



MESTRADO EM
EDUCAÇÃO PARA A SAÚDE

UNIVERSIDADE DO PORTO
FACULDADE DE MEDICINA
FACULDADE DE PSICOLOGIA E CIÊNCIAS DA EDUCAÇÃO

Carlos Daniel Macedo Ferreira

Quality of websites providing information on breast and prostate cancer in Portuguese language

Porto, 2011

Carlos Daniel Macedo Ferreira

**Quality of websites providing information on
breast and prostate cancer
in Portuguese language**

Dissertação de candidatura ao grau de
Mestre em Educação para a Saúde
apresentada à Faculdade de Medicina da Universidade do Porto

Investigação realizada no Instituto de Saúde Pública da Universidade do Porto (ISPUP)
sob a orientação do Professor Doutor Nuno Lunet e Co-orientação da Doutora Susana Silva

Porto, 2011

Agradecimentos

À minha família, por me ter ajudado desde início. Mesmo quando eu não sei, eles descobrem sempre como ajudar.

À Vera, pelo apoio e carinho que teve comigo, e que tão importante foi neste caminho. Estiveste sempre lá. Porque há sempre uma perspectiva mais simples em tudo.

Ao Professor Milton Severo, pela ajuda e pela disponibilidade que sempre mostrou.

À Professora Susana Silva, pela profundidade na reflexão que me incutiu, pela pertinência das observações e por se manter sempre disponível para me ajudar.

Ao Professor Nuno Lunet. Pelo muito que me ensinou, pela atenção e paciência. Porque investigar pode ser uma descoberta aliciante, feita com entusiasmo. Pelo rigor que emprestou a esta investigação, pelo detalhe, porque tudo é importante. Pelo tempo infinito. O tempo que teimava em fugir. “Não há nada que substitua o tempo. O tempo é tudo, Daniel”. É assim mesmo, Professor.

Index

1	Abstract	1
2	Resumo	3
3.	Introduction	6
3.1.	Health literacy	6
3.2.	The Internet as a source of information on cancer	10
3.2.1.	Quality of websites providing health information on the Internet	13
	3.2.1.1. Formal assessment	13
	3.2.1.2. Analysis of the contents	17
	3.2.1.2.1. Readability	18
4.	Aims	20
5.	Methods	21
5.1.	Selection of the websites for analysis	21
5.2.	General characterization of the websites	23
5.3.	Assessment of the formal quality of the websites according to the Discern tool, the principles of the Health On Net code, the eEurope guidelines for health related websites and the American Medical Association guidelines for medical and health information websites	28
5.4.	Analysis of the contents related to screening of breast and prostate cancer	39

5.4.1.	Specific contents on cancer screening	39
5.4.2.	Readability of the contents	43
6.	Results	44
6.1.	Websites selected for analysis	44
6.2.	General characteristics of the websites	47
6.3.	The formal quality of the websites according to the Discern tool, the principles of the Health On Net code, the eEurope guidelines for health related guidelines and the American Medical Association guidelines for medical and health information sites	49
6.3.1.	Reliability of the assessment instruments	55
6.4.	Quality of the contents related to screening of breast and prostate cancer and readability	59
7.	Discussion	65
8.	Conclusion	71
9.	References	72

Index of figures

- Figure 1 Factors and contexts that influence health knowledge, decisions and actions.
- Figure 2 A simplified model of mechanisms linking health literacy to health outcomes.
- Figure 3 Incidence and mortality by cancer (per 100 000 inhabitants), in Portuguese women and men, all ages, in 2008.
- Figure 4 Navigation process for analysis of the websites.
- Figure 5 Standardized procedure followed to classify the websites according to the main subject.
- Figure 6 Standardized procedure followed to classify the websites according to the country of origin.
- Figure 7 Standardized procedure followed to identify the websites according to the intended audience.
- Figure 8 Standardized procedure followed to identify the websites according to the affiliation.
- Figure 9 Standardized procedure followed to identify the media used by the website.
- Figure 10 Procedure followed to the analysis of information on breast cancer screening.
- Figure 11 Procedure followed to the analysis of information on prostate cancer screening.
- Figure 12 Selection of the Internet search results for breast cancer.
- Figure 13 Selection of the Internet search results for prostate cancer.
- Figure 14 Time spent in the evaluation of 54 websites with each of four instruments, according to the order of use of the instruments.

Figure 15 Time spent in the evaluation of 54 websites with each of four instruments, according to the order by which the websites were evaluated.

Figure 16 Association between the scores obtained with the instruments based on the HON code and eEurope guidelines and the ISPUP score.

Figure 17 Quality of the contents on screening in breast cancer websites (n=35).

Figure 18 Quality of the contents on screening in prostate cancer websites (n=43).

Index of tables

- Table 1 Principles of the HON code.
- Table 2 eEurope guidelines for health related websites.
- Table 3 AMA principles for medical and health information sites.
- Table 4 Flesch Reading Ease formula to measure readability and Index Fernandez-Huerta.
- Table 5 Flesch Reading Ease score and school level needed do understand texts.
- Table 6 HON code of conduct – principles and quality criteria.
- Table 7 eEurope guidelines and quality criteria.
- Table 8 AMA guidelines and quality criteria.
- Table 9 General characteristics of the websites selected for analysis.
- Table 10 Compliance of the websites providing information on cancer treatments (n=81) with the Discern criteria.
- Table 11 Compliance of the websites with the principles of the HON code.
- Table 12 Compliance of the websites with the eEurope guidelines for health related websites.
- Table 13 Compliance of the websites with the AMA guidelines for medical and health information sites.
- Table 14 Exploratory factor analysis and internal consistency conducted for the instruments based on the HON code, EEU guidelines and combination of the two scales.
- Table 15 ISPUP-score criteria (adapted from HON code and eEurope guidelines for health related websites).
- Table 16 Quality and readability of the contents on breast cancer screening and prostate cancer screening.

Table 17 Quality of the contents on breast cancer screening according to websites' order of appearance, country of origin, specific for cancer and ISPUP score.

Table 18 Quality of the contents on prostate cancer screening according to websites' order of appearance, country of origin, specific for cancer and ISPUP score.

List of abbreviations

DSC	Discern tool
HON	Health On Net
eEu	eEurope guidelines for health related websites
AMA	American Medical Association
JAMA	Journal of the American Medical Association
URL	Uniform resource locator
PDF	Portable document file
ISPUP	Instituto Saúde Pública da Universidade do Porto
PSA	Prostate specific antigen
DRE	Digital rectal examination
CFI	Comparative fit index
BSE	Breast self-exam
CBE	Clinical breast exam
MRI	Magnetic resonance imaging
EU	European Union
U.S.	United States

1. Abstract

The easy access, fast dissemination of information and availability of different sources make the Internet a privileged vehicle to search and to provide information. Its utilization has increased in the past few years, and currently it is one of the most frequently used sources of both general and health-related information.

Topics related to oncological diseases correspond to an important proportion of the Internet searches on health issues. Specifically breast and prostate cancers are leading causes of oncological morbidity and are among the cancers with the highest relative survival, which contributes for a high interest on these topics by the general population, and in particular by patients and their families.

The overall quality of the websites and the accuracy of the contents on specific topics must be assured, since this is an easily accessible and widely used resource and has the potential to shape health behaviors of a large number of people worldwide.

This dissertation aimed to evaluate the quality of websites providing information on female breast and prostate cancers in Portuguese language, by assessing their compliance with different sets of criteria of formal quality, and evaluating the contents on breast and prostate cancers screening regarding the agreement with the best available evidence on this topic and the readability of the texts.

We searched in the Google® search engine the terms *cancro da prostata* and *cancro da mama*, in September 15 and 16, 2011, respectively. We selected the first 200 results from each independent search and only pages providing information on Portuguese language were included. We identified 47 websites addressing breast cancer and 67 websites with information on prostate cancer. The websites were classified according to main subject, country of origin, intended audience, affiliation and media used to provide information. For the evaluation of their compliance with pre-defined criteria to characterize formal quality we used the Discern tool, and the Health On Net code, eEurope guidelines for health related websites and the AMA guidelines for medical and health information sites, and we propose a simpler instrument, based on the Health On Net code, eEurope guidelines for health related websites – the ISPUP score.

Most of the sites were health-related (71.9%), but only 19.3% addressed exclusively cancer, and 10.5% covered specially issues related to breast or prostate cancer. Approximately half the websites providing information in the Portuguese language were not Portuguese. Nearly all websites (91.2%) were directed to the general population, while 14.1% targeted specifically cancer patients.

Most of the sites did not comply with the quality criteria selected for the analysis of their formal characteristics. The agreement with the criteria of the Discern tool was low. Regarding the HON code, the total agreement was higher with the principles “financial

disclosure”, “transparency” and “justifiability”. Only 7.9% of the websites had the HON Seal. Less than half of the websites complied with each criterion of the eEurope guidelines for health related websites, except for the accountability (82.5%). The highest proportion of websites in total agreement with the AMA principles governing websites was observed for those related with intersite navigation (100%), functional intrasite links (98.8%) and funding and sponsorship (83.2%), and the lowest for those referring to navigation on content (5.3%) and site viewing (3.5%).

Thirty-five of the websites providing information on breast cancer (74%) addressed breast cancer screening, while 43 of those addressing prostate cancer (64%) provided information on screening. We assessed the contents related to screening regarding the methods, effectiveness, potential harms, periodicity, eligibility and instructions on how to be screened. The criteria to assess the accuracy of this information and its adequacy to the Portuguese setting were defined in accordance to the evidence summarized by the United States Preventive Task Force, the European Union Advisory Committee on Cancer Prevention and the specific recommendations for cancer screening in Portugal.

Eighty percent of the websites mentioned mammography as a method for breast cancer screening, but the correct information on the effectiveness of mammography screening was provided in only 14%. Most websites did not provide information on the potential harms of screening.

Among the websites addressing prostate cancer screening, 95.3% mentioned the assessment of the Prostate Specific Antigen (PSA) as a possible screening test, but the correct information on its effectiveness was given in less than 10% of the websites. The possible harms of the prostate cancer screening were correctly referred in a very low percentage of the websites.

We assessed the readability of the websites in the sections with screening-related information, using the Index Fernandez-Huerta. In general, the readability was good, with no significant differences between the breast and prostate cancer websites (median index values for breast vs. prostate cancer websites: 73.1 vs. 69.7, P=0.144).

There were no significant or meaningful associations between formal quality and the quality of the contents on screening.

In conclusion, there is a large margin for improving the quality of the websites providing information on breast and prostate cancer in Portuguese language. Regarding screening, websites providing information on breast cancer tend to provide more accurate information than those addressing prostate cancer.

The present work provides a framework for the standardized assessment of different dimensions for the quality of websites providing information on breast or prostate cancer that may be used for the monitoring of the quality of the health information provided in the Internet.

2. Resumo

A Internet é um meio privilegiado para procurar e disponibilizar informação, por permitir a sua rápida disseminação, permitir o recurso a diversas fontes e por ser facilmente acessível. A utilização da Internet tem aumentado nos últimos anos, sendo que actualmente é uma das fontes mais utilizadas para a procura de informação em geral, e em particular de informação relacionada com a saúde.

Uma grande parte da informação sobre saúde procurada na Internet diz respeito às doenças oncológicas. O cancro da mama e o cancro da próstata são as causas mais importantes de morbilidade oncológica, situando-se entre os cancros com a maior sobrevida relativa, o que leva a um maior interesse sobre estes cancros, pela população em geral, e especificamente pelos doentes e suas famílias.

A qualidade global dos *websites* e a validade dos conteúdos disponibilizados em assuntos específicos deve ser assegurada, uma vez que este recurso é frequentemente utilizado e facilmente acessível, e tem o potencial de moldar comportamentos em saúde de um grande número de pessoas, por todo o mundo.

Esta dissertação teve como objectivo avaliar a qualidade dos *websites* que disponibilizam informação sobre cancro da mama e cancro da próstata, em língua portuguesa, considerando diferentes conjuntos de critérios de qualidade formal, e avaliando os seus conteúdos relativamente ao rastreio do cancro da mama e do cancro da próstata tendo em conta a melhor informação disponível, bem como a legibilidade dos seus textos.

A pesquisa foi efectuada no motor de pesquisa Google®, com as expressões *cancro da próstata* e *cancro da mama*, nos dias 15 e 16 de Setembro de 2011, respectivamente. Foram seleccionados os primeiros 200 resultados de cada pesquisa, e só foram incluídas na análise *websites* escritos em língua portuguesa. Identificamos 47 *websites* com informação sobre cancro da mama e 67 *websites* com informação sobre cancro da próstata. Os *websites* foram classificados de acordo com o assunto principal que abordavam, o seu país de origem, o público-alvo, a sua afiliação e os meios pelo qual disponibilizavam informação. Para avaliar a concordância com critérios pré-definidos que caracterizam a qualidade formal, usámos o *Discern tool*, o *Health On Net code*, as guidelines para *websites* relacionados com saúde do *eEurope* e as guidelines para *websites* com informação médica ou de saúde da *American Medical Association (AMA)*, e propusemos um instrumento mais simples, baseado no *Health On Net code* e nas guidelines para *websites* relacionados com saúde do *eEurope* – o ISPUP score.

A maioria dos *websites* eram relacionados com saúde (71,9%), mas apenas 19,3% tinham como assunto principal o cancro, sendo que 10,5% eram exclusivos do cancro da mama ou do cancro da próstata. Aproximadamente metade dos *websites* que disponibilizavam informação em português não tinha origem em Portugal. Quase todos os

websites (91,2%) eram direcionados para a população em geral, enquanto 14,1% tinham como alvo especificamente os doentes com cancro.

A maioria dos *websites* não cumpria com os critérios de qualidade seleccionados para a análise das suas características formais. A concordância com os critérios do *Discern tool* foi baixa. Relativamente ao *HON code*, a *websites* concordância total foi mais alta com os princípios de “clareza financeira”, “transparência” e de “justificação”. Apenas 7,9% dos *websites* tinham o selo da *HON*. Menos de metade dos *websites* estavam de acordo com cada um dos critérios das guidelines para *websites* relacionados com saúde do *eEurope*, com excepção para o critério da “responsabilidade” (82,5%). A maior proporção de *websites* em acordo total com as guidelines para *websites* com informação médica ou de saúde da *AMA*, foi observada para os princípios relacionados com a “navegabilidade intersite” (100%), com a “funcionalidade dos links intrasite” (98,8%) e com o “financiamento e patrocínio” (83,2%), e a menor foi para os princípios relacionados com a “navegação nos conteúdos” (5,3%) e com a “visualização do website” (3,5%).

Trinta e cinco dos *websites* que disponibilizavam informação sobre o cancro da mama (74%) abordavam o rastreio do cancro da mama, enquanto 43 dos *websites* relacionados com o cancro da próstata (64%) abordavam o rastreio. A avaliação dos conteúdos relacionados com o rastreio teve em conta os métodos para o rastreio, a efectividade, os potenciais efeitos negativos, a periodicidade, a elegibilidade e as instruções para se ser rastreado. Os critérios para se avaliar a correção da informação e a sua adequação para o contexto português, foi definido de acordo com a informação sumariada pela *United States Preventive Services Task Force*, o *European Union Advisory Committee on Cancer Prevention*, e as recomendações específicas para o rastreio do cancro em Portugal. Oitenta por cento dos *websites* mencionaram a mamografia como método de rastreio do cancro da mama, mas apenas 14% *websites* referiam a informação correcta relativamente à sua efectividade. A maioria dos *websites* não mencionava qualquer potencial efeito negativo do rastreio.

Relativamente aos *websites* com informação sobre rastreio do cancro da próstata, 95,3% mencionaram a avaliação do Antigénio Específico da Próstata (PSA) como um possível método de rastreio, mas a informação correcta sobre a sua efectividade foi disponibilizada em apenas 10% dos *websites*. Os potenciais efeitos negativos do rastreio do cancro da próstata foram correctamente abordados numa percentagem muito reduzida dos *websites*.

A avaliação relativa à legibilidade dos textos foi efectuada nas secções com informação relativa ao rastreio, usando a fórmula de Fernandez-Huerta. Em geral, a legibilidade dos *websites* foi boa, sem diferenças significativas entre os *websites* com informação sobre o cancro da mama e cancro da próstata (mediana dos valores do índice para *websites* sobre cancro da mama vs. próstata: 73,1 vs. 69,7, P=0,144).

Não houve associação significativa ou importante entre a qualidade formal dos websites e a qualidade dos seus conteúdos relativamente ao rastreio.

Em conclusão, há uma larga margem para melhorar a qualidade dos websites que disponibilizam informação sobre o cancro da mama e cancro da próstata em língua portuguesa. Relativamente ao rastreio, os websites relacionados com o cancro da mama, tendem a disponibilizar informação mais precisa do que os websites relacionados com o cancro da próstata.

Este trabalho disponibiliza uma estrutura para a avaliação padronizada de diferentes dimensões da qualidade de websites relacionados com cancro da mama e cancro da próstata, que pode ser usada para monitorizar a qualidade da informação em saúde disponível na Internet.

