2.41; OR = 0.91), surrogate seekers were less likely to do so (z = -5.50; OR = 0.74). Women (z = 4.24; OR = 1.34), individuals with worse health (z = -4.43; OR = 0.79), and more knowledgeable persons (z = 2.22; OR = 1.15) were more likely to seek *emotional support*. Individuals with higher formal education, in contrast, were less likely to do so (z = -3.90; OR = 0.95).

Figure 5: Use of online information sources



Notes: Different colors represent the outcome variables. The analyses are based on n = 9 269 respondents who used the Internet for health-related purposes in at least one of the three domains under investigation. Upper panel: Proportions of all users of health-related online information in the three domains information on health problems, second opinions, and information on treatments. The order of the countries on the x-axis follows the Internet adoptions rates in Figure 1. Sums exceed 100% because of multiple responses. Lower panel: Coefficients from multilevel logit models with fixed slopes and intercepts varying between countries. Data points show the fixed effect estimates (odds ratios). Horizontal lines represent two standard errors.

Next, respondents indicated whether they used several general-purpose and health-specific Internet services. *Search engines* were most popular in all countries with an across-country average of 84 percent and a range between 68 and 95 percent of the relevant subpopulations (see Figure 5a). The other general-purpose services, *websites of newspapers and magazines* and *social networks*, were both on average used by one fifth of the relevant subpopulations with a range from 10 percent to one third.

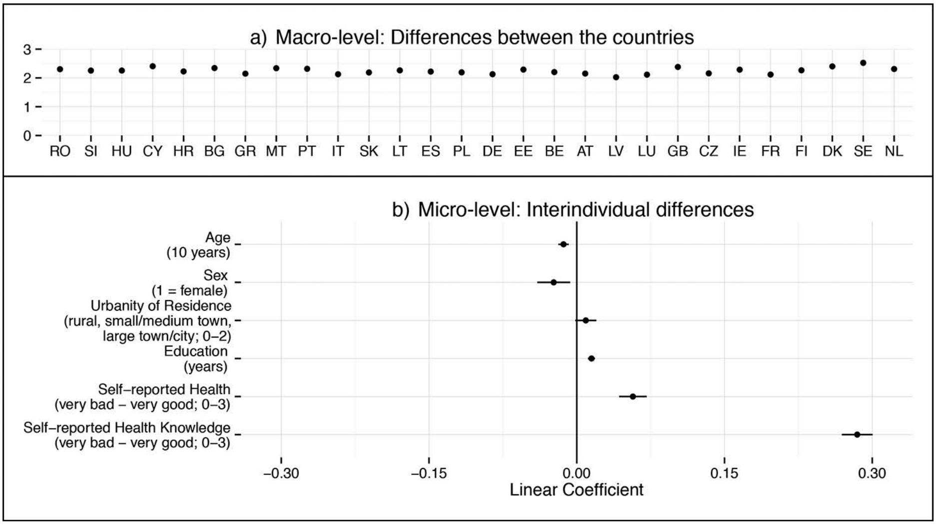
Among the health-specific sources, *dedicated websites, blogs and forums* were most popular with shares between 23 and 64 percent and an across-country average of 44 percent. They were the second most important type of sources in almost all countries, with proportions of the relevant subpopulations above 40 percent in 21 out of the 28 countries. *Websites from official health organizations* showed the greatest variation between countries. On average, 32 percent of the relevant subpopulations used these websites, but the proportions ranged from 76 percent in Sweden and 66 percent in the UK to 9 percent in Hungary, 12 percent in Slovenia, and 13 percent in Lithuania (see below for a discussion). *Websites of patient organizations* were on average used by 16 percent of the relevant subpopulations. Over one quarter of the relevant subpopulations of the UK, Netherlands, and Finland accessed these websites. In contrast, they were hardly visited at all in some south-eastern countries. *Specific applications for mobile devices* were not very commonly used overall with an across-country average of 12 percent .

Figure 5b presents the interindividual differences. Given the almost universal appeal of *search engines*, their users unsurprisingly showed only a weak profile. Older individuals (z = -3.91; OR = 0.92) and surrogate information seekers (z = -5.61; OR = 0.68), who used almost all sources less likely, were less likely to use search engines. Additional indicators for the role of search engines as the least specialized resources were the correlations with health status (z = 2.73; OR = 1.15) and health knowledge (z = -2.18; OR = 0.88). Individuals with better health (and thus presumably with less experience in the domains under investigation) and less health knowledge were somewhat more likely to use search engines. *Online newspapers and magazines* were preferred by individuals who were more educated (z = 3.30; OR = 1.04) and more knowledgeable in the health domain (z = 4.20; OR = 1.24). *Social networks* as channels for health-related communication appealed to younger (z = -6.75; OR = 0.89), female (z = 4.49; OR = 1.30), less educated (z = -6.55; OR = 0.93), and less healthy (z = -3.59; OR = 0.85) individuals who lived in less urban areas (z = -2.00; OR = 0.93).

The users of *dedicated websites, blogs or forums* were younger (z = -12.50; OR = 0.84) and more educated (z = 6.15; OR = 1.06). Individuals with worse health were also more likely to visit these websites (z = -2.20; OR = 0.92). The individual profile of users of *websites from official health institutions* was similar, with younger (z = -7.15; OR = 0.89), more educated (z = 6.42; OR = 1.07), and less healthy (z = -2.12; OR = 0.92) individuals being among their more likely visitors. In addition, individuals who were more knowledgeable in the health domain were more likely to use official online sources (z = 7.10; OR = 1.39). The somewhat similar *websites of patient organizations* were also used more likely by less healthy (z = -4.45; OR = 0.81), more educated (z = 2.36; OR = 1.03), and more knowledgeable (z = 3.59; *OR* = 1.22) individuals. In contrast to the other two more specific sources, age was not an important predictor for visiting patient organizations' websites, but women were more likely than men to visit them (z = 2.64; *OR* = 1.17). *Health-specific mobile applications* were more likely used by younger (z = -6.38; *OR* = 0.87) individuals who were less educated (z = -2.69; *OR* = 0.96), but had more health-related knowledge (z = 3.61; *OR* = 1.27).

4.4 Online health literacy

Figure 6: Online health literacy



Notes: The analyses are based on n = 13 134 respondents who used the Internet for health-related purposes. Upper panel: Mean scores of all users of general health-related online information. The order of the countries on the x-axis follows the Internet adoptions rates in Figure 1. Lower panel: Coefficients from a multilevel linear model with fixed slopes and intercepts varying between countries. Data points show the fixed effect estimates (linear coefficients). Horizontal lines represent two standard errors.

The mean scores for self-reported online health literacy hardly varied between the EU member countries (see Figure 6a). Overall, the users of online health information were quite confident in their abilities. The average score across all countries was 2.3 on the 0-to-3-scale, with all national means between 2 and 2.5.

Individual characteristics had substantial effects on online health literacy (see Figure 6b). Online health literacy was a function of self-reported health knowledge (t = 36.38; b = 0.28). The more individuals were confident in their general health-related knowledge, the more confident they felt in regard to finding, evaluating, and using online health information. Education as an indicator of cognitive abilities and general knowledge was also positively related to online health literacy (t = 8.23; b = 0.01). Users of online health information who were older (t = -5.07; b = -0.01) and female (t = -2.81; b = -0.02) were somewhat less confident in their online health literacy. Somewhat counter-intuitively, individuals with better health reported to have better online health literacy (t = 8.15; b = 0.06; see below for a discussion).

5. Discussion

In most parts, the results confirmed for the situation in the EU in late 2014 what was well-known from earlier research and other world regions, particularly from the USA. The Internet can today be considered an established channel for health communication, but informational inequalities were still persistent. Supra-national policies like the EU's "eHealth Action Plan 2012–2020" and national efforts have to take these disparities into account. Two general patterns emerged in the results. First, micro-level differences between the individuals appeared to be more important than macro-level divides between the countries. Second, first-level divides in Internet use were more pronounced than second-level uses and skills divides, both in the micro- and the macro-level analyses. The frequently cited finding that "the digital divide shifts to differences in usage" (van Deursen & van Dijk, 2014, p. 507) could not be fully supported for the health domain. This result does not imply that the second-level divides are less important, but we also should not yet file away the traditional research on first-level divides.

5.1 Macro-level digital divides

In comparison of all macro-level analyses, Internet access divides turned out to be more severe compared to uses and skills divides. National Internet adoption rates followed the known patterns of EU integration and economic wealth (Cruz-Jesus et al., 2012; Kyriakidou et al., 2011; Seri et al., 2014). A north-west vs. south-east gap was still detectable in late 2014. This results suggests that the promotion of Internet access in the respective countries remains a major task. Smaller differences and no obvious geographical patterns remained when analyzing only the subpopulations of Internet users or users of health-related Internet services.

A remarkable exception is the varying popularity of online resources that are provided by official health institutions. The finding seems to reflect differences in institutional trust and has relevant implications for the ability of national health institutions to reach out to their target groups. In most northern and western countries, they are able reach substantial parts of the online health information seekers. Their websites might therefore play a major role in public health education. In contrast, the official health institutions in most of the newer, post-communist EU member states were not (yet) able to establish their online services as a trustworthy source of information. Similar yet less extreme patterns were found for the popularity of patient organizations' websites. Not only government agencies, but also civil organizations have to catch up in the newer EU member states.

5.2 Micro-level digital divides

In agreement with the literature, the sociodemographic factors age, sex, urbanity of place of residence, and education were identified as predictors across multiple outcomes. Personal health status and health knowledge were especially important for explaining differences in the more specific facets of health-related online activities. Contrary to a recent US-American study (Prestin et al., 2015), worse health also decreased the odds to use the Internet at all. The individual-level divides in Internet access and health-related Internet use are most meaningful when considered in combination. The divides based on age, education, and place of residence formed "double divide[s]" (Renahy et al., 2008): Older and less educated individuals who lived in less urban areas were less likely to use the Internet at all, and if they did, they were less likely to use it for health-related purposes. On the contrary, the access divides based on gender and health status excluded individuals from online health information who otherwise would have quite likely used such information.

When we restricted the analyses to users of health-related online activities, we found smaller yet substantial differences. Older individuals attributed themselves less online health literacy, were less likely to draw on testimonials of other patients, and were less likely to consult several types of sources. The results regarding some newer services (mobile apps, social networks) merely reflect that individuals do not follow every new ICT development with increasing age. However, the results that older individuals were less likely to turn to health-specific websites in general, and websites of official health institutions in particular, are more worrisome. These sources provide relatively high-quality information. Especially official health organizations should put more effort into tailoring their online services to the needs of older target groups.

Women were more active than men in almost every aspect under investigation. They were more likely to use the Internet for health-related purposes, and they were more likely to search for information in all three domains. Among these more involved users, women were again more likely to seek each type of content. Finally, women were more likely to be found among the users of online services that highlight communication among patients and peers: social networks and patient organizations' websites. The stronger inclination of women to engage in health-related topics is well-known (Galdas, Cheater, & Marshall, 2005), and it was also visible in their health-related online activities. In this respect, it seems crucial to address the access divide that still excludes women from ICTs. This is even more important as women are also more likely to serve as surrogate information seekers (Sadasivam et al., 2013).

Naturally, personal health status was an important predictor across multiple outcomes. Individuals with worse health showed higher odds of general and domain-specific health-related Internet use and were more likely to refer to specialized sources. Additionally, worse health was an important predictor for online activities that go beyond mere factual information seeking, for example consulting testimonials by other patients, seeking emotional support, and communicating in social networks. There clearly is demand from individuals with health problems for a variety of health-related online content and services, making the first-level divide based on health status to appear all the more distressing. Somewhat counter-intuitively, worse health status was associated with lower levels of online health literacy. This finding might be explained by the respective informational needs. Individuals with worse health presumably searched for more specific information on their health problems that is harder to find and more complex to evaluate.

The effects of general education and domain-specific knowledge agree with the mechanisms that were described in the digital divides (e.g., Hargittai & Hsieh, 2013; Norris, 2001) and knowledge gap (e.g., Kwak, 1999; Viswanath & Finnegan, 1996) literature. More informationally privileged individuals, who presumably also have better access to and understanding of other information resources, have greater potential to benefit from the relatively new and rapidly expanding ICTs. Lower general education and health-related knowledge, in turn, were decisive risk factors in multiple ways. They were associated with lower odds of Internet access, several health-related Internet activities, and use of health-dedicated Internet sources, as well as with lower online health literacy. Even if there were no access divides, the benefits of ICT use will most likely be unequally distributed based on digital skills and health literacy (Bickmore & Paasche-Orlow, 2012). The promotion of relevant skills particularly among the less educated remains a major task to reduce second-level digital divides.

5.3 Limitations and directions for further research

As with every empirical work, our study has limitations. We built on data from a large-scale representative survey that was conducted as part of an EU policy program. The survey implemented a cross-sectional design, which is a limitation that we share with many micro-level digital divides studies. The lack of longitudinal evidence is most critical for the findings on the role of health-related knowledge, which is most likely both cause and consequence of benefiting from online health information seeking. Our results are also affected by the typical constraints of a secondary analysis of data that was not primarily collected for scientific purposes. The available measures in the survey are far from perfect. First and foremost, the lack of a measure for economic status is obviously a drawback. Other important characteristics, particularly health status and health-related knowledge, were measured with rather superficial one-item self-reports. Similar criticism applies to some measures of health-related online activities and online health literacy. Self-reports of online activities often diverge from real behavior (Scharkow, 2016). The validity of the findings may be enhanced by using passive measures of health-related online activities and test-based assessments of health-related Internet skills (van Deursen & van Dijk, 2011).

All limitations of the data source notwithstanding, we believe that researchers are obligated to use any available data that complement empirical evidence on relevant issues. This especially applies to comparative perspectives on health information seeking in Europe, where more specialized surveys like the US-American HINTS (e.g., Finney Rutten et al., 2012) are not available. By using the Eurobarometer data, we were able to update the partly dated literature with a comparative perspective instead of the predominant single-country analyses. The major advantage of the present study is its breadth, both in its geographical reach and in tapping many different health-related online activities. Users and non-users of health-related Internet services in general as well as the users of more specific activities showed distinct individual profiles.

Research should continue to investigate users and non-users of relevant services more in depth. Detailed research on user motivations, activities, and benefits promises important insights that are needed to improve health-related Internet content and services. Likewise, research on the risk groups – individuals who were identified to less likely benefit from the possibilities of online health communication – might deepen the understanding of non-use reasons and direct toward potential interventions. Finally, the social and health-related consequences for both users and non-users of online health information and communication deserve continuous scholarly attention.

References

Andreassen, H., Bujnowska-Fedak, M., Chronaki, C., Dumitru, R., Pudule, I., Santana, S., … Wynn, R. (2007). European citizens' use of E-health services: A study of seven countries. *BMC Public Health*, *7*(1), 53. doi:10.1186/1471-2458-7-53

Bachl, M. (2016). Gesundheitsbezogene Internet-Nutzung in Europa: Second-level Digital Divides? [Health-related internet use in Europe: Second-level digital divides?]. In A.-L. Camerini, R. Ludolph, & F. B. Rothenfluh (Eds.), *Gesundheitskommunikation im Spannungsfeld zwischen Theorie und Praxis [Health communication between theory and practice]* (pp. 211–223). Baden-Baden: Nomos.

Ball-Rokeach, S. J., & Jung, J.-Y. (2008). Digital divide. In W. Donsbach (Ed.), *The international encyclopedia of communication* (pp. 1304–1308). Malden: Blackwell Publishing.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01

Bickmore, T. W., & Paasche-Orlow, M. K. (2012). The role of information technology in health literacy research. *Journal of Health Communication*, *17*(sup3), 23–29. doi:10.1080/10810730.2012.712626

Cline, R. J. W., & Haynes, K. M. (2001). Consumer health information seeking on the Internet: The state of the art. *Health Education Research*, *16*(6), 671–692. doi:10.1093/her/16.6.671

Cruz-Jesus, F., Oliveira, T., & Bacao, F. (2012). Digital divide across the European Union. *Information & Management*, *49*(6), 278–291. doi:10.1016/j.im.2012.09.003

European Commission. (2012). *eHealth Action Plan 2012-2020 - Innovative healthcare for the 21st century*. Brussels: European Union. Retrieved from http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52012DC0736

European Commission. (2015). Flash Eurobarometer 404: European citizens' digital health literacy. *GESIS Data Archive, Cologne*, *ZA5948 Datafile Version 1.0.0*. doi:10.4232/1.12194

European Commission, Directorate-General for Communications Networks, Content and Technology. (2014). *European citizens' digital health literacy*. Brussels: European Union. doi:10.2759/86596

Finney Rutten, L. J., Davis, T., Beckjord, E. B., Blake, K., Moser, R. P., & Hesse, B. W. (2012). Picking up the pace: Changes in method and frame for the Health Information National Trends Survey (2011–2014). *Journal of Health Communication*, *17*(8), 979–989. doi:10.1080/10810730.2012.700998

Fox, S., & Duggan, M. (2013). *Health Online 2013*. Washington, D.C.: Pew Research Center's Internet & American Life Project. Retrieved from http://www.pewinternet.org/files/old-media//Files/Reports/PIP_HealthOnline.pdf

Galdas, P. M., Cheater, F., & Marshall, P. (2005). Men and health help-seeking behaviour: Literature review. *Journal of Advanced Nursing*, *49*(6), 616–623. doi:10.1111/j.1365-2648.2004.03331.x

Gibbons, M. C., Fleisher, L., Slamon, R. E., Bass, S., Kandadai, V., & Beck, J. R. (2011). Exploring the potential of Web 2.0 to address health disparities. *Journal of Health Communication*, *16*(sup1), 77–89. doi:10.1080/10810730.2011.596916

Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday*, *7*(4). doi:10.5210/fm.v7i4.942

Hargittai, E., & Hsieh, Y.-l. P. (2013). Digital inequality. In W. H. Dutton (Ed.), *Oxford handbook of Internet studies* (pp. 129–150). Oxford: Oxford University Press.

Higgins, O., Sixsmith, J., Barry, M. M., & Domegan, C. (2011). *A literature review on health information-seeking behaviour on the web: A health consumer and health professional perspective*. Stockholm: European Centre for Disease Prevention and Control. doi:10.2900/5788

Kontos, E. Z., Emmons, K. M., Puleo, E., & Viswanath, K. (2010). Communication inequalities and public health implications of adult social networking site use in the United States. *Journal of Health Communication*, *15*(sup3), 216–235. doi:10.1080/10810730.2010.522689

Kreps, G. L., & Neuhauser, L. (2010). New directions in eHealth communication: Opportunities and challenges. *Patient Education and Counseling*, *78*(3), 329–336. doi:10.1016/j.pec.2010.01.013

Kummervold, P. E., Chronaki, C. E., Lausen, B., Prokosch, H. U., Rasmussen, J., Santana, S., … Wangberg, S. C. (2008). eHealth trends in Europe 2005-2007: A population-based survey. *Journal of Medical Internet Research*, *10*(4), e42. doi:10.2196/jmir.1023

Kwak, N. (1999). Revisiting the knowledge gap hypothesis: Education, motivation, and media use. *Communication Research*, *26*(4), 385–413. doi:10.1177/009365099026004002

Kyriakidou, V., Michalakelis, C., & Sphicopoulos, T. (2011). Digital divide gap convergence in Europe. *Technology in Society*, *33*(3–4), 265–270. doi:10.1016/j.techsoc.2011.09.001

Lumley, T. (2004). Analysis of complex survey samples. *Journal of Statistical Software*, *9*(1), 1–19. doi:10.18637/jss.v009.i08

Norman, C. D., & Skinner, H. A. (2006a). eHEALS: The eHealth literacy scale. *Journal of Medical Internet Research*, *8*(4), e27. doi:10.2196/jmir.8.4.e27

Norman, C. D., & Skinner, H. A. (2006b). eHealth literacy: Essential skills for consumer health in a networked world. *Journal of Medical Internet Research*, *8*(2), e9. doi:10.2196/jmir.8.2.e9

Norris, P. (2001). *Digital divide: Civic engagement, information poverty, and the Internet worldwide* (online ed.). Cambridge, UK: Cambridge University Press. Available online: http://www.hks.harvard.edu/fs/pnorris/Books/Digital%20Divide.htm

Prestin, A., Vieux, S. N., & Chou, W.-Y. S. (2015). Is online health activity alive and well or flatlining? Findings from 10 years of the Health Information National Trends Survey. *Journal of Health Communication*, *20*(7), 790–798. doi:10.1080/10810730.2015.1018590

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available online: https://www.R-project.org/

Renahy, E., Parizot, I., & Chauvin, P. (2008). Health information seeking on the Internet: A double divide? Results from a representative survey in the Paris metropolitan area, France, 2005-2006. *BMC Public Health*, *8*(1), 69. doi:10.1186/1471-2458-8-69

Sadasivam, R. S., Kinney, R. L., Lemon, S. C., Shimada, S. L., Allison, J. J., & Houston, T. K. (2013). Internet health information seeking is a team sport: Analysis of the Pew Internet Survey. *International Journal of Medical Informatics*, *82*(3), 193–200. doi:10.1016/j.ijmedinf.2012.09.008

Santana, S., Lausen, B., Bujnowska-Fedak, M., Chronaki, C., Prokosch, H.-U., & Wynn, R. (2011). Informed citizen and empowered citizen in health: Results from an [sic!] European survey. *BMC Family Practice*, *12*(1), 20. doi:10.1186/1471-2296-12-20

Scharkow, M. (2016). The accuracy of self-reported Internet use – A validation study using client log data. *Communication Methods and Measures, 10*(1), 13–27. doi:10.1080/19312458.2015.1118446

Seri, P., Bianchi, A., & Matteucci, N. (2014). Diffusion and usage of public e-services in Europe: An assessment of country level indicators and drivers. *Telecommunications Policy*, *38*(5–6), 496–513. doi:10.1016/j.telpol.2014.03.004

van Deursen, A., & van Dijk, J. (2011). Internet skills and the digital divide. *New Media & Society*, *13*(6), 893–911. doi:10.1177/1461444810386774

van Deursen, A., & van Dijk, J. (2014). The digital divide shifts to differences in usage. *New Media & Society*, *16*(3), 507–526. doi:10.1177/1461444813487959

van Dijk, J. (2009). One Europe, digitally divided. In A. Chadwick & P. N. Howard (Eds.), *Routledge handbook of Internet politics* (pp. 288–304). New York: Routledge.

Viswanath, K., & Finnegan, J. R. (1996). The knowledge gap hypothesis: Twenty-five years later. In B. R. Burleson (Ed.), *Communication Yearbook* (Vol. 19, pp. 187–227). Thousand Oaks: Sage.