3. Introduction

3.1. Health literacy

Literacy involves a complex set of abilities to access, understand and use the dominant symbolic systems of a culture, allowing personal and community development that varies across populations and settings. In contemporary western societies the information available at media or other sources of electronic texts has been included in these systems. Thus, nowadays literacy includes skills for accessing, understanding and appraising information disseminated through the internet.⁽¹⁾

At the individual level, literacy can be positively influenced by personal ability, level of education and parents' level of education and socioeconomic status. Age can be other important determinant, once literacy tends to increase from the teenage years to the middle forties, when it starts to decline.^(2, 3)

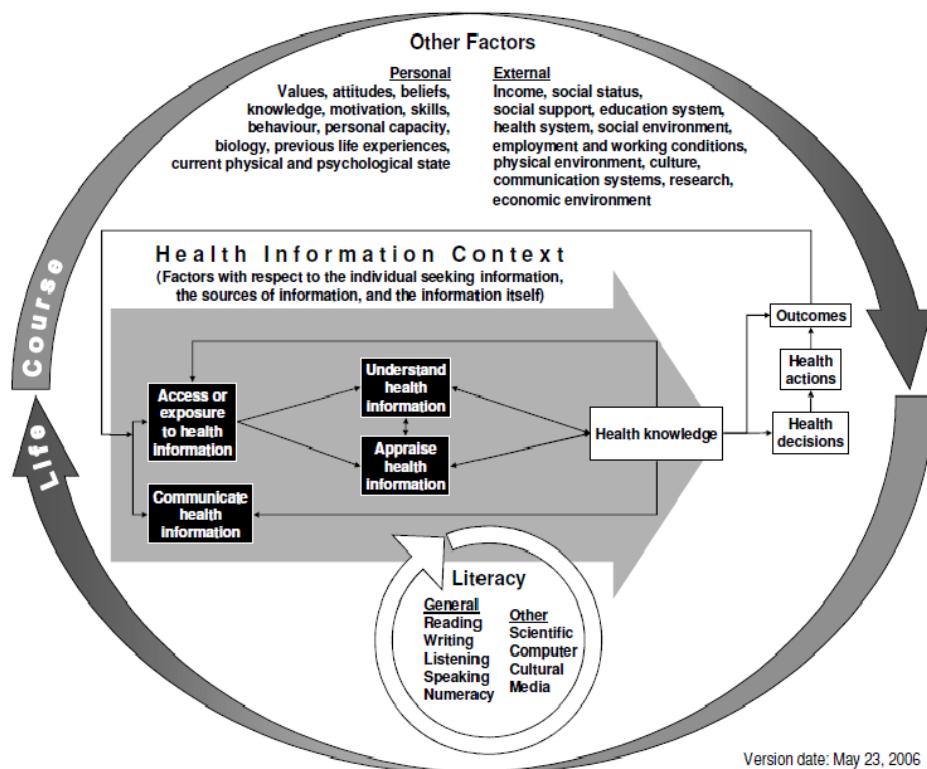
General literacy plays an important role in the health status of individuals and populations.^(4, 5) This is supported by data collected within a cohort of 3260 Medicare managed-care enrollees in 4 US metropolitan areas, that reveal a strong association between low levels of literacy (measured by the reading fluency) and increased mortality rates.⁽⁶⁾ Other study shows the relation between low levels of literacy and poorer health outcomes, including poorer knowledge about health and health care issues, higher hospitalization indexes, lower adherence to general health recommendations and higher prevalence of chronic diseases.⁽⁷⁾

Health literacy can be defined as "the degree to which people are able to access, understand, appraise and communicate information to engage with the demands of different health contexts in order to promote and maintain good health across the life-course."⁽²⁾ There are distinct dimensions of health literacy: functional health literacy (corresponding to the basic skills in reading and writing for effective functioning in health context); interactive health literacy (includes more advanced cognitive literacy and social skills that enable active participation in health care); critical health literacy (the ability to critically analyze and use information to participate in actions that overcome structural barriers to health).⁽⁸⁾ Quantitative literacy, also called health numeracy⁽⁹⁾ refers to the ability to perform simply mathematical operations^(10, 11) and its growing importance is related with the dissemination of information about probabilities concerning risks and protecting factors on health and about survival and mortality curves.⁽¹⁰⁻¹²⁾

Health knowledge can be influenced by the health information context and general literacy levels. These determinants, as well as the access or exposure to health information, can affect the possibility or capability to understand and appraise health information, influencing health knowledge, and health decisions or outcomes as a consequence.

Nevertheless, there are other important factors affecting health knowledge, that can be personal or external to the individual, and that may change during a life course, as shown in figure 1.⁽⁸⁾

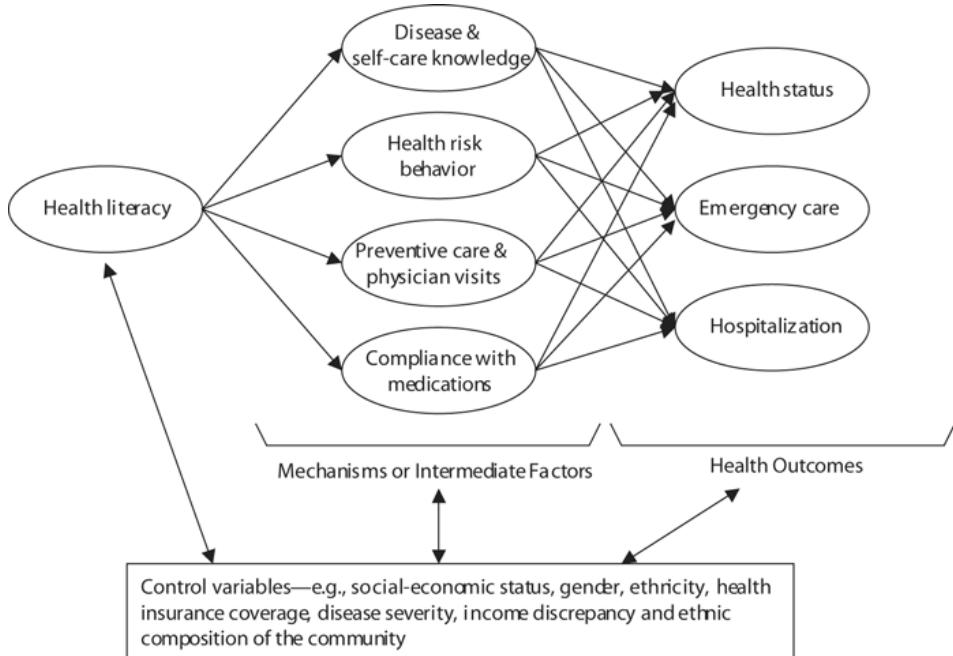
Figure 1 – Factors and contexts that influence health knowledge, decisions and actions.



Source: Kanj, 2009⁽⁸⁾

Low levels of health literacy had been associated with lack of knowledge about diseases or poorer understanding of medical information,^(6, 7, 13-15) poor self-management skills in patients with chronic disease,^(7, 16) poorer physical and psychological health, higher health care costs,⁽¹⁶⁾ less adherence to preventative measures,^(6, 7, 15, 17) patient exclusion from clinical trials and other studies,^(7, 18) more adverse reactions,^(7, 18, 19) and increased mortality.^(16, 20, 21) Some studies showed that individuals with low health literacy had poor knowledge on screenings or symptoms of cancer, and increased risk of later stage of cancer diagnosis,⁽²²⁻²⁷⁾ as well as increased cancer risk and less frequent participation in cancer control programs.^(24, 26, 28-30) Figure 2 summarizes the determinants of health literacy and their influence on health outcomes.

Figure 2 – A simplified model of mechanisms linking health literacy to health outcomes.



Source: Kwan, 2006⁽²⁾

Health literacy is becoming one of the most important factors to empower people in the field of health,^(5, 31, 32) being a key instrument for health promotion.⁽³³⁾ It is expected that the improvement of health literacy contributes for people taking more benefit from their use of the health care system, by enhancing awareness of the rights and duties of health consumers, allowing them to discuss more knowledgeably with care providers the diagnostic and treatment options, and to be involved on political decisions on health-related subjects.⁽⁸⁾ While the discussion about assessing health literacy is very important,^(10, 34) research is needed on how to improve health literacy.

Cancer Literacy

A concept linking literacy to health issues in general, had already been purposed. Nevertheless, there was not a concept for a specific context, like cancer. Assuming that people can have skills in general health, but not in a specific health issues, Diviani and Schulz propose the concept of cancer literacy, defined as “all the knowledge a person needs to possess to understand the information and advice the health system has to offer with regard to preventing, diagnosing and treating cancer.”⁽³⁵⁾ A person who is very literate in general health may not be as well literate in cancer specific conditions.⁽³⁵⁾

It is difficult to select the amount of relevant information related to cancer that should be given to lay people, both from the health professionals' and patients' perspectives.^(35, 36) This decision is frequently made by health professionals, drawing on their patients' information-seeking behaviour.⁽³⁷⁻³⁹⁾ However, some studies showed that patients need more information than the one given by their doctors.^(37, 38, 40-43) While Degner et al found slight agreement among Canadian women with breast cancer regarding the overall profile of information needs,⁽³⁹⁾ Meredith et al. found identical information needs among west Scotland cancer patients.⁽⁴¹⁾

Furthermore, health professionals must take into account that patients' need for information varies according to the phase of the disease trajectory.⁽⁴⁴⁻⁴⁷⁾ In the diagnosis patients prefer information about the stage of disease, treatment options and side-effects of treatments in the diagnosis and treatment stages; after the treatment, they required information about recovery.⁽⁴⁸⁾

It has been shown that information given to patients can increase individuals' knowledge of oral cancer^(49, 50) and skin cancer⁽²⁹⁾ and men's adherence to prostate cancer screening.⁽³⁰⁾ It also reduces the stress caused by decision about treatment options in men with prostate cancer,⁽⁵¹⁾ facilitates coping in women with breast cancer⁽³⁹⁾ and reduces their levels of anxiety or distress.⁽³⁶⁾ For the global improvement of cancer literacy it is important to define a core set of information that must be available for the general population and cancer patients.^(35, 52) For example, in 1995 the British National Health System Information Authority prepared a document where it was suggested the elaboration of a core information package for being disseminate through the health care services among cancer patients,^(45, 52) aiming to equalize the flow, the type and the quality of information provision.

3.2. The Internet as a source of information on cancer

The use of the Internet has become more frequent over the last years. In 2005, 67% of the American used Internet for general issues.⁽⁵³⁾ The use of the Internet as a source of health information has also been increasing.⁽⁵⁴⁻⁵⁸⁾ A recent study, conducted in U.S. adult population (2011), by the Pew Internet and American Life Project, suggests that 80% of Internet users search for health information online and traditional sources of health-related information are being replaced by the Internet.⁽⁵⁹⁾ A study conducted in the United States, by the American Medical Association (2005), showed that the use of the Internet for health issues has been increasing. In March 2002, 62% of the population had already used the Internet to search for health information or medical counseling, while in November 2004, the percentage of individuals that had used Internet to search for health information, which extended to subjects like information on diet, fitness or medication, raised to 80% of the population.⁽⁵³⁾

The most recent data published by Portuguese National Institute of Statistics showed that 48.8% of the Portuguese households had personal computer and 46.0% had access to Internet in 2008.⁽⁶⁰⁾ A study conducted in the Portuguese population in 2005, with a representative sample, constituted by 2001 individuals, showed that 30% of the population used the Internet for health-related issues.⁽⁵⁸⁾ Among individuals who use the Internet on a regular basis, 62% used it regarding health issues.⁽⁵⁸⁾ It also has been demonstrated that among those who use Internet for health related reasons, 14% accessed Internet with this intention at least monthly, which constitutes an increasing of 8.2% when comparing with 2003.^(54, 58) In Portugal we are assisting at an increasing access to the Internet (49% of the population in 2005 to 52.3% in 2007), and also to an increasing use of the Internet for health related reasons (29.2% to 38.3% in the same period).⁽⁵⁴⁾ These results were obtained in a study conducted by WHO/Europe and the Health Consumers Trends Survey, between 2005 and 2007, involving interviews to 3001 individuals, that also presented results that show a similar trend in the countries that participated in the project (Portugal, Denmark, Germany, Greece, Latvia, Norway and Poland).⁽⁵⁴⁾ When asked by the perceived importance of different sources of information, despite being the preferred source of information, health professionals decreased in their importance as a source of health information (nearly 10%), while family and friends gained importance. In the group of regular Internet users, Internet became the most important source of information.⁽⁵⁴⁾

These results are in accordance with another research conducted in 1999 that showed that individuals preferred the information obtained the Internet complemented by that obtained with the health professional, rather than the traditional information given by the health professional alone.⁽⁶¹⁾

The importance of the use of the Internet as a source of information, besides the general aspect of being an easily accessed source of information, is associated with the fact that it empowers people and allows them assume more responsibility for their health.^(54, 58)

When looking for general information on the Internet, people can use different tools like search engines, subject directories, searchable databases, social networking, social bookmarking or real-time search.⁽⁶²⁾ However, most individuals choose the search engines,⁽⁶³⁾ where one or more words – the search expression – are written and the resulting web pages are displayed (for example, websites, images or information). Drawing on users' search expression, the engine analyses its index and provides a list of best-matching web pages according to its criteria, usually with a short summary of the content. In 2009, the most important search engine of the U.S. was the Google[®], with indexation of trillions of web pages.^{(64) (64)} In August 2010, 71.6% of the searches conducted on the Internet in the U.S. through search engines were conducted with Google[®].⁽⁶⁵⁾

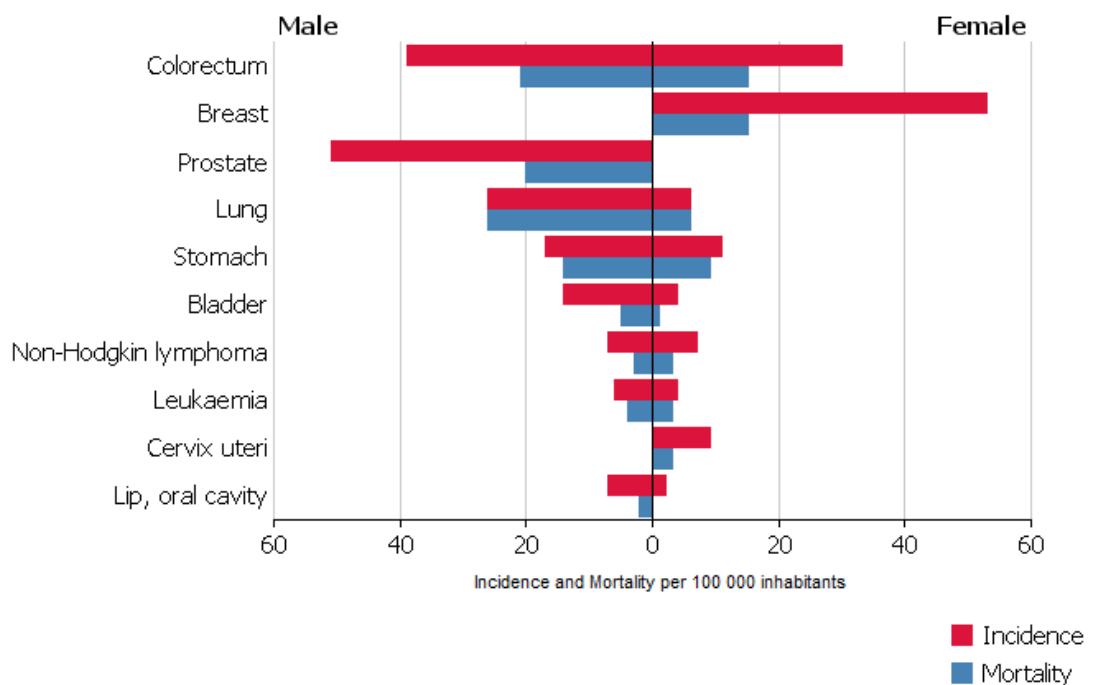
The order by which the Google[®] search engine provides the search results is based on an algorithm that takes into account the number of links and the page ranks⁽⁶⁶⁾ of the linking sites.⁽⁶⁷⁾ The page rank is not a measure of the quality of the website, but a measure of the website's popularity.⁽⁶⁸⁾

Cancer patients use different sources when searching for cancer information, such as health professionals, television, radio, books, brochures, pamphlets, newspapers, magazines or the Internet.^(36, 37, 39, 69-71) Physicians and other health professionals are recognized by patients as the best source of information on cancer.^(70, 71) The Internet is referred by cancer patients as less reliable than the experts, but used in a more regular basis.^(70, 72) Health care providers should be aware of it, because the quality of the information collected on the Internet has been described as poor,⁽⁷³⁾ and may have negative repercussions on preventive strategies, diagnosis and treatment adherence.

Breast and prostate cancers are major causes of oncological morbidity.^(74, 75) In Portugal, estimates for 2008 showed that breast cancer is the most incident (5333 new cases) and the second cause of oncological death (1537 deaths) among women,⁽⁷⁵⁾ while in men prostate cancer is also the most frequent (5140 cases) and the third cause of cancer death (2021 deaths)⁽⁷⁴⁾ (figure 3). Breast and prostate cancers have relative survivals that are among the highest.⁽⁷⁶⁾

Breast and prostate cancer patients are among the cancer patients that use internet more frequently for searching information about their disease⁽⁷¹⁾ (23.8% and 11.5%, respectively).⁽⁷⁷⁾

Figure 3 – Incidence and mortality by cancer (per 100 000 inhabitants), in Portuguese women and men, all ages, in 2008.