Viswanath, K., & Kreuter, M. W. (2007). Health disparities, communication inequalities, and eHealth. *American Journal of Preventive Medicine*, *32*(5, Supplement), S131–S133. doi:10.1016/j.amepre.2007.02.012

Weidmann, N. B., & Gleditsch, K. S. (2013). *Cshapes: CShapes dataset and utilities*. Available online: http://CRAN.R-project.org/package=cshapes

Zschorlich, B., Gechter, D., Janßen, I. M., Swinehart, T., Wiegard, B., & Koch, K. (2015). Gesundheitsinformationen im Internet: Wer sucht was, wann und wie? [Health information on the Internet: Who is searching for what, when and how?]. *Zeitschrift Für Evidenz, Fortbildung Und Qualität Im Gesundheitswesen*, *109*(2), 144–152. doi:10.1016/j.zefq.2015.03.003

Appendix: Country-level proportions for Figures 1–5

Table A1: Country-level proportions for Figures 1 & 2

Country	Internet access	General health-related Internet use
Austria (AT)	0.83	0.73
Belgium (BE)	0.83	0.68
Bulgaria (BG)	0.73	0.72
Cyprus (CY)	0.73	0.75
Czech Republic (CZ)	0.85	0.66
Germany (DE)	0.82	0.71
Denmark (DK)	0.93	0.79
Estonia (EE)	0.83	0.71
Spain (ES)	0.81	0.70
Finland (FI)	0.89	0.79
France (FR)	0.86	0.75
Great Britain (GB)	0.85	0.73
Greece (GR)	0.73	0.79
Croatia (HR)	0.73	0.79
Hungary (HU)	0.73	0.80
Ireland (IE)	0.85	0.82
Italy (IT)	0.75	0.81
Lithuania (LT)	0.79	0.70
Luxembourg (LU)	0.85	0.72
Latvia (LV)	0.84	0.74
Malta (MT)	0.74	0.67
Netherlands (NL)	0.94	0.80
Poland (PL)	0.82	0.78
Portugal (PT)	0.75	0.70
Romania (RO)	0.68	0.73
Sweden (SE)	0.93	0.78
Slovenia (SI)	0.71	0.77
Slovakia (SK)	0.78	0.77

Notes: Bases: Internet access: whole populations; General health-related Internet use: Internet users.

Table A2: Country-level proportions for Figure 3

Country	Information on health problems	Second opinions	Information on treatments
Austria (AT)	0.60	0.11	0.24
Belgium (BE)	0.56	0.12	0.26
Bulgaria (BG)	0.45	0.14	0.18
Cyprus (CY)	0.34	0.10	0.15
Czech Republic (CZ)	0.50	0.14	0.27
Germany (DE)	0.60	0.11	0.25
Denmark (DK)	0.61	0.06	0.23
Estonia (EE)	0.52	0.04	0.25
Spain (ES)	0.57	0.14	0.21
Finland (FI)	0.71	0.02	0.27
France (FR)	0.57	0.07	0.22
Great Britain (GB)	0.61	0.11	0.24
Greece (GR)	0.39	0.10	0.07
Croatia (HR)	0.48	0.11	0.20
Hungary (HU)	0.49	0.04	0.17
Ireland (IE)	0.56	0.08	0.22
Italy (IT)	0.42	0.13	0.21
Lithuania (LT)	0.38	0.09	0.29
Luxembourg (LU)	0.51	0.11	0.23
Latvia (LV)	0.50	0.12	0.22
Malta (MT)	0.34	0.14	0.23
Netherlands (NL)	0.62	0.04	0.27
Poland (PL)	0.50	0.08	0.20
Portugal (PT)	0.51	0.12	0.26
Romania (RO)	0.41	0.13	0.29
Sweden (SE)	0.61	0.05	0.26
Slovenia (SI)	0.32	0.06	0.21
Slovakia (SK)	0.40	0.04	0.25

Notes: Base: Internet users who used the Internet for health-related purposes.

Table A3: Country-level proportions for Figure 4

Country	Information	Patients' reports	Emotional support
Austria (AT)	0.97	0.41	0.13
Belgium (BE)	0.97	0.50	0.20
Bulgaria (BG)	0.98	0.45	0.15
Cyprus (CY)	0.98	0.24	0.09
Czech Republic (CZ)	0.97	0.55	0.16
Germany (DE)	0.96	0.36	0.11
Denmark (DK)	0.96	0.26	0.09
Estonia (EE)	0.94	0.29	0.09
Spain (ES)	0.99	0.41	0.13
Finland (FI)	0.99	0.35	0.14
France (FR)	0.96	0.48	0.08
Great Britain (GB)	0.98	0.35	0.14
Greece (GR)	0.97	0.22	0.09
Croatia (HR)	0.98	0.48	0.11
Hungary (HU)	0.95	0.22	0.11
Ireland (IE)	0.96	0.31	0.11
Italy (IT)	0.95	0.36	0.13
Lithuania (LT)	0.95	0.28	0.09
Luxembourg (LU)	0.95	0.47	0.17
Latvia (LV)	0.96	0.49	0.10
Malta (MT)	0.95	0.32	0.19
Netherlands (NL)	0.96	0.40	0.08
Poland (PL)	0.96	0.45	0.11
Portugal (PT)	0.95	0.38	0.18
Romania (RO)	0.96	0.43	0.17
Sweden (SE)	0.98	0.32	0.09
Slovenia (SI)	0.87	0.31	0.07
Slovakia (SK)	0.93	0.26	0.05

Notes: Base: Internet users who used the Internet for health-related purposes in at least one of the three domains under investigation.

Table A4: Country-level proportions for Figure 5

Country	Search engine	Online newspaper/ Magazine	Specific, dedicated websites, blogs, forums	Social Networks
Austria (AT)	0.95	0.22	0.45	0.16
Belgium (BE)	0.91	0.31	0.48	0.21
Bulgaria (BG)	0.86	0.18	0.47	0.34
Cyprus (CY)	0.68	0.14	0.30	0.32
Czech Republic (CZ)	0.90	0.27	0.48	0.24
Germany (DE)	0.93	0.16	0.42	0.14
Denmark (DK)	0.85	0.14	0.42	0.13
Estonia (EE)	0.90	0.10	0.32	0.13
Spain (ES)	0.90	0.24	0.64	0.24
Finland (FI)	0.92	0.26	0.50	0.23
France (FR)	0.85	0.23	0.59	0.13
Great Britain (GB)	0.94	0.21	0.51	0.21
Greece (GR)	0.69	0.19	0.34	0.17
Croatia (HR)	0.76	0.31	0.51	0.29
Hungary (HU)	0.77	0.15	0.42	0.23
Ireland (IE)	0.85	0.16	0.42	0.11
Italy (IT)	0.82	0.22	0.52	0.14
Lithuania (LT)	0.82	0.19	0.33	0.22
Luxembourg (LU)	0.86	0.23	0.43	0.20
Latvia (LV)	0.91	0.21	0.42	0.26
Malta (MT)	0.71	0.11	0.33	0.23
Netherlands (NL)	0.91	0.12	0.48	0.13
Poland (PL)	0.83	0.22	0.60	0.33
Portugal (PT)	0.89	0.28	0.47	0.25
Romania (RO)	0.70	0.20	0.43	0.27
Sweden (SE)	0.84	0.13	0.40	0.16
Slovenia (SI)	0.69	0.22	0.33	0.16
Slovakia (SK)	0.83	0.15	0.23	0.13

Table A4 (continued): Country-level proportions for Figure 5

Country	Patient organizations' websites	Specific, dedicated mobile apps	Websites of official health organizations
Austria (AT)	0.16	0.07	0.35
Belgium (BE)	0.24	0.17	0.43
Bulgaria (BG)	0.18	0.18	0.23
Cyprus (CY)	0.08	0.14	0.18
Czech Republic (CZ)	0.16	0.15	0.27
Germany (DE)	0.13	0.06	0.23
Denmark (DK)	0.21	0.08	0.36
Estonia (EE)	0.12	0.03	0.15
Spain (ES)	0.21	0.20	0.40
Finland (FI)	0.26	0.12	0.45
France (FR)	0.17	0.14	0.42
Great Britain (GB)	0.28	0.20	0.66
Greece (GR)	0.09	0.15	0.30
Croatia (HR)	0.13	0.13	0.31
Hungary (HU)	0.10	0.04	0.09
Ireland (IE)	0.17	0.11	0.38
Italy (IT)	0.15	0.12	0.36
Lithuania (LT)	0.11	0.06	0.13
Luxembourg (LU)	0.22	0.21	0.44
Latvia (LV)	0.16	0.10	0.17
Malta (MT)	0.15	0.06	0.19
Netherlands (NL)	0.28	0.11	0.36
Poland (PL)	0.15	0.14	0.23
Portugal (PT)	0.21	0.19	0.43
Romania (RO)	0.09	0.18	0.24
Sweden (SE)	0.14	0.09	0.76
Slovenia (SI)	0.06	0.06	0.12
Slovakia (SK)	0.10	0.05	0.18

Notes: Base: Internet users who used the Internet for health-related purposes in at least one of the three domains under investigation.

Replication Code for: Online Health Information Seeking In Europe

Marko Bachl

19 July 2016

This code replicates all analyses presented in the article *Online Health Information Seeking In Europe*. The data set from the Flash Eurobarometer 404 (http://dx.doi.org/10.4232/1.12194) is publicly availeable via GESIS.

Packages and Software Versions

```
library(knitr)
library(haven)
library(stringi)
library(survey)
library(srvyr)
library(ggplot2); theme_set(theme_minimal())
library(scales)
library(tidyr)
library(cshapes)
library(ggmap)
library(lme4)
library(broom)
library(RColorBrewer)
library(dplyr)

sessionInfo()
```

```
## R version 3.2.3 (2015-12-10)
## Platform: x86_64-apple-darwin13.4.0 (64-bit)
## Running under: OS X 10.11.4 (El Capitan)
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] grid      stats     graphics  grDevices utils     datasets  methods
## [8] base
##
## other attached packages:
##  [1] dplyr_0.4.3        RColorBrewer_1.1-2 broom_0.4.0
##  [4] lme4_1.1-11        Matrix_1.2-4       ggmap_2.6.1
##  [7] cshapes_0.5-1      plyr_1.8.3         maptools_0.8-39
## [10] sp_1.2-2           tidyr_0.4.1        scales_0.4.0
## [13] ggplot2_2.1.0      srvyr_0.1.0        survey_3.30-3
## [16] stringi_1.0-1      haven_0.2.0        knitr_1.12.3
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.3         nloptr_1.0.4        formatR_1.3
##  [4] tools_3.2.3         digest_0.6.9        nlme_3.1-125
##  [7] evaluate_0.8.3      gtable_0.2.0        lattice_0.20-33
## [10] png_0.1-7           psych_1.5.8         DBI_0.3.1
## [13] mapproj_1.2-4       yaml_2.1.13         parallel_3.2.3
## [16] proto_0.3-10        stringr_1.0.0       RgoogleMaps_1.2.0.7
## [19] maps_3.1.0          R6_2.1.2            jpeg_0.1-8
## [22] foreign_0.8-66      rmarkdown_0.9.5     RJSONIO_1.3-0
## [25] minqa_1.2.4         reshape2_1.4.1      magrittr_1.5
## [28] MASS_7.3-45         splines_3.2.3       htmltools_0.3
## [31] mnormt_1.5-3        assertthat_0.1      colorspace_1.2-6
## [34] geosphere_1.5-1     munsell_0.4.3       rjson_0.2.15
```

Custom Functions

```
# Remove stuff from foreign::read.spss() or haven::read_sav() (makes dplyr::filter()
 happy)
remove_label = function(x) {
  attr(x, "value.labels") <- NULL
  attr(x, "label") <- NULL
  attr(x, "labels") <- NULL
  class(x) <- NULL
  x
}
```

Data sets

Simple data

```
d = read_sav("ZA5948_v1-0-0.sav") # Data set as downloaded from GESIS has to be in th
e same directory as the *.rmd file
d[] = lapply(d, remove_label)
d = d %>%
  mutate(land = stri_trim(isocntry),
         q1 = ifelse(q1 == 8, NA, ifelse(q1 == 7, 6, q1)) * -1 + 6,
         q2 = ifelse(q2 > 6, NA, q2) * -1 + 6)
```

Weighted data

```
d_w = d %>%
  mutate(dv_inet_acc = ifelse(q1 == 0, 0, 1),
         dv_health_use = ifelse(q2 > 0, 1, 0),
         dv_noreason = q26.1,
         dv_noaccess = q26.2,
         dv_noability = as.numeric(apply(select(., starts_with("q26")[3:8]), 1, funct
ion(x) any(x == 1))),
         dv_literacy = rowMeans(mutate_each(select(., starts_with("q21")[c(1,3:6)]),
funs(ifelse(. > 4, NA, .))), na.rm=T) * -1 + 4,
         dv_spec_ill = q3.2,
         dv_spec_2op = q3.3,
         dv_spec_trt = q3.4,
         dv_uses_info = as.numeric(apply(select(., one_of("q8.1", "q8.2", "q8.3", "q
8.4", "q8.6", "q12.1", "q12.2", "q12.3", "q12.4", "q12.6", "q12.7", "q16.1", "q16.2",
 "q16.3", "q16.4", "q16.6", "q16.7")), 1, function(x) mean(x, na.rm=T) > 0)),
         dv_uses_experience = as.numeric(apply(select(., one_of("q8.5", "q12.5", "q1
6.5")), 1, function(x) mean(x, na.rm=T) > 0)),
         dv_uses_support = as.numeric(apply(select(., one_of("q8.7", "q12.8",
"q16.8")), 1, function(x) mean(x, na.rm=T) > 0)),
         dv_srcrs_search = as.numeric(apply(select(., one_of("q10.1", "q14.1", "q18.
1")), 1, function(x) mean(x, na.rm=T) > 0)),
         dv_srcrs_journ = as.numeric(apply(select(., one_of("q10.2", "q14.2",
"q18.2")), 1, function(x) mean(x, na.rm=T) > 0)),
         dv_srcrs_special = as.numeric(apply(select(., one_of("q10.3", "q14.3", "q18.
3")), 1, function(x) mean(x, na.rm=T) > 0)),
         dv_srcrs_social = as.numeric(apply(select(., one_of("q10.4", "q14.4", "q18.
4")), 1, function(x) mean(x, na.rm=T) > 0)),
         dv_srcrs_patient = as.numeric(apply(select(., one_of("q10.5", "q14.5", "q18.
5")), 1, function(x) mean(x, na.rm=T) > 0)),
         dv_srcrs_app = as.numeric(apply(select(., one_of("q10.6", "q14.6",
"q18.6")), 1, function(x) mean(x, na.rm=T) > 0)),
         dv_srcrs_official = as.numeric(apply(select(., one_of("q10.7", "q14.7", "q1
8.7")), 1, function(x) mean(x, na.rm=T) > 0)),
         iv_age = d1 / 10,
         iv_female = d2 -1,
         iv_health = ifelse(q33 > 4, NA, q33) * -1 + 4,
         iv_heaknow = ifelse(q36 > 4, NA, q36) * -1 + 4,
         iv_urban = ifelse(d13 > 3, NA, d13 - 1),
         iv_edu = ifelse(d4 == 0, NA, ifelse(d4 == 97, 14, ifelse(d4 == 98, d1, ifels
e(d4 == 99, NA, d4)))),
         iv_edu = ifelse(iv_edu < 15, 14, ifelse(iv_edu > 21, 22, iv_edu))
  ) %>%
  mutate(iv_chk = apply(select(., starts_with("iv")), 1, function(x) any(is.na(x))))
%>%
  filter(!iv_chk) %>%
  group_by(land) %>%
  select(starts_with("dv"), w1, uniqid)

d_w_svy = svydesign(ids = ~uniqid, weights = ~w1, data = d_w)
```

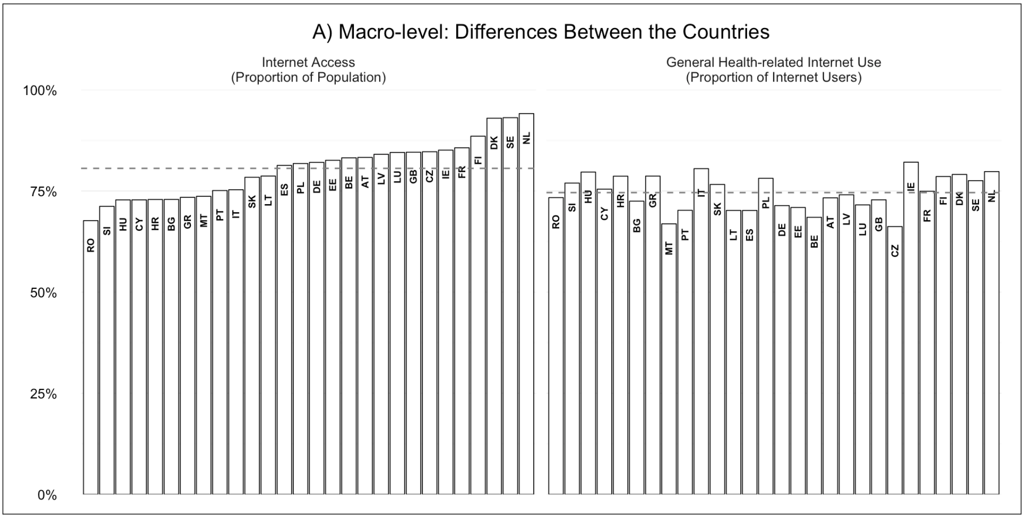
Figure 1: Differences in Internet Access and General Health-related Internet Use

Panel A

```
fig1_d1 = cbind(svyby(~dv_inet_acc, by = ~land, design = d_w_svy, FUN = svymean, na.r
m = T)[,1:2],
                dv_health_use = svyby(~dv_health_use, by = ~land, design = d_w_svy, F
UN = svymean, na.rm = T)[, 2]) %>% tbl_df() %>% mutate(land = reorder(land, dv_inet_a
cc)) %>% gather(dv, P, -land) %>% mutate(dv = factor(dv, levels = unique(dv), labels
= c("Internet Access\n(Proportion of Population)", "General Health-related Internet U
se\n(Proportion of Internet Users)")))

fig1_d2 = fig1_d1 %>% group_by(dv) %>% summarise_each(funs(mean), -land)

fig1_d1 %>% ggplot(aes(land, P, label=land)) + facet_wrap("dv", ncol=2) + scale_y_con
tinuous(expand = c(0,0), breaks = c(0,.25,.50,.75,1), labels=percent) + labs(x=NULL,
y=NULL) + scale_x_discrete(labels=NULL, breaks=NULL) + theme(panel.margin=unit(.5, "l
ines"), plot.margin=unit(rep(.5, 4), "cm"), plot.background = element_rect(color = "b
lack")) + geom_bar(stat = "identity", fill="white", color="black", size=.3) + geom_te
xt(aes(land, P-.06), size=2.4, angle = 90, fontface="bold") +
geom_hline(data=fig1_d2, aes(yintercept=P), linetype=2, color="grey50") + coord_carte
sian(ylim = 0:1) + ggtitle("A) Macro-level: Differences Between the Countries")
```



Panel B

```
d_inet_use = d %>%
  mutate(
    dv_inet_acc = ifelse(q1 == 0, 0, 1),
    iv_age = d1 / 10,
    iv_female = d2 -1,
    iv_health = ifelse(q33 > 4, NA, q33) * -1 + 4,
    iv_heaknow = ifelse(q36 > 4, NA, q36) * -1 + 4,
    iv_urban = ifelse(d13 > 3, NA, d13 - 1),
    iv_edu = ifelse(d4 == 0, NA, ifelse(d4 == 97, 14, ifelse(d4 == 98, d1, ifelse(d4
== 99, NA, d4)))),
    iv_edu = ifelse(iv_edu < 15, 14, ifelse(iv_edu > 21, 22, iv_edu))) %>%
  select(starts_with("dv_"), starts_with("iv_"), land) %>% na.omit

m_inet_acc = glmer(dv_inet_acc ~ iv_age + iv_female + iv_urban + iv_edu + iv_health +
 iv_heaknow + (1 | land), family = binomial, d_inet_use, nAGQ = 0)

d_hea_use = d %>%
  mutate(
    dv_health_use = ifelse(q2 > 0, 1, 0),
    iv_age = d1 / 10,
    iv_female = d2 -1,
    iv_health = ifelse(q33 > 4, NA, q33) * -1 + 4,
    iv_heaknow = ifelse(q36 > 4, NA, q36) * -1 + 4,
    iv_urban = ifelse(d13 > 3, NA, d13 - 1),
    iv_edu = ifelse(d4 == 0, NA, ifelse(d4 == 97, 14, ifelse(d4 == 98, d1, ifelse(d4
== 99, NA, d4)))),
    iv_edu = ifelse(iv_edu < 15, 14, ifelse(iv_edu > 21, 22, iv_edu))) %>%
  select(starts_with("dv_"), starts_with("iv_"), land) %>% na.omit

m_hea_use = glmer(dv_health_use ~ iv_age + iv_female + iv_urban + iv_edu + iv_health
+ iv_heaknow + (1 | land), family = binomial, d_hea_use, nAGQ = 0)

rbind_list(
  mutate(tidy(m_inet_acc, "fixed"), what = names(model.frame(m_inet_acc))[1]),
  mutate(tidy(m_hea_use, "fixed"), what = names(model.frame(m_hea_use))[1])) %>%
  filter(term != "(Intercept)") %>% mutate(OR = exp(estimate), conf.low = exp(estimat
e-2*std.error), conf.high = exp(estimate+2*std.error), term=factor(term, levels=rev(u
nique(term)), labels=rev(c("Age\n(10 years)", "Sex\n(1 = female)", "Urbanity of Resid
ence\n(Rural, Small/medium town,\nLarge town/city; 0-2)", "Education\n(years)", "Self
-reported Health\n(very bad - very good; 0-3)", "Self-reported Health Knowledge\n(ver
y bad - very good; 0-3)"))), what = factor(what, levels=rev(unique(what)),
labels=rev(c("Internet Access", "General Health-related\nInternet Use")))) %>%
  ggplot(aes(term, OR, ymin=conf.low, ymax=conf.high, color=what, shape=what)) + geom
_hline(yintercept=1) + geom_pointrange(position = position_dodge(width = .5)) + scale
_y_log10(expression(Odds~Ratio~(log["10"]~scale)), breaks=c(.2,1,2), limits=c(.2,2))
+ coord_flip() + scale_color_brewer("Outcome", palette="Set1") + guides(color = guide
_legend(reverse = T), shape = guide_legend(reverse = T)) + labs(x=NULL) + theme(legen
d.position="right", legend.key=element_blank(), legend.margin=unit(-1/3, "cm"), plot.
margin=unit(rep(.5, 4), "cm"), legend.key.height=unit(1, "cm"), plot.background = ele
ment_rect(color = "black")) + ggtitle("B) Micro-level: Interindividual Differences")
+ scale_shape_discrete("Outcome")
```