SOURCE: GLOBOCAN, 2008^(74, 75)

Both female breast and prostate cancers share the potential for control through secondary prevention. Breast cancer screening has been recommended while prostate cancer screening, besides being possible, remains controversial.^(78, 79) Therefore, contents on screening should be part of the information available on cancer subjects.⁽³⁵⁾

3.2.1. Quality of websites providing health information on the Internet

To avoid possible negative repercussions of misleading health information available on the Internet, some official medical entities established criteria,^(80, 81) guidelines,⁽⁸²⁾ principles or codes of conduct, for assuring its quality. Although these instruments assess elements of structural quality, they cannot evaluate the accuracy and updating of the contents. Taking into account this limitation of existing formal assessments, it is important to evaluate the quality of the health information accessible through the Internet, regardless websites' structural aspects.^(83, 84)

3.2.1.1. Formal assessment

We present a summarized description of four different instruments, which can be used to assess the formal quality of the websites.

The Discern tool

The Discern tool was the first standardized index for evaluating the quality of written health information on treatment choices, developed by investigators of the Imperial College School of Medicine (London) and University of Oxford (Oxford). It includes 15 quality criteria, all of them defined as “an essential feature or standard that is an important part of good quality information on treatment choices”^(80, 81). Each criterion is rated on a 5 point Likert scale, ranging from a “definite NO” (n=1) to a “definite YES” (n=5), if it is completely fulfilled.⁽⁸⁰⁾ This instrument has “Hints”, guiding the users when rating the criterion,⁽⁸⁰⁾ and it is divided into three sections: questions 1 to 8 assess the reliability of the information; questions 9 to 15 focus on specific details of the information on treatment choices; a final question (16th) is an intuitive appreciation of the overall trustworthiness of the website, and it is not accounted to compute the overall score. This final question could be rated as “low” (1), if the majority of the questions scored two or one; “moderate” (3), if the majority of the questions scored three or the information evaluated scored high or low values in a similar number of criteria; or “high” (5) if the majority of the questions scored four or five.⁽⁸⁰⁾

The Discern tool, however, cannot assess if the contents provided are in accordance with the evidence available on any specific subject. Prior knowledge on the topics addressed is not a requirement for using this tool; therefore, someone who searches for web information on different health topics is able to use it. Information providers may use the Discern tool as a checklist for relevant formal aspects of written health information on treatment choices, or

as a training tool for health professionals to improve communication and shared decision-making skills.⁽⁸¹⁾

The Health On Net Foundation code

The Health On Net Foundation (HON) is a non-profit and non-governmental organization, accredited by the Economic and Social Council of the United Nations. The HON was created in 1996 and is supported by the Geneva Economic and Health Department, the EU and the Genève University Hospitals. It aims “to encourage the dissemination of quality health information for patients and professionals and the general public, and to facilitate access to the latest and most relevant medical data through the use of the Internet”.⁽⁸⁵⁾

This organization proposes 8 principles for assessing the quality of all websites providing health information (table 1). The HON code was the first instrument to assess transparency in this field.

It provides a benchmark for web publishers, and websites can request the authorization to exhibit the “HON seal”. The HON Foundation team conducts the evaluation of the website according to the HON Code, but the extent to which each website must be in accordance with the principles it is not known.⁽⁸⁶⁻⁸⁸⁾ If a submitted website is in accordance with the HON code, it is assigned a seal (which can be displayed in the websites), that cannot be interpreted as a certification of the accuracy of its contents.

Table 1 – Principles of the HON code.

Principle	Description
1 – Authority	Indication of the qualification of the authors
2 – Complementarity	Indication that the website provides supportive information that should not replace the doctor-patient relationship
3 – Confidentiality	Respect for the privacy and confidentiality of personal data submitted to the website by the visitor
4 – Attribution	Citation of the source(s) of published information
5 – Justifiability	Claiming of benefits or performances of some specific medical treatment, commercial product or service, supported by clear references to scientific research results and/or published articles
6 – Transparency	Accessible presentation and accurate email contact
7 – Financial disclosure	Identification of the funding sources
8 – Advertising policy	Clear distinction of advertising from editorial content

SOURCE: The HON Foundation, 2011⁽⁸⁹⁾

The eEurope guidelines for health related websites

Due to the recognition that searching for health information on the Internet by European citizens was increasing, as well as the use of rating tools for assessing the reliability of that information, the European Council, in 2000, decided to develop a core set of quality criteria for health related websites (table 2).⁽⁹⁰⁾ The development of those quality criteria did not intend to establish a mandatory set of rules for all health related websites, but to “draw up a commonly agreed set of simple quality criteria on which Member States, as well as public and private bodies, may draw in the development of quality initiatives for health related websites. These criteria should be applied in addition to relevant Community law.”⁽⁹⁰⁾

Overall, the proposed quality criteria take into account the following issues:

- A document should tell suppliers how to comply with key criteria, but also educate users for what they ought to expect from a good health website;
- Both passive information-giving websites and more interactive websites that allow for transactions between service or information providers and users (e.g. information, products and services) should be included;
- The quality criteria should facilitate compliance with EU Directives, other current guidelines and relevant technical standards.

Table 2 – eEurope guidelines for health related websites.

Guideline	Description
1 – Transparency and honesty	Transparency of the provider of the site Transparency of purpose and objective of the site Clear definition of the target audience Transparency of all sources of funding for site
2 – Authority	Clear statement of sources for all information provided and date of publication of the source Clear information of the names and credentials of all human/institutional providers of information
3 – Privacy and data protection	Clear definition of the privacy and data protection policy and system for the process of personal data, including process invisible to users, in accordance with community Data Protection legislation
4 – Updating and information	Clear and regular updating of the site, with date of up-date clearly displayed for each page and/or item as relevant. Regular checking of relevance of information
5 – Accountability	Possibility of the user feedback, and appropriate oversight responsibility (ensuring that partnering or linking to other websites is undertaken only with trustworthy individuals and organizations, who themselves comply with relevant codes of good practice)
6 – Editorial policy	Clear statement describing what procedure was used for selection of content
7 – Accessibility	Attention to guidelines on physical accessibility as well as general findability, searchability, readability, usability, etc

SOURCE: eEurope 2002, 2000⁽⁸²⁾

American Medical Association guidelines for medical and health information sites

The American Medical Association (AMA) guidelines are primarily directed to all the websites associated with AMA,⁽⁹¹⁾ but they also intend to produce some guidance to any medical website, being a support for websites creators.

The development of these guidelines, which started in 1999, was based on other previous guidelines, and intended to set up the criteria in four major areas, which were identified as needing for quality standards. The AMA guidelines evaluate issues of content, advertising and sponsorship, privacy and confidentiality and e-commerce. The AMA guidelines are organized in four principles, each one concerning different aspects of the website (table 3).

Table 3 – AMA principles for medical and health information sites.

Principle	Description
1 – Content	High quality medical and health information must be provided by the site. Information about navigational directions and tools needed to judge the quality, reliability, objectivity, sources, and funding of content should be given.
2 – Advertising and Sponsorship	Advertising can be present, but concerns must be taken on ethical issues and on the type of products that can be advertised on the sites.
3 – Privacy and Confidentiality	Protection of the right of the individual to choose whether to allow personal information to be collected by the host website or by third parties, as well as to know the kind of information collected and for what purposes that information would be used.
4 – E-Commerce	Assurance that partners and purchaser of information, products and services on the site, will have access to secure, efficient transactions for online and remote customer fulfillment.

SOURCE: Winker *et al*, 2000⁽⁹²⁾

3.2.1.2. Analysis of the contents

Health-related contents disseminated in the Internet should be accurate and update in a regular basis,⁽⁹³⁾ and there are different guidelines for the assessment of its formal quality. However, there are no instruments specially devoted to the assessment of the accuracy of the contents.^(94, 95) Such an evaluation needs to be conducted case by case, taking into account the best available evidence on each health topic and the local health policies.⁽⁹⁶⁾

Several attempts were made to construct an instrument that could be easily used by individuals to assess the quality of the contents of a website. The University of Edinburgh consider the audience, authority, accuracy, currency, characteristics of the URL and reliability of the sources as criteria to assess the quality of the contents provided by a website. Others, like the University at Albany (State University of New York), considered as essential information about the expertise of the author, the correct source of the information, the independence from other sources, the level and depth of information provided, the presentation of the information, the updating of the website in general and of the contents, the maintenance of the functional links, the information about websites' contact and the characteristics of the URL.⁽⁹⁷⁾ The Midland Technical College considers as quality criteria for the contents of websites aspects such as accuracy, authority, currency, objectivity, purpose and scope, intended audience, format and appearance, and privacy.⁽⁹⁸⁾ These examples of instruments or criteria; demonstrate that some aspects like the accuracy, intended audience, authority, currency or characteristics of the URL, tend to be consensual. Although, other aspects seem to be particular of some instruments, which demonstrate that there is no gold standard to assess the quality of the contents provided by websites.

The accuracy of information, considered as important to judge the quality of the contents provided by a website,⁽⁹⁹⁾ may be related with the actual correctness of the information that is required to be up-to-date, factual, detailed, exact and comprehensive.⁽¹⁰⁰⁾ Another dimension considered in the evaluation of the quality of the website was the credibility, referred as a "question of trust."⁽¹⁰⁰⁾ The credibility of the information, defined as "...judgments made by a perceiver concerning the believability of a communicator"⁽¹⁰¹⁾ may be confirmed in different ways. In a qualitative study using focus groups, naturalistic observation of consumers searching the world wide web in a usability laboratory, and in-depth interviews, conducted in Heidelberg, Germany (2002) results showed that the quality of the information was assured if the website was affiliated to official authorities; if it was understandable for people in general and written for a professional related to the subject; include citations of scientific journals; or if it shows the photograph of the author.⁽¹⁰²⁾

The quality of contents available on the Internet, related with health issues, has been considered poor.^(66, 83, 103-107) Regulation of the websites and education by information providers have been considered to improve the general quality of the websites.^(95, 107)

3.2.1.2.1. Readability

The readability refers to the facility with which a text is read,⁽¹⁰⁸⁾ being an important aspect of a website's quality of the contents.⁽¹⁰⁹⁾ The information provided on a website is available to people with different characteristics⁽¹¹⁰⁻¹¹²⁾ and therefore the texts must be written taking into account the expected literacy level of the intended audience.

Readability can be measured by instruments specifically designed for that purpose, which are based in characteristics such as sentence length and words list or sentence length and word length.⁽¹¹³⁾ The underlying basis of the readability is that longer words and longer sentences difficult the reading of a written text.

Most of these instruments aim to evaluate the readability for texts written in English, such as Flesch Easy Reading Readability,⁽¹¹³⁻¹¹⁶⁾ Flesch-Kincaid,^(117, 118) SMOG Grading,^(119, 120) FOG,⁽¹⁰⁹⁾ and Automated Readability Index.⁽¹⁰⁹⁾ These instruments cannot be directly applied to Portuguese texts because the measure of readability is based on structural aspects of the texts that vary according to the language.

In average, the length of English words is higher than that of Portuguese words. English words have an average of two syllables, while the Portuguese have 2.5 syllables. The application of the existing instruments to Portuguese written texts would always yield a score reflecting a poorer readability than the one that actually occurs.⁽¹²¹⁾

One of the most used formulas for assessing readability is the Easy Flesch Reading Readability, which is based on the average sentence length in words and in syllables (table 4).

Table 4 – Flesch Reading Ease formula to measure readability and Index Fernandez-Huerta.

Instruments	Formulas
Flesch Reading Ease	$206.835 - 84.6 \text{ (average number of syllables per word)} - 1.015 \text{ (average number of words per sentence)}$
Index Fernandez-Huerta	$206.84 - 0.6 \text{ (average number of syllables per word)} - 1.02 \text{ (average number of words per sentence)}$

SOURCE: Barboza, et al.⁽¹⁰⁹⁾

The application of the formula results on a score, ranging from 0 to 100, representing the worst level (very difficult to read) and the better level of readability, respectively.

The scores obtained with the application of the formula, are grouped in seven categories, which reflect the school level that an individual must have to understand the text (table 5).^(113, 115)

Some attempts were made to validate the Easy Flesch Reading Readability for texts written in Portuguese,^(109, 115) but there is no specific instrument to be used with the Portuguese language yet. The Index Fernandez-Huerta was created to measure the readability of texts written in Spanish, based on the Easy Flesch Reading Readability formula. The Spanish language share with the Portuguese the same Latin basis, so it is the best formula to be applied in Portuguese texts, at the moment.⁽¹⁰⁹⁾ In this case, the score obtained cannot be interpreted in terms of the level of schooling necessary for the text to be readable, as possible with the Easy Flesch Reading Readability formula,⁽¹¹³⁾ since the Portuguese school curriculum is different from the American one. We know that the original formula ranges from 0 to 100, where 0 is the extreme of most difficulty on reading and 100 represents the extreme of easiest reading. Based on this aspect, a readability level about 70 was expected to correspond to a text readable by most people in Spanish language.⁽¹⁰⁹⁾

Taking this into account, and based in a previous work in Portuguese language, where it was measured the readability level of governmental Brazilian websites,⁽¹⁰⁹⁾ we assumed that this would be also a good value for readability in Portuguese language.

Table 5 – Flesch Reading Ease score and school level needed do understand texts.

Flesch score	School level
0-30	college graduate
30-50	college
50-60	10 th to 12 th grade (high school)
60-70	8 th and 9 th grade
70-80	7 th grade
80-90	6 th grade
90-100	5 th grade

4. Aims

The easy access, fast dissemination of information and availability of different sources make the Internet a privileged vehicle to search and to provide information.^(73, 122) Its use has increased in the past few years in western societies, and currently it is one of the most frequently used sources of both general and health related information.^(64, 123, 124)

The overall quality of the health websites and the accuracy of their contents must be assured, since Internet, a widely used resource,^(66, 122) has the potential to shape health behaviors and knowledge of a large number of people worldwide.^(83, 105, 125)

Information related to oncological diseases correspond to an important proportion of the Internet searches on health issues.⁽⁵⁹⁾ Breast and prostate cancers are leading causes of oncological morbidity,^(74, 75) and are among the cancers with the highest relative survival,⁽⁷⁶⁾ which contributes to a high interest in these subjects, by population in general, and patients and their families in particular. Breast and prostate cancer patients are the ones that use internet more frequently to search for information related to their disease.⁽⁷¹⁾

Furthermore, these cancers have specificities regarding the potential for control through secondary prevention, and the extent to which organized and opportunistic screening is conducted in different settings. Thus, these cancers may constitute a good model for designing a framework of website quality assessment that may be extended to other conditions for which screening is recommended or effectively conducted, regardless of the available evidence on its effectiveness.

This dissertation aimed to evaluate the quality of websites providing information on female breast and prostate cancers in Portuguese language, through the accomplishment of the following specific objectives:

- To describe the formal characteristics of the websites, according to the principles defined by the “Discern tool”, “Health On Net code”, the “eEurope guidelines for health related websites” and the “American Medical Association guidelines for medical and health information sites”.
- To evaluate the contents of the websites on breast or prostate cancer screening, its agreement with the best available evidence on this topic, and its readability score.

5. Methods

5.1. Selection of the websites for analysis

We searched the World Wide Web to identify web pages in Portuguese language that addressed prostate or breast cancers, in the 15th and September 16th 2011, respectively, using the Google® search engine, with the expressions *cancro da prostata* and *cancro da mama*, respectively.

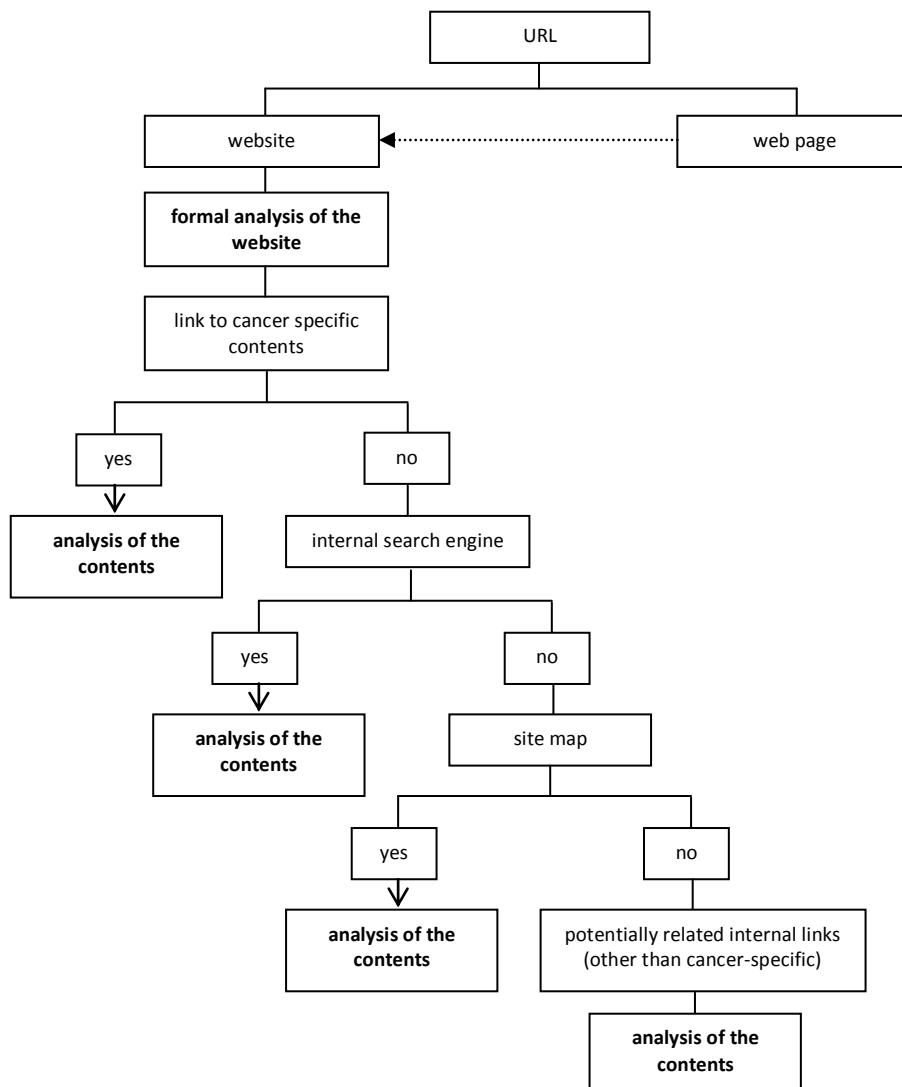
We saved the 200 first results from each search for further analysis, including information on the URL (Uniform Resource Locator) of each web page, and registered its rank in the search. The websites were initially screened to assess eligibility, by applying the following exclusion criteria:

- inaccessible websites due to non-functioning URL;
- websites not providing information in Portuguese language;
- repeated websites (corresponding to different web pages from the same website);
- websites providing information on breast or prostate cancer only in the format of downloadable files (e.g. slideshows, portable document files), or only through audio or videos (e.g. YouTube videos);
- scientific articles (whether or not located in medical websites);
- blogs or forums;
- general encyclopedias;
- websites only providing information about female breast or prostate cancers in the form of news;
- websites with no specific information on female breast or prostate cancers (e.g. advertising only, male breast cancer).