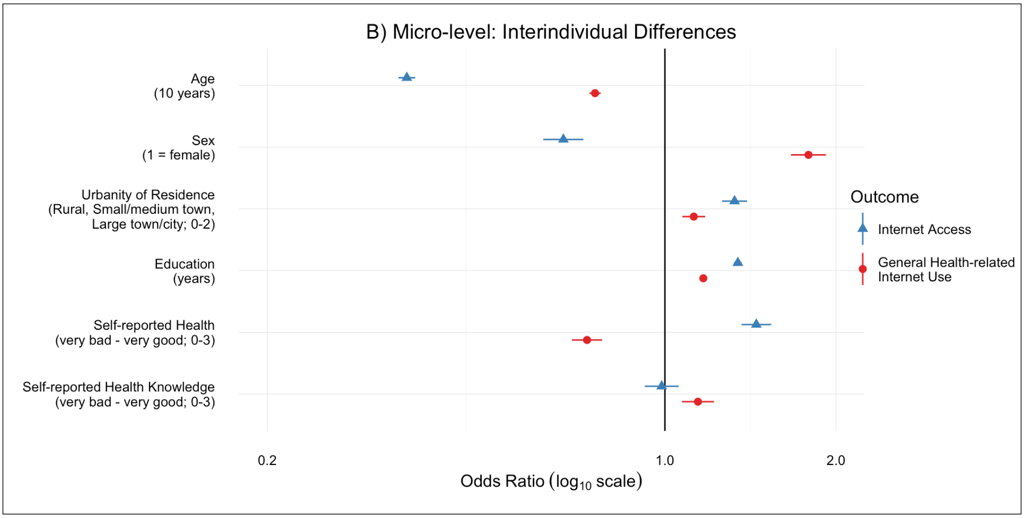


Figure 2: Differences in Internet Access and General Health-related Internet Use Between Countries: Geographical Distribution

```
all_maps = cshp(date = as.Date("2012-06-30"))
europe = all_maps[all_maps$ISO1AL2 %in% unique(d$land), ]
europe$land = as.character(europe$ISO1AL2)
europe = fortify(europe, region="land")
europe_data = cbind(svyby(~dv_inet_acc, by = ~land, design = d_w_svy, FUN = svymean,
na.rm = T)[,1:2],
                    dv_health_use = svyby(~dv_health_use, by = ~land, design = d_w_sv
y, FUN = svymean, na.rm = T)[, 2]) %>% tbl_df()
europe = left_join(europe, europe_data, by = c("id" = "land")) %>% tbl_df
frame_map = fortify(all_maps) %>% tbl_df() %>% filter(long > -11.9 & long < 35.1 & la
t > 33.9 & lat < 71.1)
europe %>% gather(what, P, -(1:7)) %>% mutate(what = factor(what, levels = unique(wha
t), labels=c("Internet Access\n(Proportion of Population)", "General Health-related I
nternet Use\n(Proportion of Internet Users)"))) %>% ggplot() + geom_polygon(data=fram
e_map, aes(long, lat, group), fill="white") + geom_polygon(aes(x=long, y=lat, group=g
roup, fill=P)) + coord_map(xlim = c(-12,35), ylim = c(34,71)) + theme_nothing(legend
= T) + scale_fill_gradient("Proportion", guide = "colorbar") + facet_wrap("what") + g
eom_path(data=frame_map, aes(long, lat, group))
```

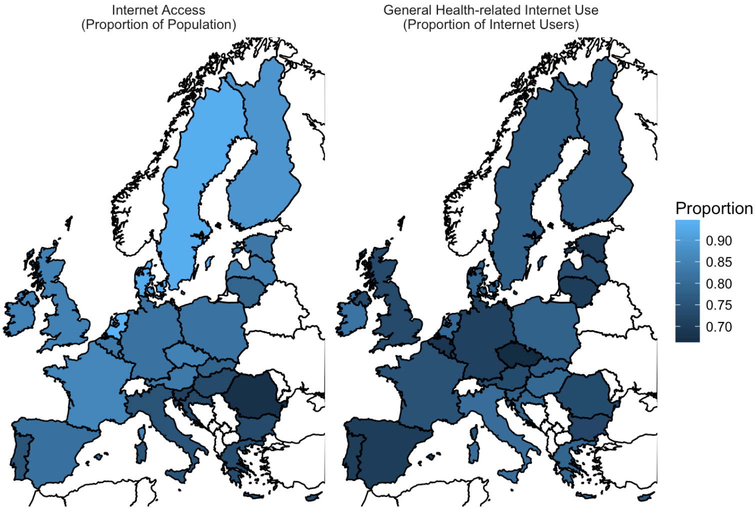
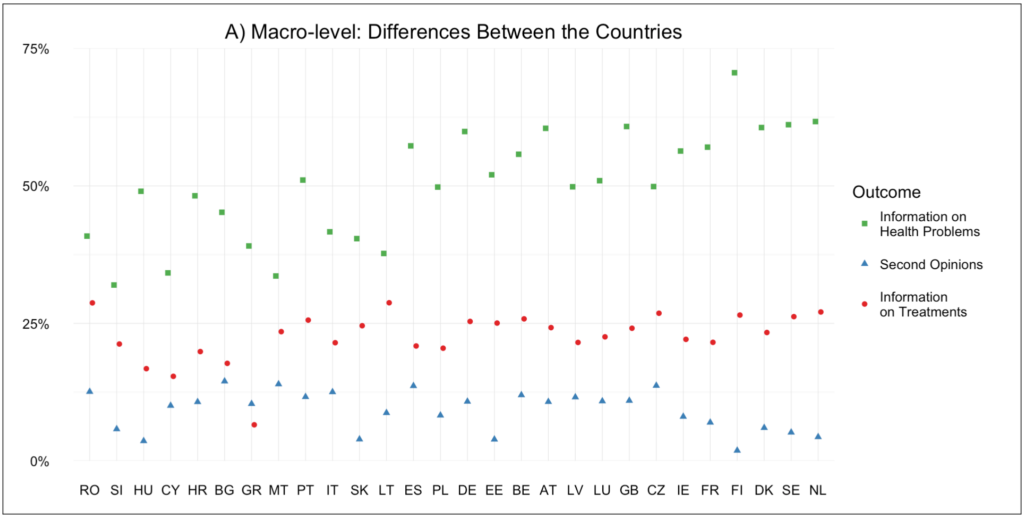


Figure 3: Domain-specific Uses of Health-related Online Information

Panel A

```
cbind(svyby(~dv_spec_ill, by = ~land, design = d_w_svy, FUN = svymean, na.rm = T)
[,1:2],
      dv_spec_2op = svyby(~dv_spec_2op, by = ~land, design = d_w_svy, FUN = svymean,
na.rm = T)[, 2],
      dv_spec_trt = svyby(~dv_spec_trt, by = ~land, design = d_w_svy, FUN = svymean,
na.rm = T)[, 2],
      dv_inet_acc = svyby(~dv_inet_acc, by = ~land, design = d_w_svy, FUN = svymean,
na.rm = T)[, 2]) %>% tbl_df() %>% mutate(land = reorder(land, dv_inet_acc)) %>% selec
t(-dv_inet_acc) %>% gather(dv, P, -land) %>% mutate(dv = factor(dv, levels = unique(d
v), labels = c("Information on\nHealth Problems", "Second Opinions", "Information\non
 Treatments"))) %>% ggplot(aes(land, P, color=dv, shape=dv)) + geom_point(position =
position_dodge(width = .3)) + scale_y_continuous(expand = c(0,0), breaks = c(0,.25,.5
0,.75,1), labels=percent, limits=c(0,.75)) + labs(x=NULL, y=NULL) + theme(panel.margi
n=unit(.5, "lines"), plot.margin=unit(c(.5,.5,.5,.5), "cm"),
legend.key.height=unit(1, "cm"), plot.background = element_rect(color = "black")) + s
cale_color_manual("Outcome", values=rev(brewer.pal(3, name="Set1"))) + scale_shape_ma
nual("Outcome", values=c(15,17,16)) + ggtitle("A) Macro-level: Differences Between th
e Countries")
```



A) Macro-level: Differences Between the Countries
75%
50%
25%
0%
RO SI HU CY HR BG GR MT PT IT SK LT ES PL DE EE BE AT LV LU GB CZ IE FR FI DK SE NL
Outcome
Information on Health Problems
Second Opinions
Information on Treatments

Panel B

```
d_spec = d %>%
  mutate(
    dv_spec_ill = q3.2,
    dv_spec_2op = q3.3,
    dv_spec_trt = q3.4,
    iv_age = d1 / 10,
    iv_female = d2 -1,
    iv_health = ifelse(q33 > 4, NA, q33) * -1 + 4,
    iv_heaknow = ifelse(q36 > 4, NA, q36) * -1 + 4,
    iv_urban = ifelse(d13 > 3, NA, d13 - 1),
    iv_edu = ifelse(d4 == 0, NA, ifelse(d4 == 97, 14, ifelse(d4 == 98, d1, ifelse(d4
== 99, NA, d4)))),
    iv_edu = ifelse(iv_edu < 15, 14, ifelse(iv_edu > 21, 22, iv_edu))) %>%
  select(starts_with("dv_"), starts_with("iv_"), land) %>% na.omit

m_spec_ill = glmer(dv_spec_ill ~ iv_age + iv_female + iv_urban + iv_edu + iv_health +
 iv_heaknow + (1 | land), family = binomial, d_spec, nAGQ = 0)

m_spec_2op = glmer(dv_spec_2op ~ iv_age + iv_female + iv_urban + iv_edu + iv_health +
 iv_heaknow + (1 | land), family = binomial, d_spec, nAGQ = 0)

m_spec_trt = glmer(dv_spec_trt ~ iv_age + iv_female + iv_urban + iv_edu + iv_health +
 iv_heaknow + (1 | land), family = binomial, d_spec, nAGQ = 0)

rbind_list(
  mutate(tidy(m_spec_ill, "fixed"), what = names(model.frame(m_spec_ill))[1]),
  mutate(tidy(m_spec_2op, "fixed"), what = names(model.frame(m_spec_2op))[1]),
  mutate(tidy(m_spec_trt, "fixed"), what = names(model.frame(m_spec_trt))[1])) %>%
  filter(term != "(Intercept)") %>% mutate(OR = exp(estimate), conf.low = exp(estimat
e-2*std.error), conf.high = exp(estimate+2*std.error), term=factor(term, levels=rev(u
nique(term)), labels=rev(c("Age\n(10 years)", "Sex\n(1 = female)", "Urbanity of Resid
ence\n(Rural, Small/medium town,\nLarge town/city; 0-2)", "Education\n(years)", "Self
-reported Health\n(very bad - very good; 0-3)", "Self-reported Health Knowledge\n(ver
y bad - very good; 0-3)"))), what = factor(what, levels=rev(unique(what)),
labels=rev(c("Information on\nHealth Problems", "Second Opinions", "Information\non T
reatments")))) %>%
  ggplot(aes(term, OR, ymin=conf.low, ymax=conf.high, color=what, shape=what)) + geom
_hline(yintercept=1) + geom_pointrange(position = position_dodge(width = .5)) + scale
_y_log10(expression(Odds~Ratio~(log["10"]~scale)), breaks=c(.2,1,2),
limits=c(.2,2.02)) + coord_flip() + scale_color_brewer("Outcome", palette="Set1") + g
uides(color = guide_legend(reverse = T), shape = guide_legend(reverse = T)) +
labs(x=NULL) + theme(legend.position="right", legend.key=element_blank(), legend.marg
in=unit(-1/3, "cm"), plot.margin=unit(rep(.5, 4), "cm"), legend.key.height=unit(1, "c
m"), plot.background = element_rect(color = "black")) + ggtitle("B) Micro-level: Inte
rindividual Differences") + scale_shape_discrete("Outcome")
```