After the identification of the eligible websites, when the URL corresponded to a website's main page, it was the departure point for the identification of breast or prostate cancer related information searching the whole website. Otherwise, when the URL corresponded to a web page other than the website main page, we navigated towards the latter, and then a more comprehensive screening of the website was conducted for identification of all relevant pages (providing information on breast or prostate cancers), as depicted in figure 4.

The general characterization of the website was accomplished in any of its pages. The analysis of the contents on cancer screening was conducted in the websites that specifically addressed this topic, regardless of the extent of the available information.

Figure 4 – Navigation process for analysis of the website.

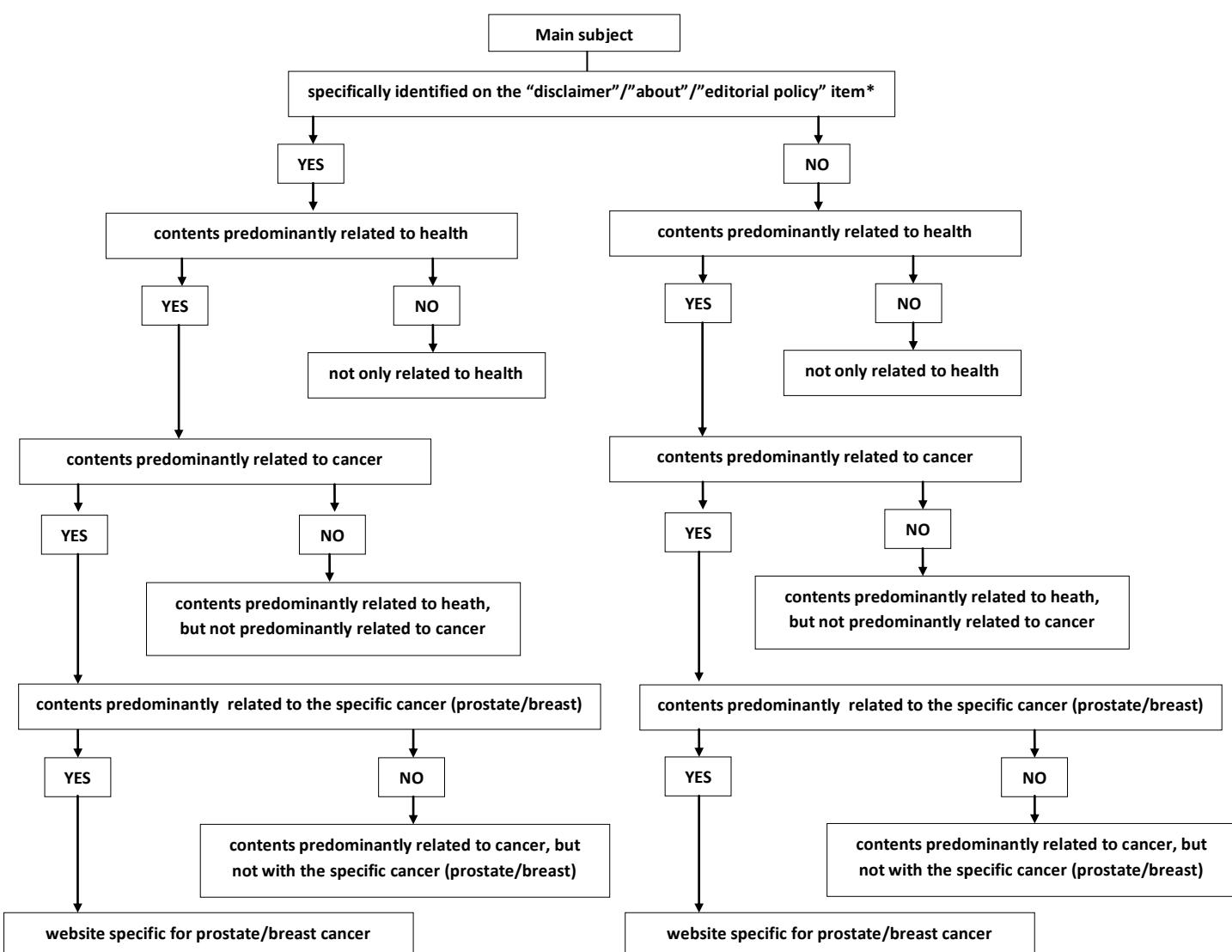


5.2. General characterization of the websites

The general characterization of the websites was accomplished through the collection of information on the following variables: main subject, country of origin, intended audience, affiliation and media used to convey the cancer-specific information.

The main subject corresponds to the subject matter predominantly addressed by its contents. The websites were classified regarding the predominant relation of its contents to health (health related/not only health related); to cancer (cancer related/not only cancer related) and to breast or prostate specific disease (breast or prostate cancer specific/not specific for breast or prostate cancer), as depicted in figure 5.

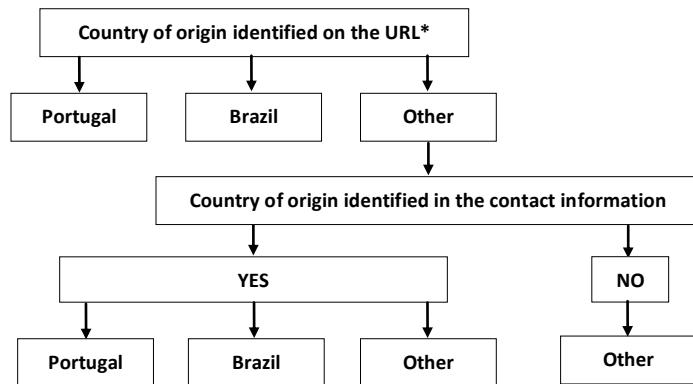
Figure 5 – Standardized procedure followed to classify the websites according to the main subject.



* In the websites that did not present these items, we searched other fields for this information.

The websites could have their origin in Portugal, Brazil, African countries where the Portuguese is the official language, or in other non Portuguese-speaking countries. The classification was achieved as described in figure 6.

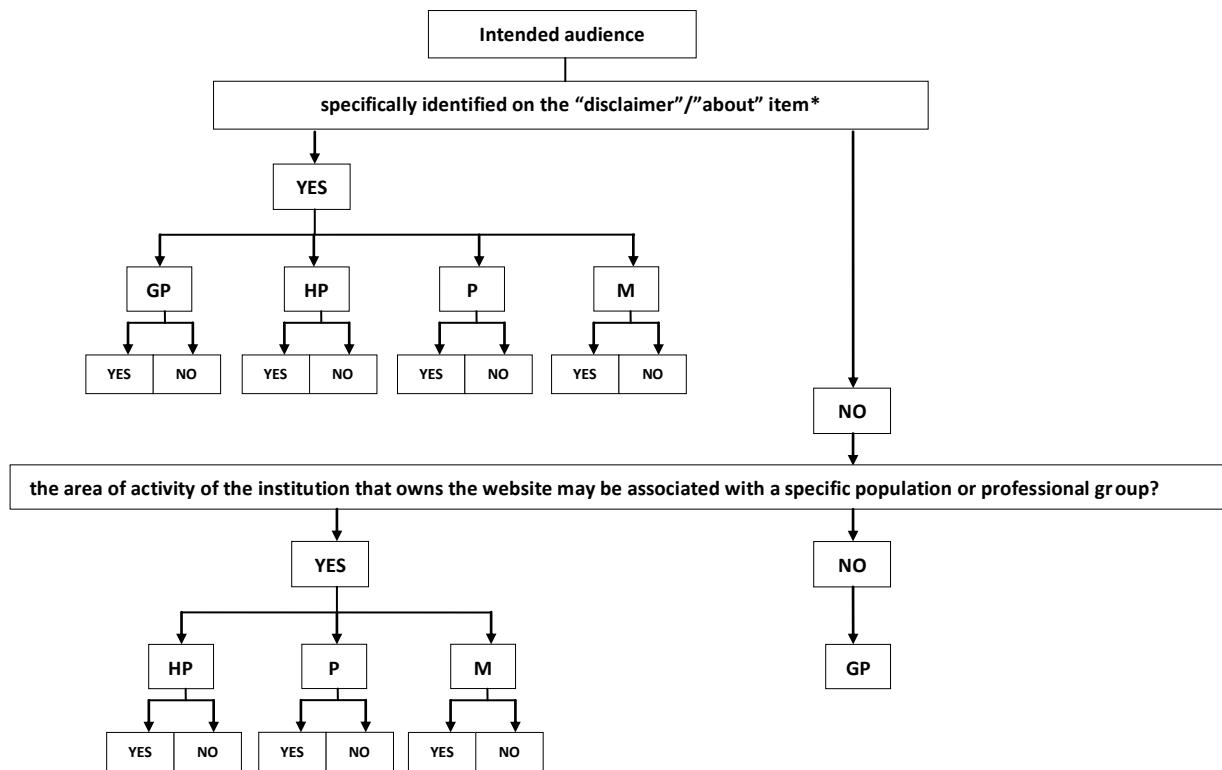
Figure 6 – Standardized procedure followed to classify the websites according to the country of origin.



* The URL allows the determination of the origin of the website. (“.pt” if Portuguese, “.br” if Brazilian, or other).

The intended audience was classified as general population (GP), patients (P), health professionals (HP) or media (M). Since the websites could target more than one population group, these categories were not mutually exclusive. Figure 7 presents the standardized procedures followed to classify the website according to its intended audience.

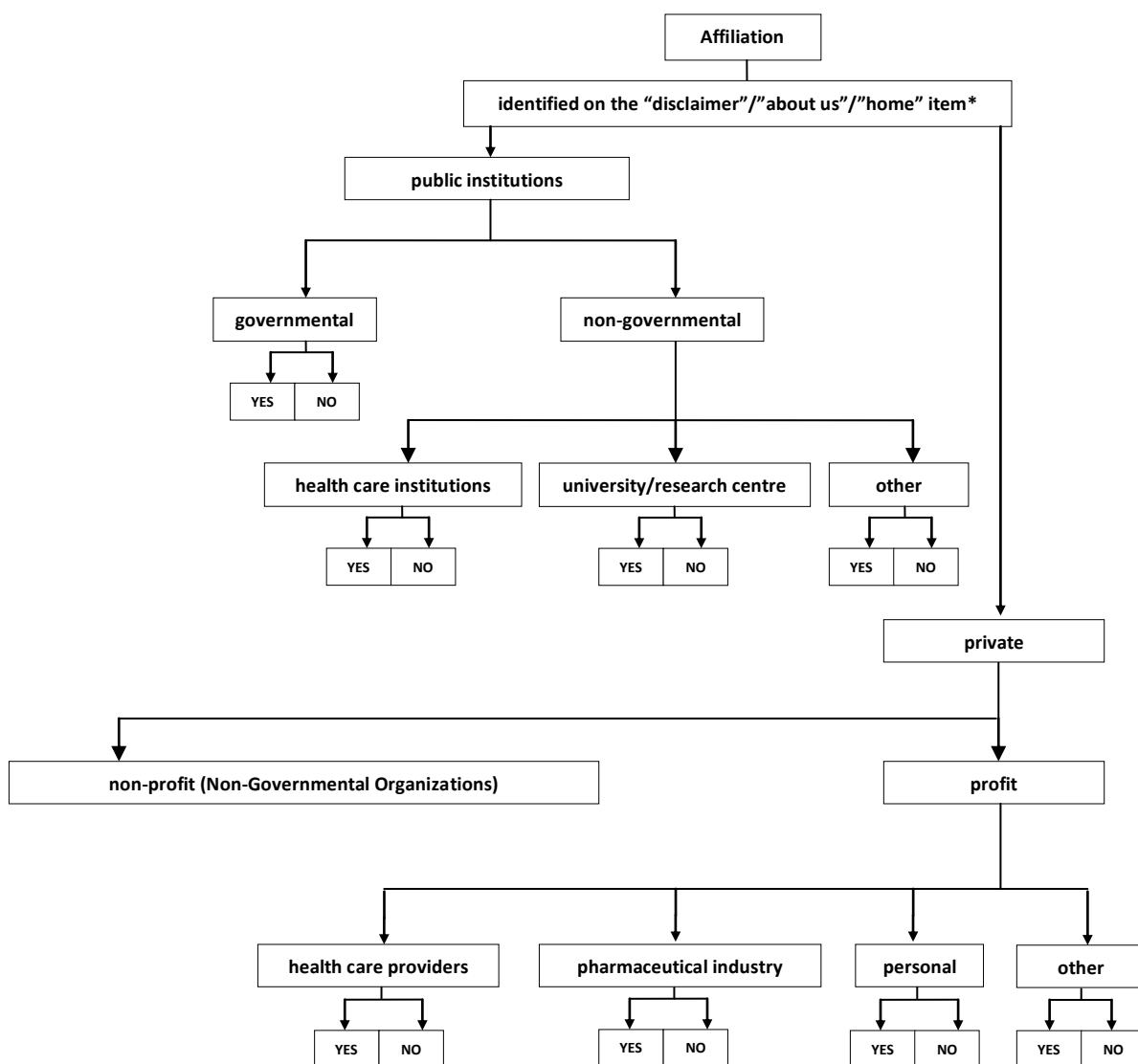
Figure 7 – Standardized procedure followed to identify the websites according to the intended audience.



* In the websites that did not present these items, we searched other fields for this information.

The affiliation of the website was primarily defined as public or private. The websites of public institutions were further classified as governmental or non-governmental. The latter institutions could be health care providers, universities, research centers or other (not further specified). Among the private institutions, we distinguished the organizations responsible for the websites based on the profit intent, and grouped them as profit or non-profit. The latter correspond to Non-Governmental Organizations (NGO). The former were further classified as health care providers, pharmaceutical industries, individual subject or other. The standardized procedures for this classification are described in figure 8.

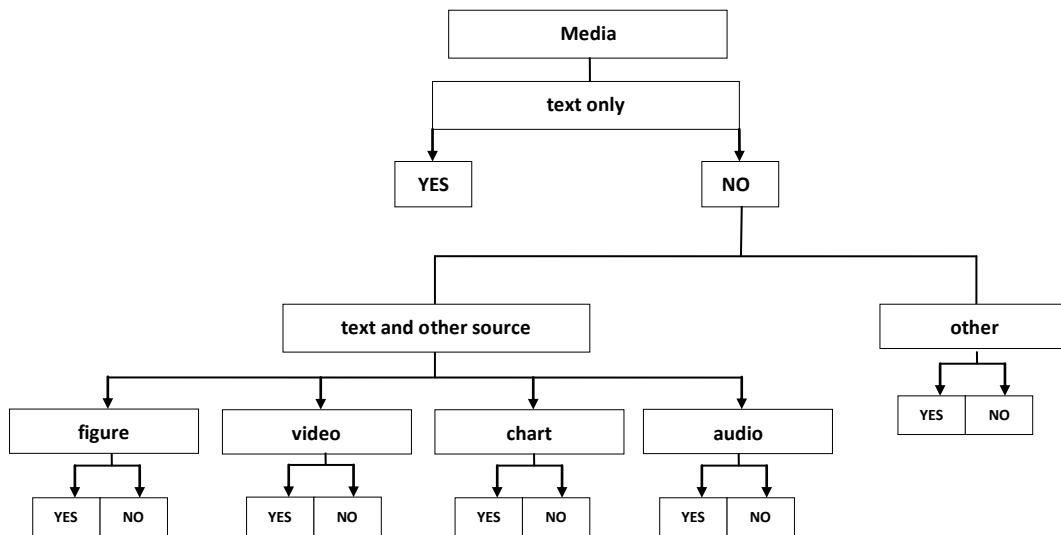
Figure 8 – Standardized procedure followed to identify the websites according to the affiliation.



* In the websites that did not present these items, we searched other fields for this information.

The media used by the websites to convey the cancer-specific information (display of contents) were classified in five mutually exclusive categories: text only; text and figures; text and video; text and charts; text and audio (figure 9).