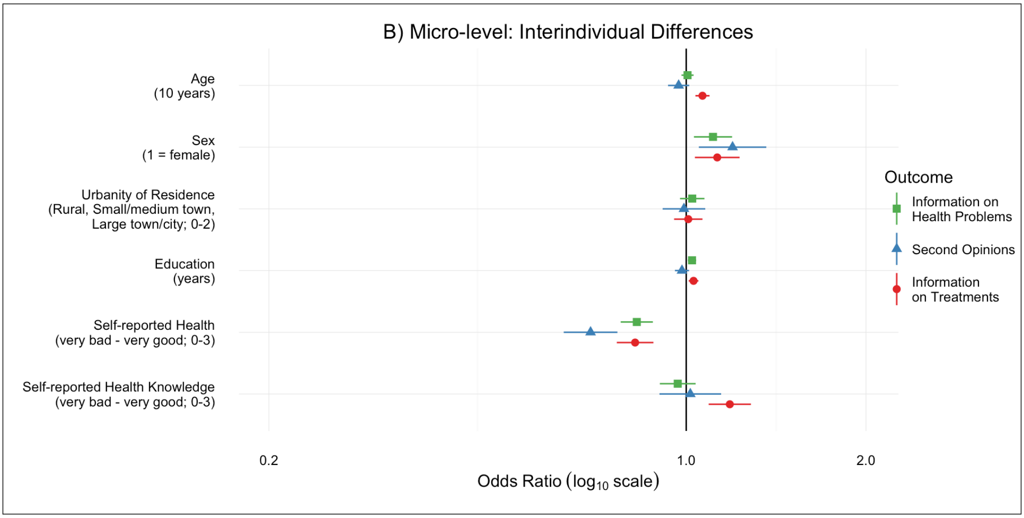
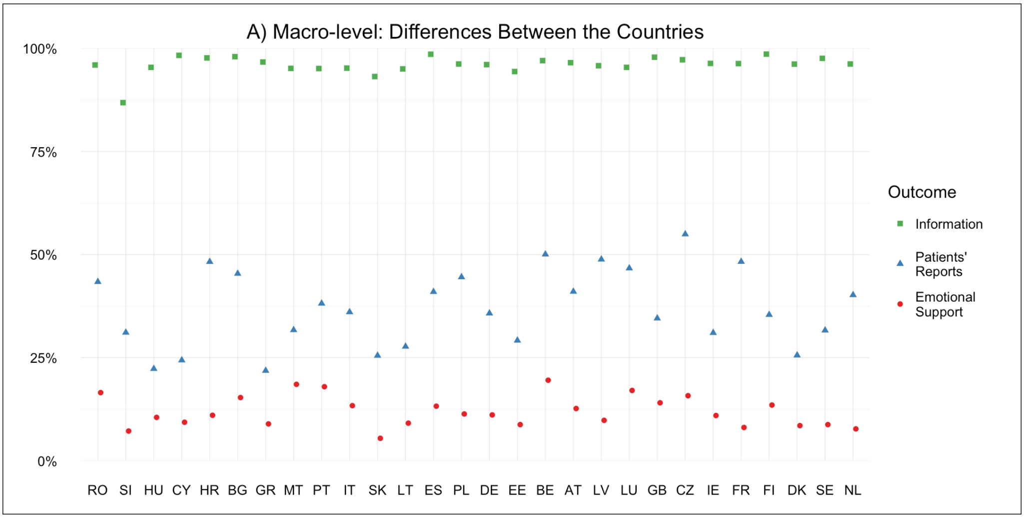


Figure 4: Use of Types of Health-related Online Information

Panel A

```
cbind(svyby(~dv_uses_info, by = ~land, design = d_w_svy, FUN = svymean, na.rm = T)
[,1:2],
      dv_uses_experience = svyby(~dv_uses_experience, by = ~land, design = d_w_svy, F
UN = svymean, na.rm = T)[, 2],
      dv_uses_support = svyby(~dv_uses_support, by = ~land, design = d_w_svy, FUN = s
vymean, na.rm = T)[, 2],
      dv_inet_acc = svyby(~dv_inet_acc, by = ~land, design = d_w_svy, FUN = svymean,
na.rm = T)[, 2]) %>% tbl_df() %>% mutate(land = reorder(land, dv_inet_acc)) %>% selec
t(-dv_inet_acc) %>% gather(dv, P, -land) %>% mutate(dv = factor(dv, levels = unique(d
v), labels = c("Information", "Patients'\nReports", "Emotional\nSupport"))) %>% ggplo
t(aes(land, P, color=dv, shape=dv)) + geom_point(position = position_dodge(width =
.3)) + scale_y_continuous(expand = c(0,0), breaks = c(0,.25,.50,.75,1), labels=percen
t, limits=c(0,1)) + labs(x=NULL, y=NULL) + theme(panel.margin=unit(.5, "lines"), plo
t.margin=unit(c(.5,.5,.5,.5), "cm"), legend.key.height=unit(1, "cm"), plot.background
 = element_rect(color = "black")) + scale_color_manual("Outcome", values=rev(brewer.p
al(3, name="Set1"))) + scale_shape_manual("Outcome", values=c(15,17,16)) +
ggtitle("A) Macro-level: Differences Between the Countries")
```


A) Macro-level: Differences Between the Countries
100%
75%
50%
25%
0%
RO SI HU CY HR BG GR MT PT IT SK LT ES PL DE EE BE AT LV LU GB CZ IE FR FI DK SE NL
Outcome
Information
Patients' Reports
Emotional Support

Panel B

```
d_uses = d %>%
  mutate(
    dv_uses_info = as.numeric(apply(select(., one_of("q8.1", "q8.2", "q8.3", "q8.4",
"q8.6", "q12.1", "q12.2", "q12.3", "q12.4", "q12.6", "q12.7", "q16.1", "q16.2", "q16.
3", "q16.4", "q16.6", "q16.7")), 1, function(x) mean(x, na.rm=T) > 0)),
    dv_uses_experience = as.numeric(apply(select(., one_of("q8.5", "q12.5",
"q16.5")), 1, function(x) mean(x, na.rm=T) > 0)),
    dv_uses_support = as.numeric(apply(select(., one_of("q8.7", "q12.8", "q16.8")),
1, function(x) mean(x, na.rm=T) > 0)),
    iv_age = d1 / 10,
    iv_female = d2 -1,
    iv_health = ifelse(q33 > 4, NA, q33) * -1 + 4,
    iv_heaknow = ifelse(q36 > 4, NA, q36) * -1 + 4,
    iv_urban = ifelse(d13 > 3, NA, d13 - 1),
    iv_edu = ifelse(d4 == 0, NA, ifelse(d4 == 97, 14, ifelse(d4 == 98, d1, ifelse(d4
== 99, NA, d4)))),
    iv_edu = ifelse(iv_edu < 15, 14, ifelse(iv_edu > 21, 22, iv_edu)),
    iv_onlyothers = as.numeric(apply(select(., one_of("q11.1", "q15.1", "q19.1")), 1,
 function(x) mean(x, na.rm=T) == 0))) %>%
  select(starts_with("dv_"), starts_with("iv_"), land) %>% na.omit

m_uses_all = lapply(names(select(d_uses, starts_with("dv_uses"))), function(i){
  frml = formula(paste(i, "~ iv_age + iv_female + iv_urban + iv_edu + iv_health + iv_
heaknow + iv_onlyothers + (1 | land)"))
  glmer(frml, data = d_uses, family = binomial, nAGQ = 0)
})
names(m_uses_all) = names(select(d_uses, starts_with("dv_uses")))

rbind_all(lapply(m_uses_all, function(i) tidy(i, "fixed"))) %>% mutate(what = rep(nam
es(m_uses_all), each = nrow(tidy(m_uses_all[[1]], "fixed")))) %>%
  filter(term != "(Intercept)") %>% mutate(OR = exp(estimate), conf.low = exp(estimat
e-2*std.error), conf.high = exp(estimate+2*std.error), term=factor(term, levels=rev(u
nique(term)), labels=rev(c("Age\n(10 years)", "Sex\n(1 = female)", "Urbanity of Resid
ence\n(Rural, Small/medium town,\nLarge town/city; 0-2)", "Education\n(years)", "Self
-reported Health\n(very bad - very good; 0-3)", "Self-reported Health Knowledge\n(ver
y bad - very good; 0-3)", "Surrogate information seeking\n(1 = only for others)"))),
what = factor(what, levels=rev(unique(what)), labels=rev(c("Information", "Patient
s'\nReports", "Emotional\nSupport")))) %>%
  ggplot(aes(term, OR, ymin=conf.low, ymax=conf.high, color=what, shape=what)) + geom
_hline(yintercept=1) + geom_pointrange(position = position_dodge(width = .5)) + scale
_y_log10(expression(Odds~Ratio~(log["10"]~scale)), breaks=c(.2,1,2), limits=c(.2,2))
+ coord_flip() + scale_color_brewer("Outcome", palette="Set1") + guides(color = guide
_legend(reverse = T), shape = guide_legend(reverse = T)) + labs(x=NULL) + theme(legen
d.position="right", legend.key=element_blank(), legend.margin=unit(-1/3, "cm"), plot.
margin=unit(rep(.5, 4), "cm"), legend.key.height=unit(1, "cm"), plot.background = ele
ment_rect(color = "black")) + ggtitle("B) Micro-level: Interindividual Differences")
+ scale_shape_discrete("Outcome")
```

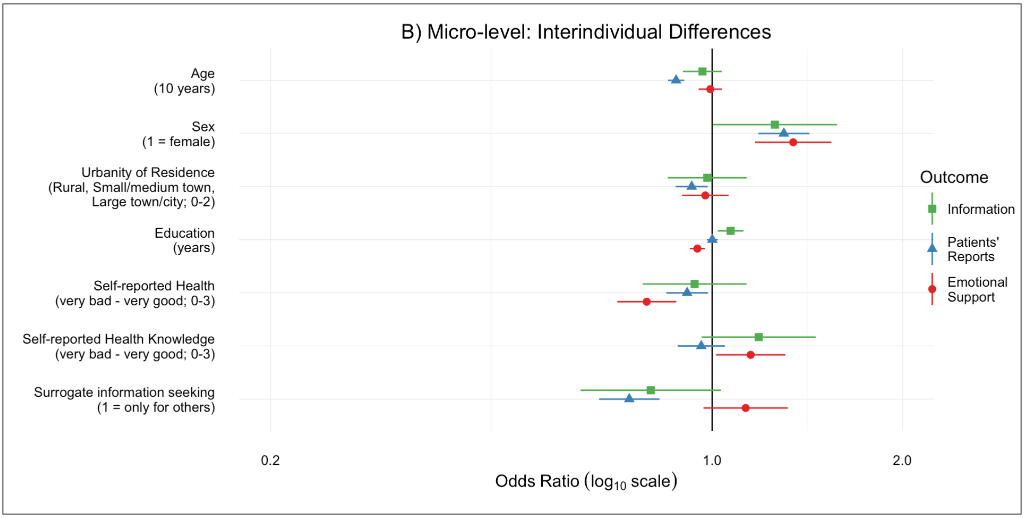
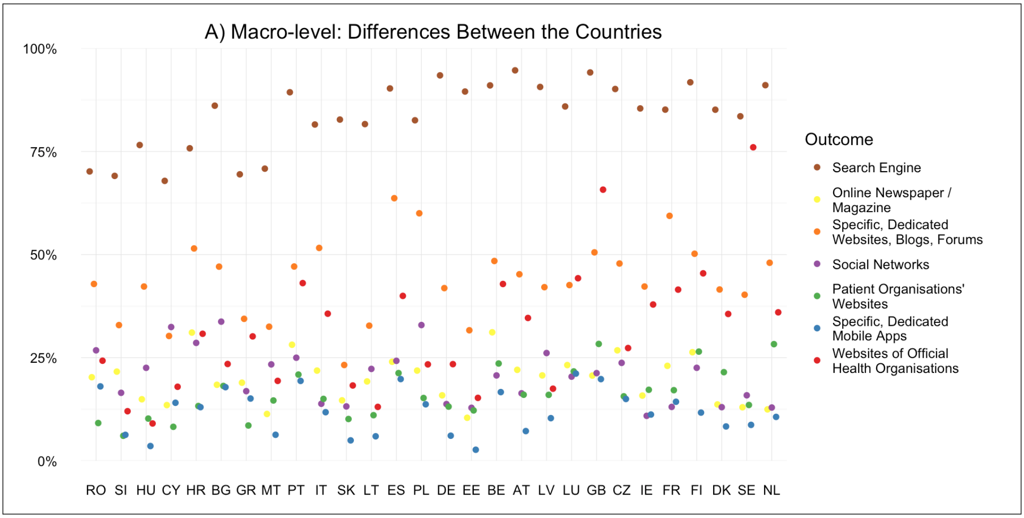