Figure 9 – Standardized procedure followed to identify the media used by the website.



The results, regarding the general assessment of the characteristics of the websites, are presented as the proportion of websites depicting each of the above characteristics, for the whole sample and according to the cancer addressed in the website (breast vs. prostate cancers) and the ranking of the websites in each of the searches (first 30 URL vs. remaining results). The latter cutoff was selected because individuals who search on the Internet tend to navigate until the third page of results.⁽³⁷⁾ The proportions were compared with the chi² or the Fisher exact test, as appropriate.

5.3. Assessment of the formal quality of the websites according to the Discern tool, the principles of the Health On Net code, the eEurope guidelines for health related websites and the American Medical Association guidelines for medical and health information websites

The formal characteristics of the websites were evaluated using the Discern tool and taking into account guidelines and principles issued by three different organizations, namely the Health On Net Code, the eEurope guidelines for health-related websites, and the American Medical Association guidelines for medical and health information sites.

The Discern tool

This instrument may be applied to assess the quality of written information about treatments^(80, 81, 106), and therefore only the websites addressing treatments breast or prostate cancer were evaluated.

The Discern tool has specific instructions for its use that clarifies the intended purpose of the questions and provides examples to support the rating of each question on a 5-point scale ranging from no to yes, and the instrument was used as suggested by the handbook instructions.⁽⁸⁰⁾

The results are presented as the median (percentile 25-percentile 75) score obtained in each of the items of this tool. The Kruskal-Wallis test was used to compare the distribution of the scores assigned to the websites addressing breast or prostate cancer.

The Health On Net Foundation code

The Health On Net Foundation (HON) defined a set of quality criteria (HON code of conduct), based in eight principles, and may assign a certificate, the HON Seal, to the websites that comply with the principles.⁽⁸⁹⁾ These tools may contribute used by the citizens to perceive the quality of websites. However, the rules for the verification of the accomplishment of the criteria listed in the code of conduct are not provided and it is also unknown the extent to which the websites need to comply with the principles to be allowed to exhibit the HON Seal. Therefore, we defined a set of criteria to classify the websites regarding the conformity with each of these items.

Each quality criterion was coded as 0 when it was not accomplished or 1 when it was fulfilled.

When more than two options were available to characterize a website according to a criterion, the code 0 or 1 could be applied to more than one option corresponding to non accomplishment or accomplishment of the criterion, respectively.

The criteria for which different levels of agreement were possible were coded with consecutive integers, from 0 (not fulfilled) to the highest degree of conformity possible (e.g. 2 or 3), in an ordinal scale. The code 9 was used to identify the websites in which the evaluation was not possible or not applicable (table 6).

For descriptive analysis we combined the results of the evaluation of the criteria corresponding to the same principle, as follows:

- when the principle corresponded to a single criterion, the same code was used for both;
- when the same principle included more than one criterion, the final score for the principle was 0 (no compliance), when none of the criteria was accomplished; 2 (total compliance), when all the criteria had the maximum possible score; or 1 (partial compliance), otherwise.

The results are presented as the proportion of websites with different degrees of compliance with each of the HON principles (non-compliance, partial compliance, total compliance). The proportions were compared with the chi² or the Fisher exact test, as appropriate.

Table 6 – HON code of conduct – principles and quality criteria.

Criteria	Accomplishment of the criteria		Code
<u>Principle 1 (p1) – Authority</u>			
The website provides general information about the organization or individual responsible for its operation and content, and a person is named as editor or principal author.	HON_1	Yes, an author's name is given	1
		No	0
<u>Principle 2 (p2) – Complementarity</u>			
A statement declaring that information provided on the website is meant to complement and not replace any advice or information from a health professional is clearly provided.	HON_3	Given by medical/health professionals whose training/credentials are listed	2
		Clear statement (e.g. a disclaimer) is made whenever medical/health information or advice is offered by non-medical professionals or organizations	1
A statement describing the intended		Some health/medical information is not attributed to an author	0
	HON_4	Yes	1
			0

mission of the website is provided.		No	0
The website clearly mentions the intended (general public, health professionals, students...).	HON_5	Yes	1
		No	0
Principle 3 (p3) – Confidentiality			
Privacy/Confidentiality policy regarding e-mail addresses, personal and medical information is displayed on the website.	HON_6	Yes	1
		No	0
Do the website and its mirrors respect the legal requirements, including those concerning medical and personal information privacy that apply in the country and state of their location?	HON_7	Don't know	2
		Yes	1
		No	0
Principle 4 (p4) – Attribution			
Is the last modification date provided for the website?	HON_8	Yes, for all the web pages of the website	3
		Yes, for each web page having health/medical content	2
		Yes, for the website as a whole	1
		No (explain as necessary below)	0
Does the website contain information from external sources?	HON_9	Yes, an HTML link (valid and regularly checked) is provided to the source data	3
		Yes, a bibliographic reference to the source data is given	2
		No, the content of my website is original, written by the editorial website team	1
		Yes, but no references to the source is made	0
Principle 5 (p5) – Justifiability			
Does the website make claims relating to the benefit or performance of a specific medical treatment, commercial product or service?	HON_10	Yes, all claims are supported by clear references to scientific research results and/or published articles	1
		Yes, my claims are based on my personal research or opinions	0
		No	1
Principle 6 (p6) – Transparency			
A valid email address for the webmaster or a link to a valid contact form is easily accessible throughout the website?	HON_11	Yes	1
		No	0
Principle 7 (p7) – Financial disclosure			
Is the source of the funding of the website clearly described? for commercial or non-commercial organizations:	HON_12a	Yes	1
		No	0
		Not applicable	9
Is the source of the funding of the website clearly described? for personal or private websites, or those hosted without charge	HON_12b	Yes	1
		No	0
		Not applicable	9
Principle 8 (p8) – Advertising policy			
The website displays advertising that	HON_13	A web page provides a description of our advertising policy	1

is a source of income		Separation between editorial content and advertising is clearly stated	1
		No explanation regarding banner advertising is given	0
		All ads banners are clearly identified as advertising with the word 'advertising'	1
		Advertising is not identified as such	0
		Not applicable	9
The website is part of a link/banner exchange	HON_14	Yes, a statement describes precisely the relationship between my website and the other websites	3
		Yes, a statement describes precisely the relationship between my website and the other websites mentioning any economic benefit derived from the exchanges	2
		Yes, but there is no specific description about our policy	1
		No	0
		Not applicable	9
The website does not display advertising	HON_15	There is a clear statement explaining that my website does not accept or host any advertisement	1
		There is no statement displayed	0
		Not applicable	9

eEurope guidelines quality criteria for health related websites

The eEurope guidelines for health related websites are recommendations of the European Council that may be followed by those who are responsible or want to create a website on health information.^(82, 90) To assess the quality of the websites according to these guidelines we defined a set of criteria that are presented in table 7. All criteria were coded as 0 or 1, when they were fulfilled or not fulfilled, respectively.

Since according to the instrument, the first four criteria were related with a single principle (p1 – “transparency and honesty”), for the descriptive analysis we presented the results for the principle, that was coded as follows: 0 (no compliance), when none of the criteria was accomplished; 2 (total compliance), when none of the criteria was accomplished; or 1 (partial compliance), otherwise.

The item related to the principle “authority” was divided in three sub-items referring to different requirements, as follows: name of the persons of the staff (6.1); the credentials of the persons of the staff (6.2); the date at which credential were received (6.3). Then we used the credentials of the staff item as reference, because it was the most complete, once only could be accomplished if the names of the staff were also displayed, and because we

thought that it would be very rare a website presents the date at which credentials were received.

One of the criteria was destined to assess the accessibility of the website, regarding physical accessibility, findability, searchability, readability and usability. This criteria is complex to evaluate and there are instruments specially created to assess this particularities that go beyond health related websites and extends to all websites in general. Nevertheless, as we considered the readability an important factor in the quality of the website we used the Index Fernandez Huerta to measure the readability. The latter, however, was evaluated in the analysis of the contents and not with the application of these guidelines.

We did not evaluate the item related to the responsible partnering of the websites, because we could not define the concept of “trustworthy individuals”, as well as the concept of “good practice”. It was impossible to clarify the concepts in a way that could allow us to evaluate those on the websites.

The results are presented as the proportion of websites with different degrees of compliance with each of the eEu criteria (non-compliance, partial compliance, total compliance). The proportions were compared with the χ^2 or the Fisher exact test, as appropriate.

Table 7 – eEurope guidelines and quality criteria.

Criteria		Accomplishment of the criteria	Code
Guideline (p1) - Transparency and honesty			
Transparency of the provider of the website – including name, physical address and electronic address of the person or organization responsible for the website	eEu_1	Yes No	1 0
Transparency of purpose and objective of the website	eEu_2	Yes No	1 0
Target audience clearly defined	eEu_3	Yes No	1 0
Transparency of all sources of funding for website (grants, sponsors, advertisers, non-profit, voluntary assistance)	eEu_4	Yes No	1 0
Guideline - Authority – source and date of information			
Clear statement of sources for all information provided and date of publication of the source	eEu_5	Yes No	1 0
Guideline - Authority – name and credentials of information providers			
Name of the human/institutional providers of information up on the website	eEu_6.1	Yes No	1 0
Credentials of the human/institutional providers of information up on the website	eEu_6.2	Yes No	1 0
Dates at which credentials were received	eEu_6.3	Yes No	1 0
Guideline - Privacy and protection data			
Privacy and data protection policy and system for the processing of personal data, including processing invisible to users, to be clearly defined in accordance with community Data Protection legislation	eEu_7	Yes No	1 0
Guideline - Updating and information			
Clear and regular updating of the website, with date of up-date clearly displayed for each web page and/or item as relevant. Regular checking of relevance of information	eEu_8	Yes No	1 0
Guideline - Accountability			
User feedback, and appropriate oversight responsibility (such as a named quality compliance office of each website)	eEu_9	Yes No	1 0
Guideline - Editorial policy			
Guideline - Clear statement describing what procedure was used for selection of content	eEu_11	Yes No	1 0

American Medical Association guidelines for medical and health information sites

The American Medical Association established principles that should reflect its policy concerning websites.^(92, 114, 126-128) In order to turn possible the application of these principles, we created a set of criteria based on those suggested by the AMA. We attributed the code 0 if the criterion was not accomplished or the code 1 when the criterion was accomplished. The code 9 was used when the criterion was not applicable, or could not be evaluated (table8).

For descriptive analysis we combined the results of the evaluation of the criteria corresponding to the same principle, as follows:

- when the guideline corresponded to a single criterion, the same code was used for both;
- when the same guideline included more than one criterion , the final score for the guideline was 0 (no compliance), when none of the criteria was accomplished, 2 (total compliance), when all the criteria had the maximum possible score, or 1 (partial compliance), otherwise.

The AMA principles governing websites have some principles that could not be assessed, namely principles related to the sponsorship, privacy policy or e-commerce. These were not evaluated because their utility is related to webmasters or people responsible for the sites in time of planning a website in accordance to AMA principles, and in order to evaluate the websites with these principles we had to have some information that were not displayed on the websites.

The results are presented as the proportion of websites with different degrees of compliance with each of the AMA guidelines (non-compliance, partial compliance, total compliance). The proportions were compared with the χ^2 or the Fisher exact test, as appropriate.

Table 8 – AMA guidelines and quality criteria.

Criteria		Accomplishment of the criteria	Code
<u>Guideline (p1) - Website ownership</u>			
Website ownership, including affiliations, strategic alliances, and significant investors, should be clearly indicated on the home screen or directly accessible from a link on the home screen	AMA_1	Yes No	1 0
<u>Guideline (p2) - Website viewing</u>			
The website should provide information about the platform(s) and browser(s) that permit optimal viewing in a location that is easy to find.	AMA_2	Yes No	1 0
<u>Guideline (p3) - Viewer access, payment and privacy</u>			
Information about restrictions on access to content, required registration, and password protection (if applicable) should be provided and easy to find. Information about payment (e.g. subscriptions, document delivery, pay per view, etc.) should be provided and easy to find. Information about privacy should be provided and easy to find.	AMA_3	Yes No	1 0
<u>Guideline (p4) – Funding and sponsorship</u>			
Funding or other sponsorship for any specific content should be clearly indicated. Content should be clearly distinguished from advertising.	AMA_4	Yes No	1 0
<u>Guideline (p5) – Quality of editorial content</u>			

A description of the editorial process and method of content review should be posted on the website.	AMA_5	Yes No	1 0
A list of staff members and other individuals (e.g. editorial board) responsible for content quality, other than anonymous peer reviewers, should be posted on the website.			
The dates that content is posted, revised, and updates should be clearly indicated.	AMA_6	Yes No	1 0
Source for specific content should be clearly identified (e.g. author byline or names of individual, organizational, departmental, institutional, agency or commercial provider/producer). Affiliations and relevant financial disclosures for authors and content producers should be clearly indicated. Reference material used to develop content should be cited in a manner appropriate for the website's audience.	AMA_7	Yes No	1 0
Guideline (p6) – Functional intrasite links			
Intrawebsite content links should be reviewed before posting and maintained and monitored. External website links should be reviewed before posting and maintained and monitored. If links are not functional, these links should be repaired in a timely manner.	AMA_8	Yes No Website don't have links	1 0 9
Guideline (p7) – Intersite navigation			
Website should not prevent viewers from returning to a previous website. Websites should not redirect the viewer to a website the viewer did not intend to visit. Website should not frame other websites without permission.	AMA_9	Yes No	1 0
Guideline (p8) – Downloading files			
If content can be downloaded in a portable document file (PDF) format, instructions regarding how to download the PDF file and how to obtain the necessary software should be provided and easy to find. A link to such software should be provided.	AMA_10	Yes No No pdf available for download	1 0 9
Guideline (p9) – Navigation of content			
Features that facilitate use of the website should be provided and easy to find, and should include a website map, or other website organizational guide, a help function or frequently-asked-questions web page, a feedback mechanism, and customer service information (if available).	AMA_11	Yes No	1 0
Each distinct website should provide a search engine or appropriate navigation tool to facilitate use. If the website provides a search engine, instructions specifying how to use the search function and how to conduct different types of search may be provided.	AMA_12	Yes No	1 0
Graphics files should include a “mouse over” indication of the graphical content. For large files, the space where the file resides should include the size of the file. As a courtesy to the viewer, when possible, when a large file can be downloaded by clicking, the viewer may be informed of the size of the file before the file begins downloading and should have the opportunity to cancel the download.	AMA_13	Yes No No graphics	1 0 9
Guideline (p10) – Advertising			
The website displays advertisements	AMA_14	Yes No	1 0
Guideline (p12) – Privacy			
The website displays its privacy policy	AMA_28	Yes No	1 0

Time spent on the use of website assessment with the Discern tool and compliance defined by the HON code, eEurope guidelines for health related websites and AMA guidelines for medical and health information sites

After a training that comprised the evaluation of 60 websites with each of the four instruments (HON code, Discern tool, eEurope guidelines for health related websites and AMA principles governing websites), 54 new websites (38 that provided information on prostate cancer and 16 that addressed breast cancer) were evaluated by the same person and the time spent on this task was registered. All websites were evaluated with each of the four instruments, systematically alternating the order by which they were used.

The time spent on the evaluation of the websites with different instruments may be used for the choice of the evaluation tool, and is being described in boxplots and compared with the Kruskal-Wallis test, according to the instrument used and the order of its application.

Reliability of the assessment of the compliance with the HON code and the eEurope guidelines for health related websites

From the resources described above for the assessment of website quality, we selected the HON code and the eEurope guidelines for health related websites for a more detailed assessment of their potential to be used as evaluation tools for standardized quantitative assessment of the quality of the websites. These guidelines are relatively easy to use, their application may be standardized, and evaluate similar aspects of the websites' quality.

The Discern tool has the advantage of having detailed instructions for its use to evaluate written materials related with health, and provides a summary measure of the assessment. However, it is limited by the fact that a large part of the instrument is devoted to the evaluation of issues related with treatments, and therefore it is not useful to assess websites providing general health information, health promotion or prevention.

The AMA principles governing websites provide a comprehensive set of criteria for the evaluation of the websites' quality. However, it is apparently more difficult to translate into an instrument for quantitative assessment because a large proportion of the items require information that is seldom available in the websites.

We evaluated the websites for each of the HON code criteria, as described in table 6. For analysis we considered the criteria from HON_1 to HON_4, HON_8 to HON_11, and the principles 3 and 7, as previously described. Each variable was coded as described in table 6 for the individual criteria and the concordance with the principles was coded as 0, 1 or 2, for no concordance, partial or complete concordance, respectively.

We also conducted an evaluation of the websites for each of the eEurope guidelines for health related websites criteria. For analysis we considered all the criteria in table 7, except the criterion 6.1, because it would be included in the analysis of the criterion 6.2. The criterion 6.2 would only be accomplished if the 6.1 criterion was also accomplished. Although the opposite could not happen, we did not consider a website as providing the credentials of the authors, when it only provided their names; it was our opinion that credentials of the providers are essential to certificate the authority of the information. Regarding the criteria 6.3, we did not include it on the evaluation because the date at which the credentials were received would be a very particular information, that most of the authors would not provide, since it is not a fundamental information to characterize an individuals' qualification.

Exploratory factor analysis (weighted least squares) on the ordinal items was used to evaluate homogeneity (e.g. to confirm if there was a single latent variable) of the item and Cronbach's alpha was used to measure the reliability.⁽¹²⁹⁾ The global goodness of fit of the underlying structure with one factor was evaluated using comparative fit index (CFI), which is recommended when N<250.⁽¹³⁰⁾

We conducted separate analysis for the HON code and the eEurope guidelines for health related websites, and the resulting scales were also evaluated together, aiming to create a new instrument, accounting for different criteria from the HON code and the eEurope guidelines for health related websites. Such an instrument is expected to be easier to apply while accounting for the main quality criteria evaluated with the other two instruments.

5.4. Analysis of the contents related to screening of breast and prostate cancer

5.4.1. Specific contents on cancer screening

A large proportion of women and men participate in screening activities for breast and prostate cancers, respectively, despite screening programs are available only for the former. We analyzed the contents of the websites on this topic, namely regarding the existence of specific information on it and its accuracy, towards an assessment of their quality in this particular aspect.

The specific information to be searched in each website was defined in accordance with the attempt of Diviani et al. to define which information on cancer should be considered essential.⁽³⁵⁾ We selected topics that cover the different options for screening and their effectiveness, the potential harms of screening, recommended periodicity, eligibility criteria for screening and instructions on how to be screened.

The criteria to assess the accuracy of the information and its adequacy to the Portuguese setting were defined in accordance with the evidence summarized by the U.S. Preventive Services Task Force (USPSTF),^(78, 131) the European Union Advisory Committee on Cancer Prevention⁽¹³²⁾ and the specific recommendations for cancer screening in Portugal.⁽¹³³⁾ For example, in Portugal there is a screening program for breast cancer, which differs slightly from the U.S. Preventive Task Force recommendations and the EU Advisory Committee on Cancer Prevention, especially regarding the age from which women should start their regular biennial mammograms,⁽¹³³⁾ and this was taken into account when defining which would be the most appropriate messages for the Portuguese population.

From each site we selected the information about screening for further analysis and then we conducted the evaluation, regarding the referred criteria.

The specific items searched, as well as the message considered the most appropriate are presented in figures 10 and 11, for breast and prostate cancer, respectively. For each item three options were possible: does not mention the subject; mentions the subject but the information is incorrect or incomplete; mentions the subject and the information is correct.

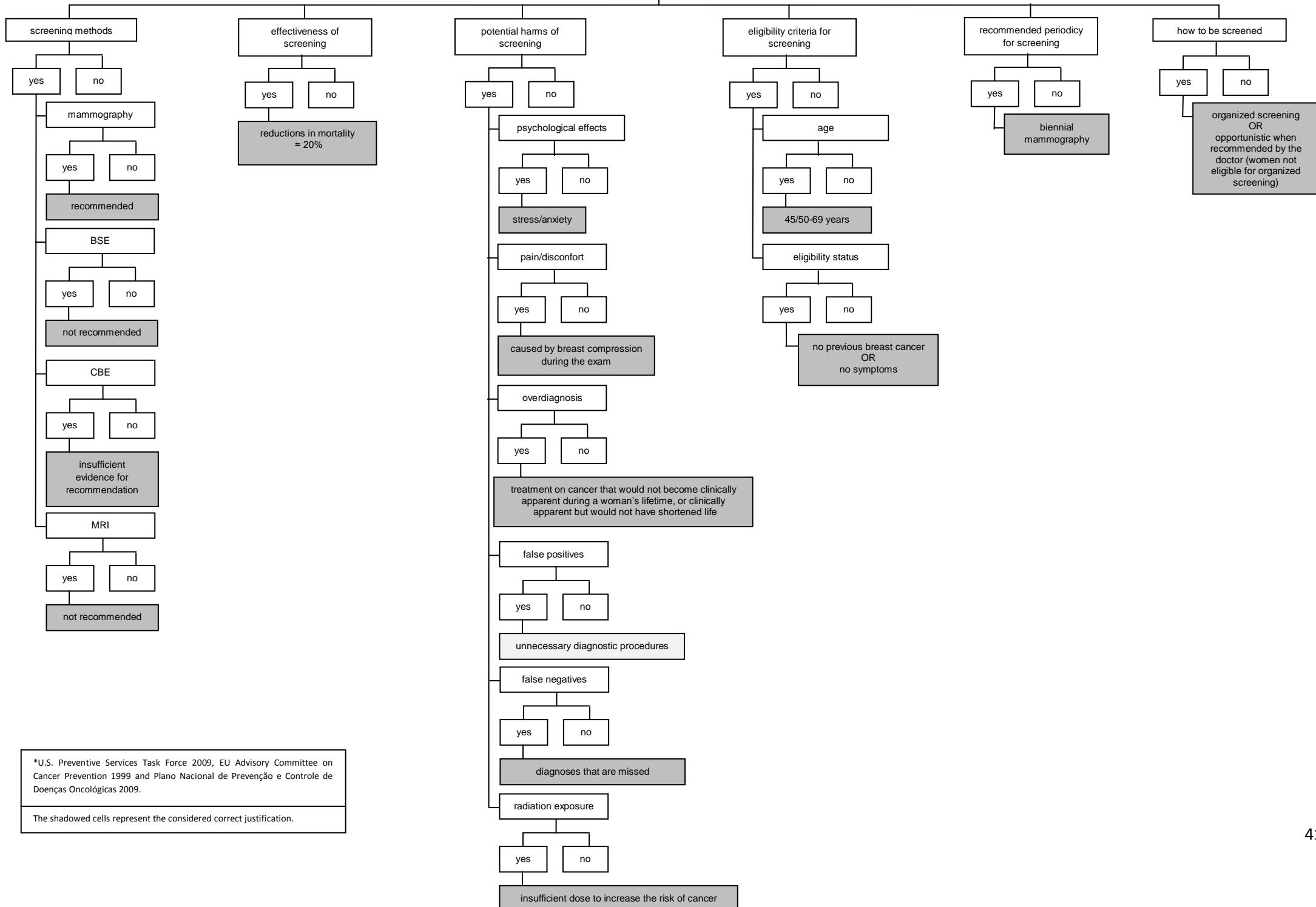
The results are presented as the proportion of websites in these categories, for each item, separately for breast and prostate cancer websites.

The items addressing similar subjects were grouped together and it was considered that the topic was not mentioned when none of the items within the topic were mentioned, mentioned with incorrect information when this applied to at least one of the items with no correct information being provided in each of them, or that the topic was mentioned correctly when this applies to at least one of the items within the same topic. This was used for comparison between breast and prostate cancer websites and, for each of them, for

comparisons according to the websites' characteristics. Proportions were compared using the chi² or the Fisher exact test, as appropriate. The distributions of the ISPUP score values obtained by the websites in each of these topic-specific categories were compared using the Kruskal-Wallis test.

Figure 10 - Procedure followed to the analysis of information on breast cancer screening.

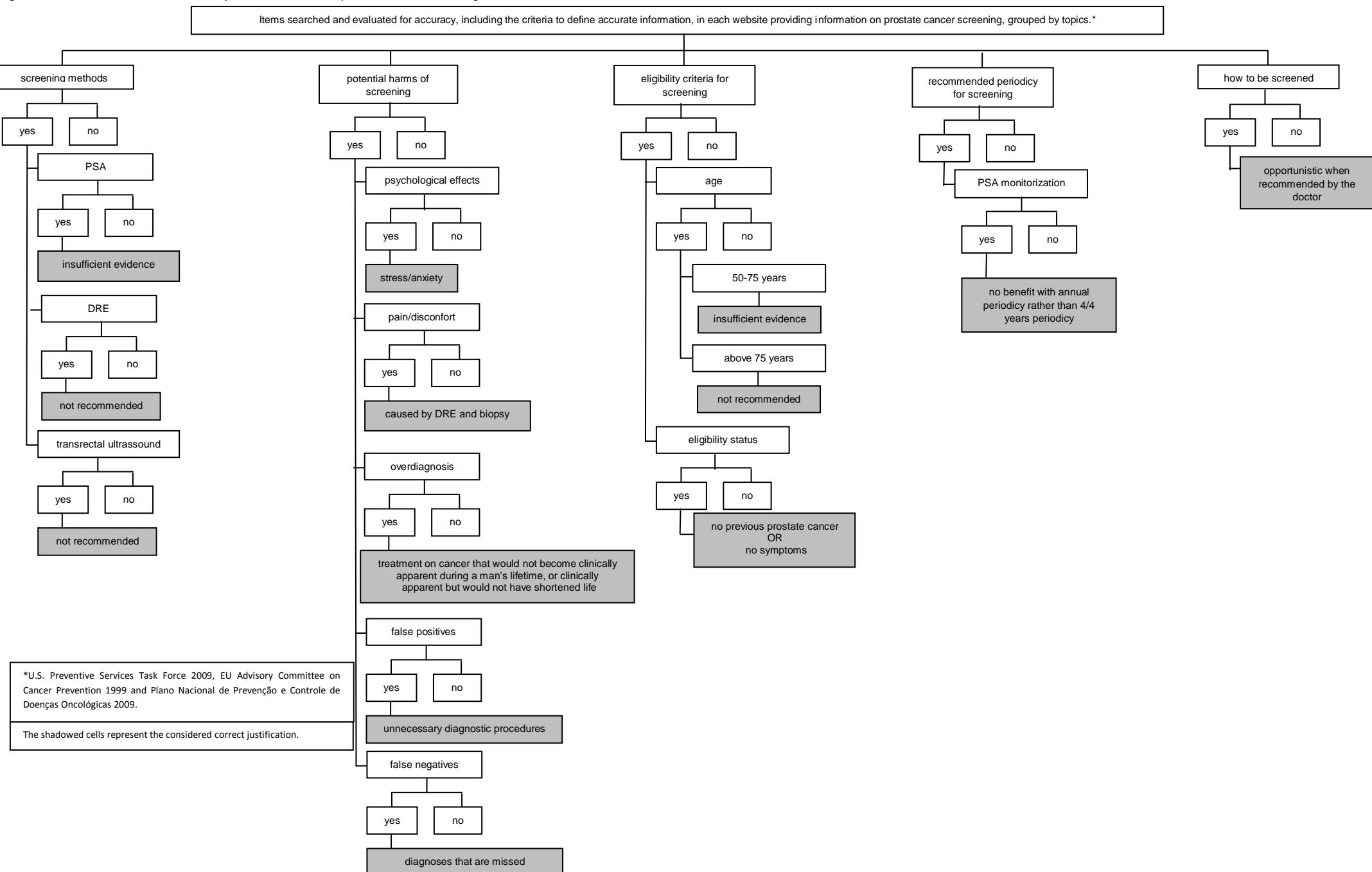
Items searched and evaluated for accuracy, including the criteria to define accurate information, in each website providing information on breast cancer screening, grouped by topics.*



*U.S. Preventive Services Task Force 2009, EU Advisory Committee on Cancer Prevention 1999 and Plano Nacional de Prevenção e Controle de Doenças Oncológicas 2009.

The shadowed cells represent the considered correct justification.

Figure 11 – Procedure followed to the analysis of information on prostate cancer screening.



*U.S. Preventive Services Task Force 2009, EU Advisory Committee on Cancer Prevention 1999 and Plano Nacional de Prevenção e Controle de Doenças Oncológicas 2009.

The shadowed cells represent the considered correct justification.

5.4.2. Readability of the contents

To assess the readability of the websites' contents addressing cancer screening we initially identified the sections with screening-related information. In the websites providing information on breast cancer we selected the text from the sections related to symptoms, diagnosis, screening and types of cancer, while in websites providing information on prostate cancer we selected information related to screening, cancer detection and diagnosis. These sections were systematically selected in all websites, to ensure comparability and because information on screening does not always appear in specific sections. These were identified as the sections more likely to address screening.

We used the Index Fernandez-Huerta to determine the readability of the contents. One of the steps in the process of readability calculation was counting the syllables, words and sentences of each part of the text. After extracting the information to a Microsoft Office Word® document, the text was analyzed using the software TextMeter®, which is an application of text statistics only for Portuguese language.⁽¹³⁴⁾ This software counts the number of words and sentences, and also provided an algorithm for counting syllables. The results in the readability index were compared between breast and prostate cancer websites and, for each of them, according to the websites' characteristics using the Kruskal-Wallis test. The association between the readability index and the ISPUP score was estimated through the Spearman correlation coefficient.

6. Results

6.1. Websites selected for analysis

In the first 200 results of two Google searches conducted to obtain information on breast and prostate cancer, we identified 47 eligible websites addressing different issues related with breast cancer in Portuguese language, and 67 websites with prostate cancer information fulfilled the eligibility criteria. Among the latter, 43 websites (64%) provided specific information on prostate cancer screening, and 35 of the former (74%) covered issues related with breast cancer screening.

The flowcharts presented as figures 12 and 13 describe the initial screening of the Internet search results for breast and prostate cancers, respectively.

Figure 12 – Selection of the Internet search results for breast cancer.

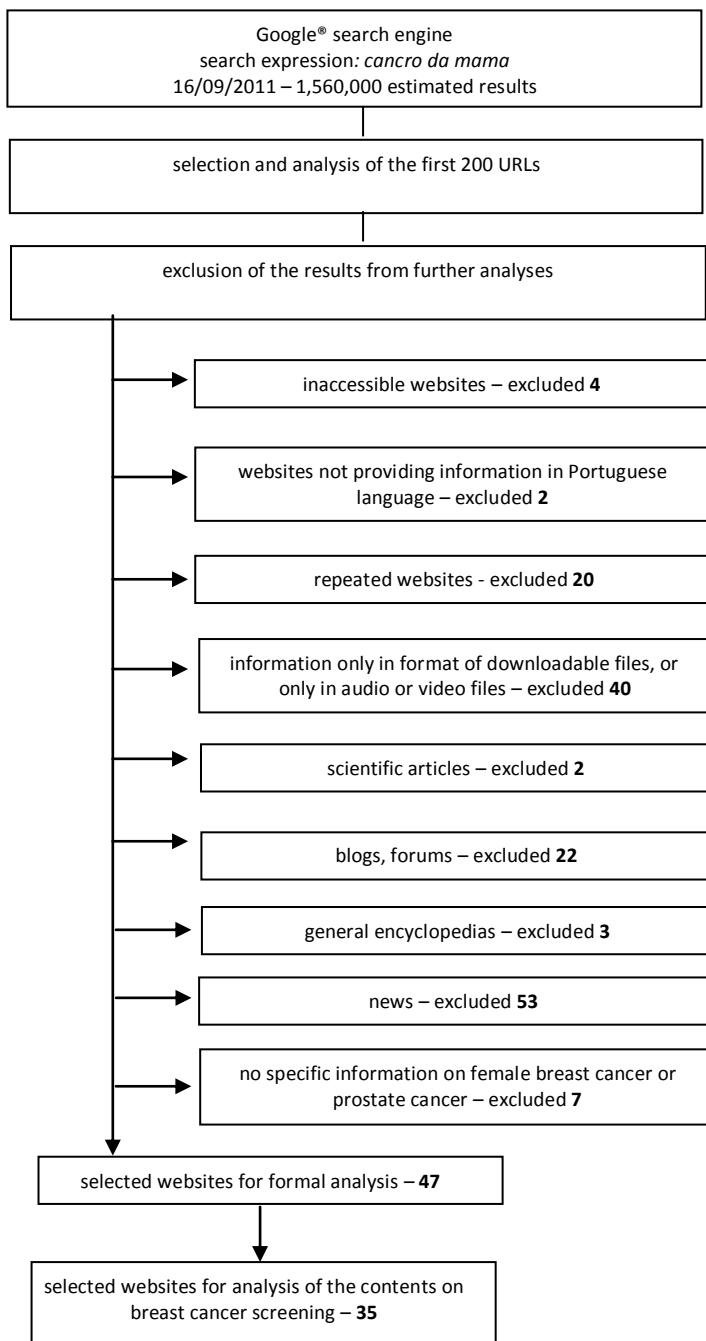
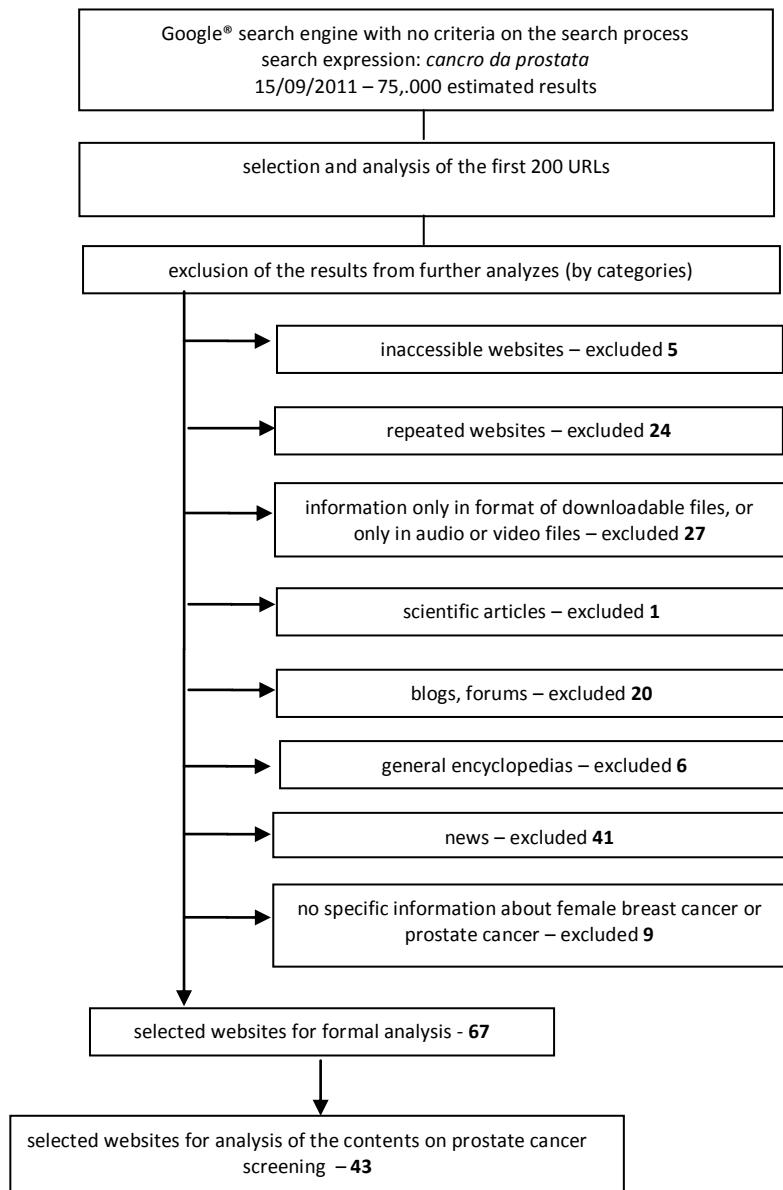


Figure 13 – Selection of the Internet search results for prostate cancer.