Figure 5: Use of Online Information Sources

Panel A

```
cbind(svyby(~dv_srcrs_search, by = ~land, design = d_w_svy, FUN = svymean, na.rm = T)
[,1:2],
                 dv_srcrs_journ = svyby(~dv_srcrs_journ, by = ~land, design =
d_w_svy, FUN = svymean, na.rm = T)[, 2],
                 dv_srcrs_special = svyby(~dv_srcrs_special, by = ~land, design = d_w
_svy, FUN = svymean, na.rm = T)[, 2],
                 dv_srcrs_social = svyby(~dv_srcrs_social, by = ~land, design = d_w_s
vy, FUN = svymean, na.rm = T)[, 2],
      dv_srcrs_patient = svyby(~dv_srcrs_patient, by = ~land, design = d_w_svy, FUN =
 svymean, na.rm = T)[, 2],
                 dv_srcrs_app = svyby(~dv_srcrs_app, by = ~land, design = d_w_svy, FU
N = svymean, na.rm = T)[, 2],
                 dv_srcrs_official = svyby(~dv_srcrs_official, by = ~land, design = d
_w_svy, FUN = svymean, na.rm = T)[, 2],
                 dv_inet_acc = svyby(~dv_inet_acc, by = ~land, design = d_w_svy, FUN
= svymean, na.rm = T)[, 2]) %>% tbl_df() %>% mutate(land = reorder(land,
dv_inet_acc)) %>% select(-dv_inet_acc) %>% gather(dv, P, -land) %>% mutate(dv = facto
r(dv, levels = unique(dv), labels = c("Search Engine", "Online Newspaper
/\nMagazine", "Specific, Dedicated\nWebsites, Blogs, Forums", "Social Networks", "Pat
ient Organisations'\nWebsites", "Specific, Dedicated\nMobile Apps", "Websites of Offi
cial\nHealth Organisations"))) %>% ggplot(aes(land, P, color=dv)) + geom_point(positi
on = position_dodge(width = .6)) + scale_y_continuous(expand = c(0,0), breaks = c(0,.
25,.50,.75,1), labels=percent, limits=c(0,1)) + labs(x=NULL, y=NULL) + theme(panel.ma
rgin=unit(.5, "lines"), plot.margin=unit(c(.5,.5,.5,.5), "cm"), legend.key.height=uni
t(.8, "cm"), plot.background = element_rect(color = "black")) + scale_color_manual("O
utcome", values=rev(brewer.pal(7, name="Set1"))) + scale_shape_manual("Outcome", valu
es=letters[7:1]) + ggtitle("A) Macro-level: Differences Between the Countries")
```


A) Macro-level: Differences Between the Countries
100%
75%
50%
25%
0%
RO SI HU CY HR BG GR MT PT IT SK LT ES PL DE EE BE AT LV LU GB CZ IE FR FI DK SE NL
Outcome
Search Engine
Online Newspaper / Magazine
Specific, Dedicated Websites, Blogs, Forums
Social Networks
Patient Organisations' Websites
Specific, Dedicated Mobile Apps
Websites of Official Health Organisations

Panel B

```
d_srcs_all = d %>%
  mutate(
    dv_srcrs_search = as.numeric(apply(select(., one_of("q10.1", "q14.1", "q18.1")),
1, function(x) mean(x, na.rm=T) > 0)),
    dv_srcrs_journ = as.numeric(apply(select(., one_of("q10.2", "q14.2", "q18.2")),
1, function(x) mean(x, na.rm=T) > 0)),
    dv_srcrs_special = as.numeric(apply(select(., one_of("q10.3", "q14.3", "q18.3")),
 1, function(x) mean(x, na.rm=T) > 0)),
    dv_srcrs_social = as.numeric(apply(select(., one_of("q10.4", "q14.4", "q18.4")),
1, function(x) mean(x, na.rm=T) > 0)),
    dv_srcrs_patient = as.numeric(apply(select(., one_of("q10.5", "q14.5", "q18.5")),
 1, function(x) mean(x, na.rm=T) > 0)),
    dv_srcrs_app = as.numeric(apply(select(., one_of("q10.6", "q14.6", "q18.6")), 1,
function(x) mean(x, na.rm=T) > 0)),
    dv_srcrs_official = as.numeric(apply(select(., one_of("q10.7", "q14.7",
"q18.7")), 1, function(x) mean(x, na.rm=T) > 0)),
    iv_age = d1 / 10,
    iv_female = d2 -1,
    iv_health = ifelse(q33 > 4, NA, q33) * -1 + 4,
    iv_heaknow = ifelse(q36 > 4, NA, q36) * -1 + 4,
    iv_urban = ifelse(d13 > 3, NA, d13 - 1),
    iv_edu = ifelse(d4 == 0, NA, ifelse(d4 == 97, 14, ifelse(d4 == 98, d1, ifelse(d4
== 99, NA, d4)))),
    iv_edu = ifelse(iv_edu < 15, 14, ifelse(iv_edu > 21, 22, iv_edu)),
    iv_onlyothers = as.numeric(apply(select(., one_of("q11.1", "q15.1", "q19.1")), 1,
 function(x) mean(x, na.rm=T) == 0))) %>%
  select(starts_with("dv_"), starts_with("iv_"), land) %>% na.omit

m_srcs_all = lapply(names(select(d_srcs_all, starts_with("dv_srcrs"))), function(i){
  frml = formula(paste(i, "~ iv_age + iv_female + iv_urban + iv_edu + iv_health + iv_
heaknow + iv_onlyothers + (1 | land)"))
  glmer(frml, data = d_srcs_all, family = binomial, nAGQ = 0)
})
names(m_srcs_all) = names(select(d_srcs_all, starts_with("dv_srcrs")))

rbind_all(lapply(m_srcs_all, function(i) tidy(i, "fixed"))) %>% mutate(what = rep(nam
es(m_srcs_all), each = nrow(tidy(m_srcs_all[[1]], "fixed")))) %>%
  filter(term != "(Intercept)") %>% mutate(OR = exp(estimate), conf.low = exp(estimat
e-2*std.error), conf.high = exp(estimate+2*std.error), term=factor(term, levels=rev(u
nique(term)), labels=rev(c("Age\n(10 years)", "Sex\n(1 = female)", "Urbanity of Resid
ence\n(Rural, Small/medium town,\nLarge town/city; 0-2)", "Education\n(years)", "Self
-reported Health\n(very bad - very good; 0-3)", "Self-reported Health Knowledge\n(ver
y bad - very good; 0-3)", "Surrogate information seeking\n(1 = only for others)"))),
what = factor(what, levels=rev(unique(what)), labels=rev(c("Search Engine", "Online N
ewspaper /\nMagazine", "Specific, Dedicated\nWebsites, Blogs, Forums", "Social Networ
ks", "Patient Organisations'\nWebsites", "Specific, Dedicated\nMobile Apps", "Website
s of Official\nHealth Organisations")))) %>%
  ggplot(aes(term, OR, ymin=conf.low, ymax=conf.high, color=what)) + geom_hline(yinte
rcept=1) + geom_pointrange(position = position_dodge(width = .6)) + scale_y_log10(exp
ression(Odds~Ratio~(log["10"]~scale)), breaks=c(.2,1,2), limits=c(.2,2)) +
coord_flip() + scale_color_brewer("Outcome", palette="Set1") + guides(color = guide_l
egend(reverse = T), shape = guide_legend(reverse = T)) + labs(x=NULL) + theme(legend.
position="right", legend.key=element_blank(), legend.margin=unit(-1/3, "cm"), plot.ma
rgin=unit(rep(.5, 4), "cm"), legend.key.height=unit(1, "cm"), plot.background = eleme
nt_rect(color = "black")) + ggtitle("B) Micro-level: Interindividual Differences") +
scale_shape_manual("Outcome", values = letters[1:7])
```

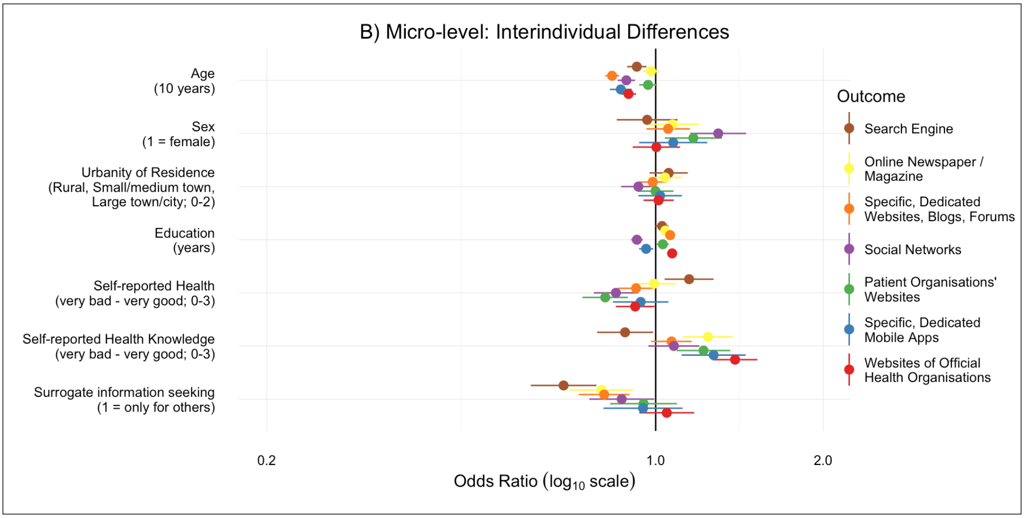
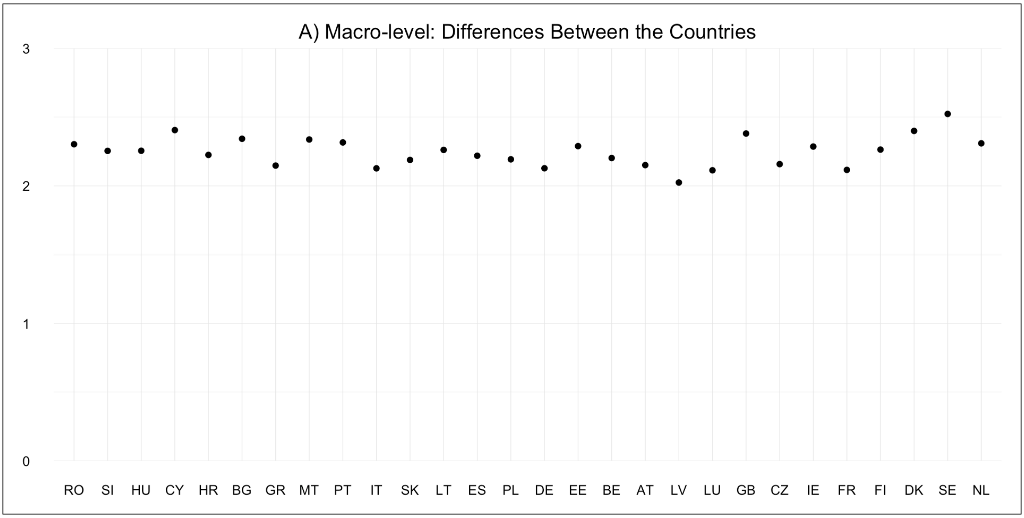


Figure 6: Online Health Literacy

Panel A

```
cbind(svyby(~dv_literacy, by = ~land, design = d_w_svy, FUN = svymean, na.rm = T)
[,1:2],
                 dv_inet_acc = svyby(~dv_inet_acc, by = ~land, design = d_w_svy, FUN
= svymean, na.rm = T)[, 2]) %>% tbl_df() %>% mutate(land = reorder(land,
dv_inet_acc)) %>% select(-dv_inet_acc) %>% ggplot(aes(land, dv_literacy)) + geom_poin
t() + scale_y_continuous(expand = c(0,0), breaks = 0:3, limits=c(0,3)) + labs(x=NULL,
 y=NULL) + theme(panel.margin=unit(.5, "lines"), plot.margin=unit(c(.5,.5,.5,.5), "c
m"), plot.background = element_rect(color = "black")) + ggtitle("A) Macro-level: Diff
erences Between the Countries")
```



Panel B

```
d_hea_lit = d %>%
  mutate(
    dv_literacy = rowMeans(mutate_each(select(., starts_with("q21")[c(1,3:6)]),
funs(ifelse(. > 4, NA, .))), na.rm=T) * -1 + 4,
    iv_age = d1 / 10,
    iv_female = d2 -1,
    iv_health = ifelse(q33 > 4, NA, q33) * -1 + 4,
    iv_heaknow = ifelse(q36 > 4, NA, q36) * -1 + 4,
    iv_urban = ifelse(d13 > 3, NA, d13 - 1),
    iv_edu = ifelse(d4 == 0, NA, ifelse(d4 == 97, 14, ifelse(d4 == 98, d1, ifelse(d4
== 99, NA, d4)))),
    iv_edu = ifelse(iv_edu < 15, 14, ifelse(iv_edu > 21, 22, iv_edu))) %>%
  select(starts_with("dv_"), starts_with("iv_"), land) %>% na.omit

m_hea_lit = lmer(dv_literacy ~ iv_age + iv_female + iv_urban + iv_edu + iv_health + i
v_heaknow + (1 | land), d_hea_lit)

tidy(m_hea_lit, "fixed") %>% filter(term != "(Intercept)") %>% mutate(conf.low = esti
mate-2*std.error, conf.high = estimate+2*std.error, term=factor(term, levels=rev(uniq
ue(term)), labels=rev(c("Age\n(10 years)", "Sex\n(1 = female)", "Urbanity of Residenc
e\n(Rural, Small/medium town,\nLarge town/city; 0-2)", "Education\n(years)", "Self-re
ported Health\n(very bad - very good; 0-3)", "Self-reported Health Knowledge\n(very b
ad - very good; 0-3)")))) %>% ggplot(aes(estimate, term)) + geom_point() +
geom_vline(xintercept=0) + geom_errorbarh(aes(xmin=conf.low, xmax=conf.high),
height=0) + scale_x_continuous("Linear Coefficient", limits=c(-.31,.31),
breaks=c(-.3, -.15,0,.15,.3)) + labs(y=NULL) + theme(plot.margin=unit(rep(.5, 4), "c
m"), plot.background = element_rect(color = "black")) + ggtitle("B) Micro-level: Inte
rindividual Differences")
```

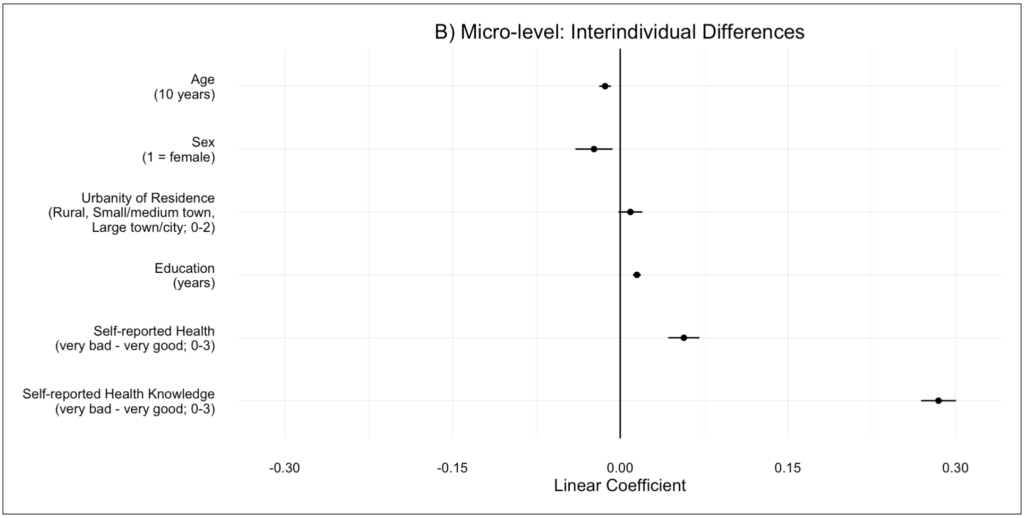


Table A1: Country-Level Proportions for Figures 1 & 2

```
full_table = d_w %>% ungroup %>% select(-dv_noreason, -dv_noaccess, -dv_noability) %
>% as_survey_design(ids = uniqid, weights = w1) %>% group_by(land) %>%
summarise_each(funs(survey_mean(., na.rm = T)), -w1, -uniqid) %>% select(-
ends_with("_se"))

full_table[, c(1:3)] %>% kable(digits = 2, col.names = c("Country", "Internet
Access", "General Health-Related Internet Use"))
```

Country	Internet Access	General Health-Related Internet Use
AT	0.83	0.73
BE	0.83	0.68
BG	0.73	0.72
CY	0.73	0.75
CZ	0.85	0.66
DE	0.82	0.71
DK	0.93	0.79
EE	0.83	0.71
ES	0.81	0.70
FI	0.89	0.79
FR	0.86	0.75
GB	0.85	0.73
GR	0.73	0.79
HR	0.73	0.79
HU	0.73	0.80
IE	0.85	0.82
IT	0.75	0.81
LT	0.79	0.70
LU	0.85	0.72
LV	0.84	0.74
MT	0.74	0.67
NL	0.94	0.80
PL	0.82	0.78
PT	0.75	0.70
RO	0.68	0.73
SE	0.93	0.78
SI	0.71	0.77

Country	Internet Access	General Health-Related Internet Use
SK	0.78	0.77

Table A2: Country-Level Proportions for Figure 3

```
full_table[, c(1, 5:7)] %>% kable(digits = 2, col.names = c("Country", "Information o
n Health Problems", "Second Opinions", "Information on Treatments"))
```

Country	Information on Health Problems	Second Opinions	Information on Treatments
AT	0.60	0.11	0.24
BE	0.56	0.12	0.26
BG	0.45	0.14	0.18
CY	0.34	0.10	0.15
CZ	0.50	0.14	0.27
DE	0.60	0.11	0.25
DK	0.61	0.06	0.23
EE	0.52	0.04	0.25
ES	0.57	0.14	0.21
FI	0.71	0.02	0.27
FR	0.57	0.07	0.22
GB	0.61	0.11	0.24
GR	0.39	0.10	0.07
HR	0.48	0.11	0.20
HU	0.49	0.04	0.17
IE	0.56	0.08	0.22
IT	0.42	0.13	0.21
LT	0.38	0.09	0.29
LU	0.51	0.11	0.23
LV	0.50	0.12	0.22
MT	0.34	0.14	0.23
NL	0.62	0.04	0.27
PL	0.50	0.08	0.20
PT	0.51	0.12	0.26
RO	0.41	0.13	0.29

Country	Information on Health Problems	Second Opinions	Information on Treatments
SE	0.61	0.05	0.26
SI	0.32	0.06	0.21
SK	0.40	0.04	0.25

Table A3: Country-Level Proportions for Figure 4

```
full_table %>% select(land, contains("uses")) %>% kable(digits = 2, col.names = c("Co
untry", "Information", "Patients' Reports", "Emotional Support"))
```

Country	Information	Patients' Reports	Emotional Support
AT	0.97	0.41	0.13
BE	0.97	0.50	0.20
BG	0.98	0.45	0.15
CY	0.98	0.24	0.09
CZ	0.97	0.55	0.16
DE	0.96	0.36	0.11
DK	0.96	0.26	0.09
EE	0.94	0.29	0.09
ES	0.99	0.41	0.13
FI	0.99	0.35	0.14
FR	0.96	0.48	0.08
GB	0.98	0.35	0.14
GR	0.97	0.22	0.09
HR	0.98	0.48	0.11
HU	0.95	0.22	0.11
IE	0.96	0.31	0.11
IT	0.95	0.36	0.13
LT	0.95	0.28	0.09
LU	0.95	0.47	0.17
LV	0.96	0.49	0.10
MT	0.95	0.32	0.19
NL	0.96	0.40	0.08
PL	0.96	0.45	0.11

Country	Information	Patients' Reports	Emotional Support
PT	0.95	0.38	0.18
RO	0.96	0.43	0.17
SE	0.98	0.32	0.09
SI	0.87	0.31	0.07
SK	0.93	0.26	0.05

Table A4: Country-Level Proportions for Figure 5

```
full_table %>% select(land, contains("srcrs")) %>% kable(digits = 2, col.names = c("C
ountry", "Search Engine", "Online Newspaper / Magazine", "Specific, Dedicated Website
s, Blogs, Forums", "Social Networks", "Patient Organisations' Websites", "Specific, D
edicated Mobile Apps", "Websites of Official Health Organisations"))
```

Country	Search Engine	Online Newspaper / Magazine	Specific, Dedicated Websites, Blogs, Forums	Social Networks	Patient Organisations' Websites	Specific, Dedicated Mobile Apps	Websites of Official Health Organisations
AT	0.95	0.22	0.45	0.16	0.16	0.07	0.35
BE	0.91	0.31	0.48	0.21	0.24	0.17	0.43
BG	0.86	0.18	0.47	0.34	0.18	0.18	0.23
CY	0.68	0.14	0.30	0.32	0.08	0.14	0.18
CZ	0.90	0.27	0.48	0.24	0.16	0.15	0.27
DE	0.93	0.16	0.42	0.14	0.13	0.06	0.23
DK	0.85	0.14	0.42	0.13	0.21	0.08	0.36
EE	0.90	0.10	0.32	0.13	0.12	0.03	0.15
ES	0.90	0.24	0.64	0.24	0.21	0.20	0.40
FI	0.92	0.26	0.50	0.23	0.26	0.12	0.45
FR	0.85	0.23	0.59	0.13	0.17	0.14	0.42
GB	0.94	0.21	0.51	0.21	0.28	0.20	0.66
GR	0.69	0.19	0.34	0.17	0.09	0.15	0.30
HR	0.76	0.31	0.51	0.29	0.13	0.13	0.31
HU	0.77	0.15	0.42	0.23	0.10	0.04	0.09
IE	0.85	0.16	0.42	0.11	0.17	0.11	0.38
IT	0.82	0.22	0.52	0.14	0.15	0.12	0.36
LT	0.82	0.19	0.33	0.22	0.11	0.06	0.13

Country	Search Engine	Online Newspaper / Magazine	Specific, Dedicated Websites, Blogs, Forums	Social Networks	Patient Organisations' Websites	Specific, Dedicated Mobile Apps	Websites of Official Health Organisations
LU	0.86	0.23	0.43	0.20	0.22	0.21	0.44
LV	0.91	0.21	0.42	0.26	0.16	0.10	0.17
MT	0.71	0.11	0.33	0.23	0.15	0.06	0.19
NL	0.91	0.12	0.48	0.13	0.28	0.11	0.36
PL	0.83	0.22	0.60	0.33	0.15	0.14	0.23
PT	0.89	0.28	0.47	0.25	0.21	0.19	0.43
RO	0.70	0.20	0.43	0.27	0.09	0.18	0.24
SE	0.84	0.13	0.40	0.16	0.14	0.09	0.76
SI	0.69	0.22	0.33	0.16	0.06	0.06	0.12
SK	0.83	0.15	0.23	0.13	0.10	0.05	0.18