6.2. General characteristics of the websites

The general characteristics of the websites selected for detailed analysis are presented in table 9.

Seven out of 10 websites providing information on these cancers were health-related, and the proportion was higher among those appearing in the first thirty results of the search (86.2% vs. 67.1%, P=0.048). Approximately 20% and 10% of the websites covered exclusively issues related with cancer or specifically breast/prostate cancer, respectively; these appeared in the first pages of the search 4 and 9 times more frequently, respectively. Nearly half the websites were from Portugal, and was more likely to find a Portuguese website in the first 30 results (79.3% vs. 43.5%, P=0.004).

The websites appearing in the first three pages of results aimed more frequently the cancer patients (44.8% vs. 3.5%, P<0.001) and less often the general population (75.9% vs. 96.5%, P=0.001).

Approximately three-quarters of the websites provided the information only on the format of text; video and audio were seldom used.

There were no statistically significant differences in the characteristics of the websites according to the cancer addressed. However, those providing information on breast cancer tended to target the general population less often (85.1% vs. 95.5%, P=0.053) and those on prostate cancer were more frequently from Brazil (35.8% vs. 21.3%, P=0.123).

Table 9 – General characteristics of the websites selected for analysis.

	Websites providing health information on cancer						
	Cancer			P	Order of appearance in search		P
	All n (%)	Breast n (%)	Prostate n (%)		≤30 n (%)	>30 n (%)	
Health-specific website							
No	32 (28.1)	13 (27.7)	19 (28.4)	0.935	4 (13.8)	28 (32.9)	0.048
Yes	82 (71.9)	34 (72.3)	48 (71.6)		25 (86.2)	57 (67.1)	
Cancer-specific website							
No	92 (80.7)	38 (80.8)	54 (80.6)	0.973	16 (55.2)	76 (89.4)	<0.001
Yes	22 (19.3)	9 (19.2)	13 (19.4)		13 (44.8)	9 (10.6)	
Cancer-specific (covering exclusively breast or prostate cancer)							
No	102 (89.5)	42 (89.4)	60 (89.6)	0.974	20 (69.0)	82 (96.5)	<0.001
Yes	12 (10.5)	5 (10.6)	7 (10.4)		9 (31.0)	3 (3.5)	
Country of origin							
Portugal	60 (52.6)	30 (63.8)	30 (44.8)	0.123	23 (79.3)	37 (43.5)	0.004
Brazil	34 (29.8)	10 (21.3)	24 (35.8)		4 (13.8)	30 (35.3)	
Other	20 (17.5)	7 (14.9)	13 (19.4)		2 (6.9)	18 (21.2)	
Intended audience							
General population	104 (91.2)	40 (85.1)	64 (95.5)	0.053	22 (75.9)	82 (96.5)	0.001
Cancer patients	16 (14.1)	9 (19.2)	7 (10.4)	0.188	13 (44.8)	3 (3.5)	<0.001
Media professionals	2 (1.8)	1 (2.1)	1 (1.5)	>0.999	2 (6.9)	0 (0.0)	0.063
Health professionals	9 (7.9)	5 (10.6)	4 (6.0)	0.363	2 (6.9)	7 (8.2)	0.817
Display of contents							
Text only	83 (72.8)	37 (78.7)	46 (68.7)	0.234	21 (72.4)	62 (72.9)	0.956
Text and figures	27 (23.7)	8 (17.1)	19 (28.4)	0.161	6 (20.7)	21 (24.7)	0.660
Text and videos	4 (3.5)	2 (4.3)	2 (3.0)	0.717	2 (6.9)	2 (2.4)	0.251
Text and charts	8 (7.1)	4 (8.5)	4 (6.0)	0.601	2 (6.9)	6 (7.1)	0.976
Text and audio	0 (0.0)	0 (0.0)	0 (0.0)	---	0 (0.0)	0 (0.0)	---

6.3. The formal quality of the websites according to the Discern tool, the Health On Net code, the eEurope guidelines for health related websites and the American Medical Association guidelines for medical and health information sites

The results of the application of the Discern instrument are presented in table 10. The median score was low for most criteria. The highest score was observed for the criteria related with the possibility of having more than one treatment option, with most websites having the maximum classification.

There were no significant differences between the websites addressing breast or prostate cancer.

Table 10 – Compliance of the websites providing information on cancer treatments (n=81) with the Discern criteria.

Rating according to the Discern criteria	Websites providing health information on cancer			
	All Median (P25-P75)	Breast cancer Median (P25-P75)	Prostate cancer Median (P25-P75)	P
Discern 1 – Clear aims	2.0 (1.0-3.0)	1.0 (1.0-3.0)	3.0 (1.0-5.0)	0.371
Discern 2 – Aims achievement	2.0 (1.0-3.0)	1.0 (1.0-3.0)	1.0 (3.0-3.0)	0.538
Discern 3 – Relevance of information	3.0 (3.0-5.0)	3.0 (3.0-5.0)	3.0 (3.0-5.0)	0.826
Discern 4 – Clear information on sources used for information	1.0 (1.0-3.0)	1.0 (1.0-3.0)	1.0 (1.0-3.0)	0.754
Discern 5 – Clear information on date of sources of information	1.0 (1.0-1.0)	1.0 (1.0-1.0)	1.0 (1.0-1.0)	0.841
Discern 6 – Balanced and unbiased information	3.0 (2.0-3.0)	3.0 (3.0-3.0)	3.0 (2.0-3.0)	0.895
Discern 7 – Information of additional source for support	1.0 (1.0-1.0)	1.0 (1.0-3.0)	1.0 (1.0-1.0)	0.431
Discern 8 – Reference of uncertain areas	1.0 (1.0-2.0)	1.0 (1.0-1.0)	1.0 (1.0-2.0)	0.259
Discern 9 – Description of how treatment works	3.0 (1.0-3.0)	3.0 (2.0-5.0)	3.0 (1.0-3.0)	0.263
Discern 10 – Description of the benefits of the treatment	1.0 (1.0-2.0)	1.0 (1.0-3.0)	1.0 (1.0-2.0)	0.534
Discern 11 – Description of risks of the treatment	2.0 (1.0-3.0)	2.0 (1.0-5.0)	2.0 (1.0-3.0)	0.574
Discern 12 – Description of what happened if no treatment were used	1.0 (1.0-1.0)	1.0 (1.0-1.0)	1.0 (1.0-1.0)	0.639
Discern 13 – Description of how treatment affects overall quality of life	1.0 (1.0-1.0)	1.0 (1.0-1.0)	1.0 (1.0-1.0)	0.977
Discern 14 – Clear possibility of more than one treatment option	5.0 (5.0-5.0)	5.0 (5.0-5.0)	5.0 (5.0-5.0)	0.758
Discern 15 – Description on support for shared decision-making	1.0 (1.0-1.0)	1.0 (1.0-2.0)	1.0 (1.0-1.0)	0.300
Discern 16 – Overall rating	2.0 (2.0-3.0)	2.0 (2.0-3.0)	2.0 (2.0-3.0)	0.611

Table 11 presents the results of the application of the HON code to websites providing information on breast and prostate cancer.

The principles financial disclosure, transparency and justifiability had the higher proportions of websites with total agreement. No significant differences were observed between websites addressing breast or prostate cancer, except for the principle authority, with which none of the breast websites was in total agreement.

Noteworthy is the fact that only 7.9% of the websites had the HON Seal.

Table 11 – Compliance of the websites with the principles of the HON code.

	Websites providing health information on cancer			
	All n (%)	Breast cancer n (%)	Prostate cancer n (%)	P
Principle 1 – Authority				
No	62 (54.4)	32 (68.1)	30 (44.8)	0.019
Yes (partial)	47 (41.2)	15 (31.9)	32 (47.8)	
Yes (total)	5 (4.4)	0 (0.0)	5 (7.5)	
Principle 2 – Complementarity				
No	35 (30.7)	15 (31.9)	20 (29.8)	0.217
Yes (partial)	62 (54.4)	22 (46.8)	40 (59.7)	
Yes (total)	17 (14.9)	10 (21.3)	7 (10.4)	
Principle 3 – Confidentiality				
No	64 (56.1)	24 (51.1)	40 (59.7)	0.526
Yes (partial)	45 (39.5)	20 (42.6)	25 (37.3)	
Yes (total)	5 (4.4)	3 (6.4)	2 (3.0)	
Principle 4 – Attribution				
No	2 (1.8)	0 (0.0)	2 (3.0)	0.153
Yes (partial)	77 (67.5)	36 (76.6)	41 (61.2)	
Yes (total)	35 (30.7)	11 (23.4)	24 (35.8)	
Principle 5 – Justifiability				
No	36 (31.6)	15 (31.9)	21 (31.3)	0.896
Yes	6 (5.3)	3 (6.4)	3 (4.5)	
NA (no claims)	72 (63.2)	29 (61.7)	43 (64.2)	
Principle 6 – Transparency				
No	21 (18.4)	8 (17.0)	13 (19.4)	0.747
Yes	93 (81.6)	39 (83.0)	54 (80.6)	
Principle 7 – Financial disclosure				
No	19 (16.7)	5 (10.6)	14 (20.9)	0.148
Yes	95 (83.3)	42 (89.4)	53 (79.1)	
Principle 8 – Advertising				
No	14 (12.3)	7 (14.9)	7 (10.4)	0.481
Yes	54 (47.4)	24 (51.1)	30 (44.8)	
NA (no advertising)	46 (40.4)	16 (34.0)	30 (44.8)	
HON Seal				
No	105 (92.1)	42 (89.4)	63 (94.0)	0.363
Yes	9 (7.9)	5 (10.6)	4 (6.0)	

Table 12 presents the results of the evaluation conducted using the eEurope guidelines.

Less than half of the websites complied with each of the items from the guidelines, except for the accountability (82.5%). Significant differences between breast and prostate cancer websites were observed only for the criterion related with the display of the credentials from the information providers, which was observed nearly 4 times more frequently among the prostate cancer related websites.

Table 12 – Compliance of the websites with the eEurope guidelines for health related websites.

	Websites providing health information on cancer			
	All	Breast cancer	Prostate cancer	P
	n (%)	n (%)	n (%)	
eEu p1 – Transparency and honesty				
No	8 (7.0)	2 (4.3)	6 (9.0)	0.559
Yes (partial)	80 (70.2)	35 (74.5)	45 (67.2)	
Yes (total)	26 (22.8)	10 (21.3)	16 (23.9)	
eEu 5 – Authority – source and date of information				
No	93 (81.6)	37 (78.7)	56 (83.6)	0.510
Yes	21 (18.4)	10 (21.3)	11 (16.4)	
eEu 6.2 - Authority – name and credentials of information providers				
No	93 (81.6)	44 (93.6)	49 (73.1)	0.005
Yes	21 (18.4)	3 (6.4)	18 (26.9)	
eEu 7 – Privacy and protection data				
No	65 (57.0)	24 (51.1)	41 (61.2)	0.282
Yes	49 (43.0)	23 (48.9)	26 (38.8)	
eEu 8 – Updating and information				
No	80 (70.2)	36 (76.6)	44 (45.7)	0.209
Yes	34 (29.9)	11 (23.4)	23 (34.3)	
eEu 9 - Accountability				
No	20 (17.5)	8 (17.2)	12 (17.9)	0.902
Yes	94 (82.5)	39 (83.0)	55 (82.1)	
eEu 11 – Editorial policy				
No	81 (71.0)	48 (71.6)	33 (70.2)	0.868
Yes	33 (29.0)	19 (28.4)	14 (29.8)	

Table 13 presents the results of the assessment of the compliance with the AMA guidelines for medical and health information sites.

The proportion of websites in total agreement with the principles was the highest for those related with intersite navigation (100%), functional intrasite links (98.8%) and funding and sponsorship (83.2%), and the lowest for those referring to site viewing (3.5%) and navigation on content (5.3%). The results show that there are no significant differences between the different websites, but regarding the principle AMA 3, related to viewer access, payment and privacy, there were significant differences, once the websites providing information on breast cancer tend to accomplish more this principle than websites providing information on prostate cancer (61.7% vs. 38.8%, P=0.016).

Table 13 – Compliance of the websites with the AMA guidelines for medical and health information sites.

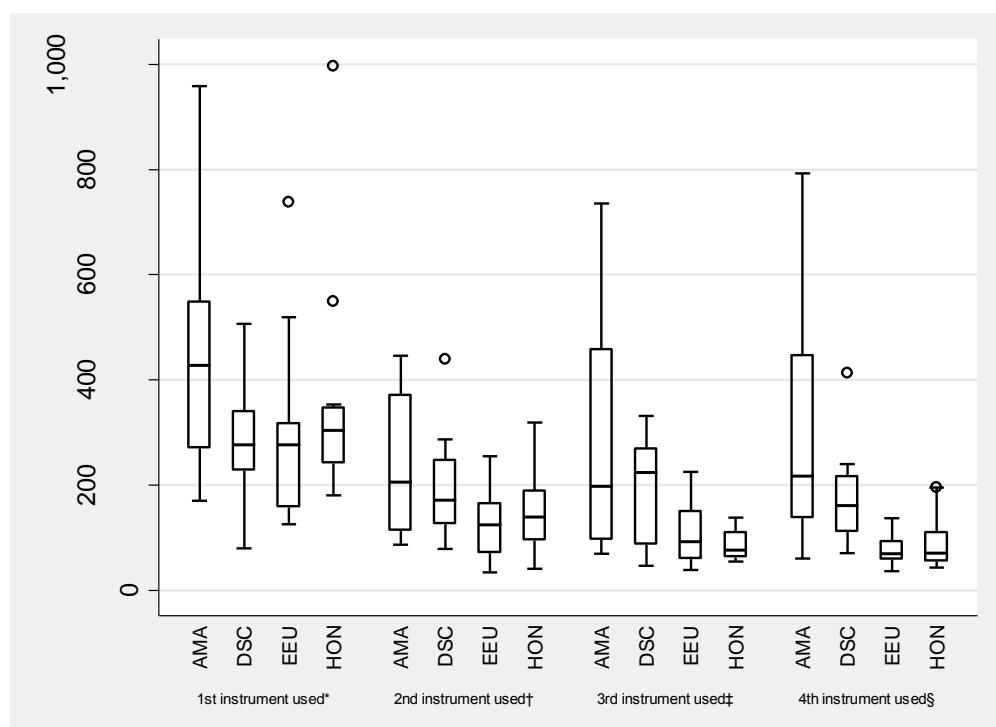
	Websites providing health information on cancer				P
	All		Breast cancer	Prostate cancer	
	n (%)	n (%)	n (%)	n (%)	
AMA p1 – site ownership	No Yes	32 (28.1) 82 (71.9)	13 (27.7) 34 (72.3)	19 (28.4) 48 (71.6)	0.935
AMA p2 – site viewing	No Yes	110 (96.5) 4 (3.5)	44 (93.6) 3 (6.4)	66 (98.5) 1 (1.5)	0.304
AMA p3 – viewer access, payment and privacy	No Yes	59 (51.8) 55 (48.2)	18 (38.3) 29 (61.7)	41 (61.2) 26 (38.8)	0.016
AMA p4 – funding and sponsorship	No Yes	19 (16.7) 95 (83.3)	5 (10.6) 42 (89.4)	14 (20.9) 53 (79.1)	0.148
AMA p5 – quality of editorial content	No Yes (partial) Yes (total)	53 (46.5) 52 (45.6) 9 (7.9))	22 (46.8) 20 (42.6) 5 (10.6)	31 (46.3) 32 (47.8) 4 (6.0)	0.629
AMA p6 – functional intrasite links	No Yes	1 (1.2) 81 (98.8)	1 (2.9) 34 (97.1)	0 (0.0) 47 (100.0)	0.427
AMA p7 – intersite navigation	No Yes	0 (0.0) 114 (100.0)	0 (0.0) 114 (100.0)	0 (0.0) 114 (100.0)	---
AMA p8 – downloading files	No Yes	11 (68.8) 5 (31.2)	6 (66.7) 3 (33.3)	5 (71.4) 2 (28.6)	>0.999
AMA p9 – navigation on contents	No Yes (partial) Yes (total)	3 (2.6) 105 (92.1) 6 (5.3)	2 (4.3) 43 (91.5) 2 (4.3)	1 (1.5) 62 (92.5) 4 (6.0)	0.619
AMA p10* – advertising	No Yes	11 (6.88) 5 (31.2)	6 (66.7) 3 (33.3)	5 (71.4) 2 (28.6)	>0.999
AMA p12* - privacy	No Yes	65 (57.0) 49 (43.0)	24 (51.1) 23 (48.9)	41 (61.2) 26 (38.8)	0.282

* We did not have all the information to complete the full evaluation on the principle. Some information only could be given by the site's webmaster, and the criterion could not be evaluated based on the information available on the website.

Time spent on the evaluation of the websites

Figure 14 depicts the time spent in the evaluation of the websites according to the order of use the instruments. When each instrument was the first to be used there were no significant differences in the time spent in the evaluation, although the median was higher for the instrument based on the AMA principles. Statistically significant differences were observed only when the instruments were used in a third or fourth evaluation of the same websites, with lower figures registered for the HON and eEurope guidelines.

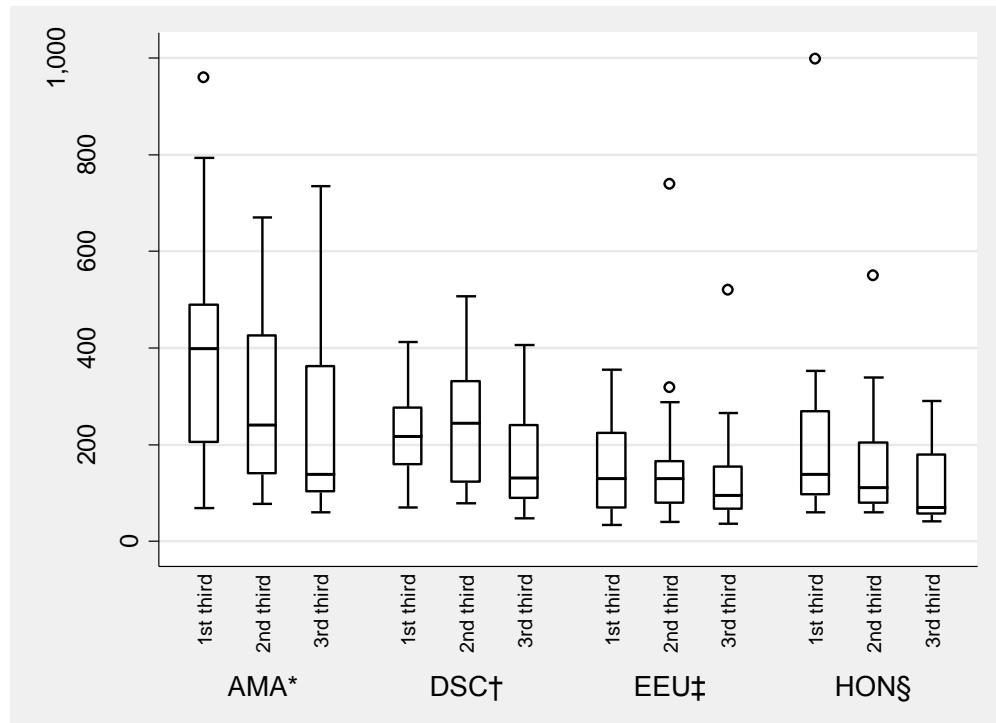
Figure 14 – Time spent in the evaluation of 54 websites with each of four instruments, according to the order of use of the instruments.



* P=0.111; † P=0.955; ‡ P=0.005; § P<0.001

Figure 15 presents time spent in the evaluation of the websites with each instrument according to the order by which the websites were evaluated. The median time necessary for the evaluations decreased from the first 18 websites (1st third) to the last 18 websites (3rd third). The variation was smaller when using the eEurope guidelines (1st third vs. 3rd third: 129.5 vs. 95 seconds) and statistically significant differences were observed only for the HON instrument.

Figure 15 – Time spent in the evaluation of 54 websites with each of four instruments, according to the order by which the websites were evaluated.



* P=0.094; † P=0.087; ‡ P=0.579; § P<0.013

6.3.1. Reliability of the assessment instruments

The initial assessment of the instrument based on the HON code showed high correlations between the variable HON_7 – respect for legal requirements – and several others and was dropped. The variables referring to the assessment of HON_2 – credentials of medical information providers, HON_5 – intended audience – and principle 8 – advertising policy – were dropped due to their strong correlation with HON_1 – name of the editor – and HON_4 – intended mission of the site – and principle 7 – financial disclosure –, respectively. The analysis including the remaining variables identified a single latent variable as the best solution. HON_10 – claims of the site – and HON_8 – date of updating – were then dropped due to a correlation <0.3 with the single factor and the Cronbach's alpha was 0.606 (table 14). The scores assigned to each variable were added up to form a score based on the HON code (HON-score).

For the instrument based on the eEu guidelines, a single latent variable was also considered as the best solution, and the Cronbach's alpha was 0.707 (table 14). The scores assigned to each variable were added up to form a score (eEu-score).

When the two instruments were combined to produce a single one, the variables referring to the items eEu_2 – purpose and objective of the website, eEu_7 – privacy and data protection, eEu_9 – accountability – and the principle 7 of HON – disclosure of funding sources – were dropped because the sample correlations were higher than 0.99 with the items HON_4 – intended mission, HON_6 – privacy policy, HON_11 – website contact – and eEu_4 – transparency of sources of funding, respectively. Strong correlations were also observed between HON_9 – sources of external information – and eEu_5 – sources of information, and between eEu_6.2 – name and credentials of information providers – and HON_1 – name of the editor, respectively, and HON_9 and eEu_6.2 were dropped (table 14). Among each highly correlated pair the variable chosen to be dropped was the one perceived as more ambiguous or difficult to assess. The scores assigned to each variable were added up to form a new score that was named ISPUP-score (table 15).

In order to evaluate their homogeneity, an exploratory factor analysis was conducted for each scale (HON-score, eEu-score and ISPUP-score). This analysis revealed a first factor that accounted more than 40% of the variance, and the first eigenvalue was 3 times larger than the second eigenvalue. The fit index met the criteria to support the 1 factor structure, the CFI ranged from 0.911 to 0.951. The Cronbach's alpha ranged from 0.61 and 0.74 (table 14).

Table 14 - Exploratory factor analysis and internal consistency conducted for the instruments based on the HON code, EEU guidelines and combination of the two scales.

Items	Factor 1	Factor 1	Factor 1
HON1	0.536		0.684
HON3	0.402		0.426
HON4	0.443		0.592
HON6	0.831		0.652
HON8	0.246		
HON9	0.427		
HON10	0.033		
HON11	0.855		0.851
P7HON	0.645		
EEU1		0.944	0.990
EEU2		0.588	
EEU3		0.565	0.527
EEU4		0.584	0.553
EEU5		0.399	0.416
EEU6.2		0.383	
EEU7		0.627	
EEU8		0.337	0.346
EEU9		0.910	
EEU11		0.758	0.766
Eigenvalue 1	3.194	4.468	4.876
Eigenvalue 2	1.414	1.534	1.607
CFI	0.911	0.951	0.928
Alpha Cronbach	0.606	0.707	0.738

CFI – Comparative Fit Index

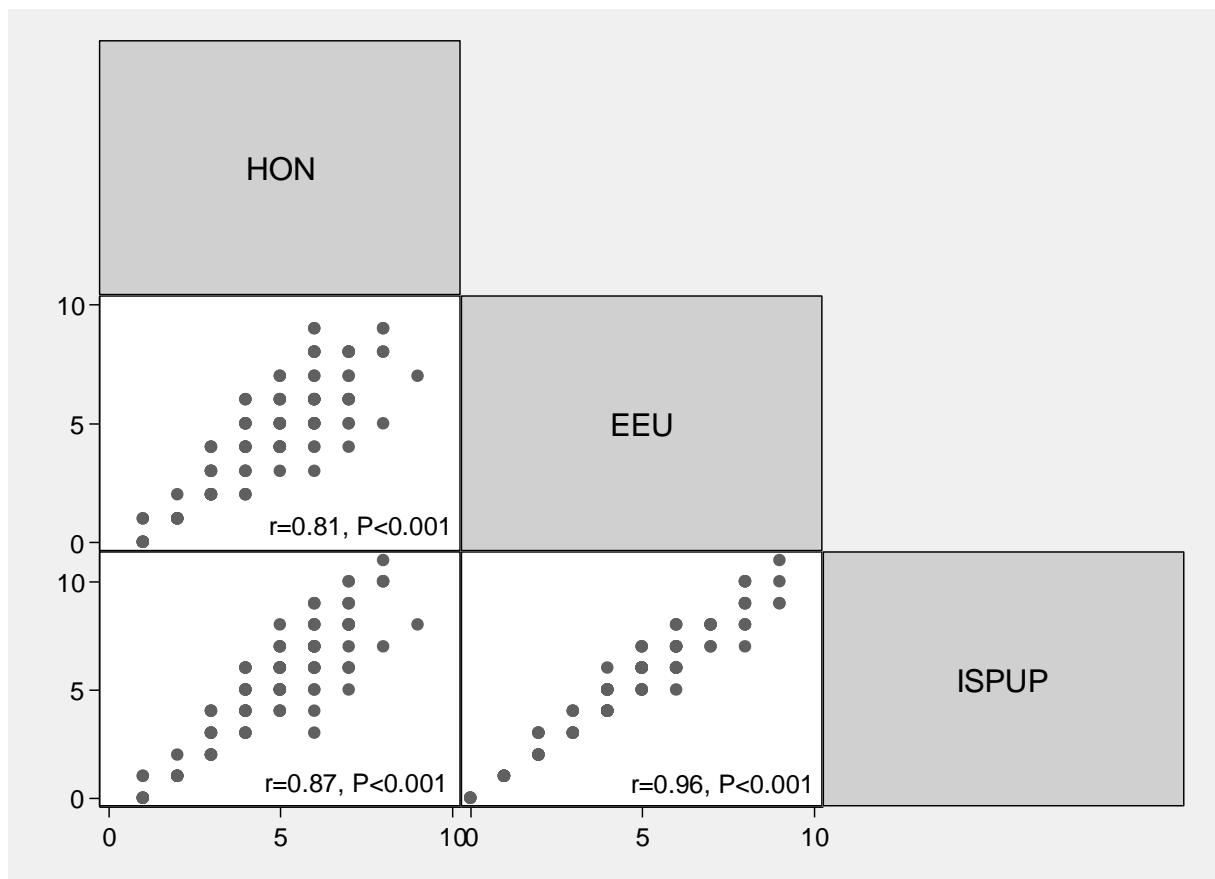
The three scores were strongly correlated (Spearman correlation coefficients >0.80), with the stronger correlations observed between the ISPUP-score and the other two (figure 16).

When comparing the 9 websites to which the HON Seal was attributed with the remaining 105, regarding the quality scores, the median scores were significantly higher in the groups of websites that display the seal, either for the HON-score (7 vs. 5, P=0.002), eEu-score (7 vs. 5, P=0.003), or the ISPUP-score (8 vs. 5, P=0.001).

Table 15 – ISPUP-score criteria (adapted from HON code and eEurope guidelines for health related websites).

Criteria	Origin of criteria
1 – Intended mission of the site	derived from “A statement describing the intended mission of the website is provided.” - HON_4 and “Transparency of purpose and objective of the website” - eEu_2
2 – Statement declaring that information provided on the website is meant to complement and not replace any advice or information from a health professional	derived from “A statement declaring that information provided on the website is meant to complement and not replace any advice or information from a health professional is clearly provided.” - HON_3
3 – Privacy policy	derived from “Privacy/Confidentiality policy regarding e-mail addresses, personal and medical information is displayed on the website.” HON_6 and “Privacy and data protection policy and system for the processing of personal data, including processing invisible to users, to be clearly defined in accordance with community Data Protection legislation” - eEu_7
4 – Website contact	derived from “A valid email address for the webmaster or a link to a valid contact form is easily accessible throughout the website?” - HON_11 and “User feedback, and appropriate oversight responsibility (such as a named quality compliance office of each website)” - eEu_9
5 – Transparency of sources of funding	derived from “Is the source of the funding of the website clearly described? for commercial or non-commercial organizations:” and “Is the source of the funding of the website clearly described? for personal or private websites, or those hosted without charge” - HON_principle7 and “Transparency of all sources of funding for website (grants, sponsors, advertisers, non-profit, voluntary assistance)” - eEu_4
6 – Credentials of sources of information	derived from “Does the website contain information from external sources?” - HON_9 and “Clear statement of sources for all information provided and date of publication of the source” - eEu_5
7 – Name of the editor	derived from “The website provides general information about the organization or individual responsible for its operation and content, and a person is named as editor or principal author.” - HON_1 and “Credentials of the human/institutional providers of information up on the website” - eEu_6.2)

Figure 16 – Association between the scores obtained with the instruments based on the HON code and eEurope guidelines and the ISPUP score.



6.4. Quality of contents related to screening of breast or prostate cancer and readability

The results presented in figure 17 refer to the evaluation of screening related contents in websites providing information on breast cancer screening.

Most websites mentioned mammography as a method for breast cancer screening (80%), although only 28% mentioned it correctly as the only recommended method for screening, and correct quantitative estimates of the effectiveness are provided in only 14%. The breast self-exam and the clinical breast exam are mentioned almost as often as the mammography, but rarely provide correct information. MRI is seldom referred.

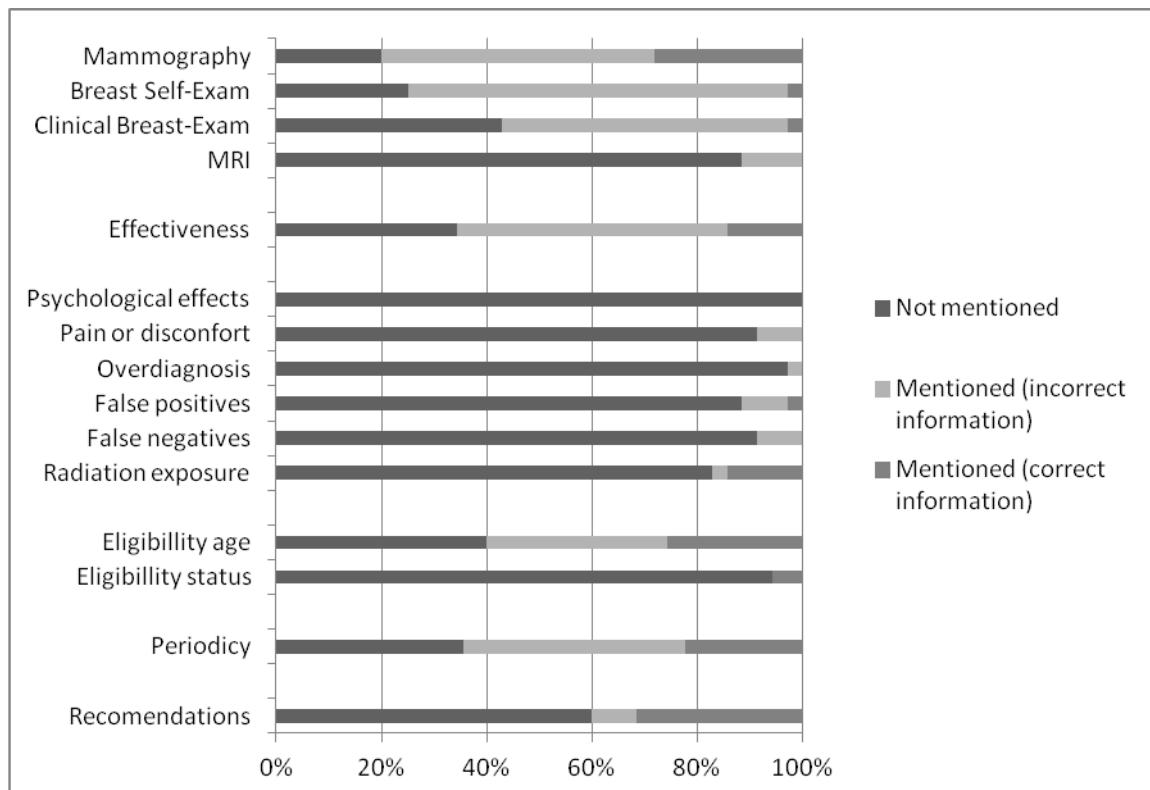
Most of the potential harms of screening were not approached by nearly all websites.

The potential for overdiagnosis, false positive and false negative results was addressed in a very low percentage of the websites, and most of the times the information was not correct. Radiation exposure was correctly mentioned in 14.3% of the websites.

Approximately one-quarter of the websites gave correct information about the eligible ages for screening, but the fact that screening is conducted in asymptomatic subjects was seldom addressed.

The adequate periodicity for screenings was mentioned in 22.2% of the websites and the recommendations on how to be screened were correct in 31.4% of the websites.

Figure 17 – Quality of the contents on screening in breast cancer websites (n=35).



The results expressed in figure 18 refer to the evaluation of screening related contents in websites providing information on prostate cancer screening.

PSA is mentioned as a possible screening test in nearly all websites, but the correct information on the PSA screening was given in less than 10%. Incorrect information on DRE was given nearly as frequently as PSA, and transrectal ultrasound was referred in nearly half the websites.

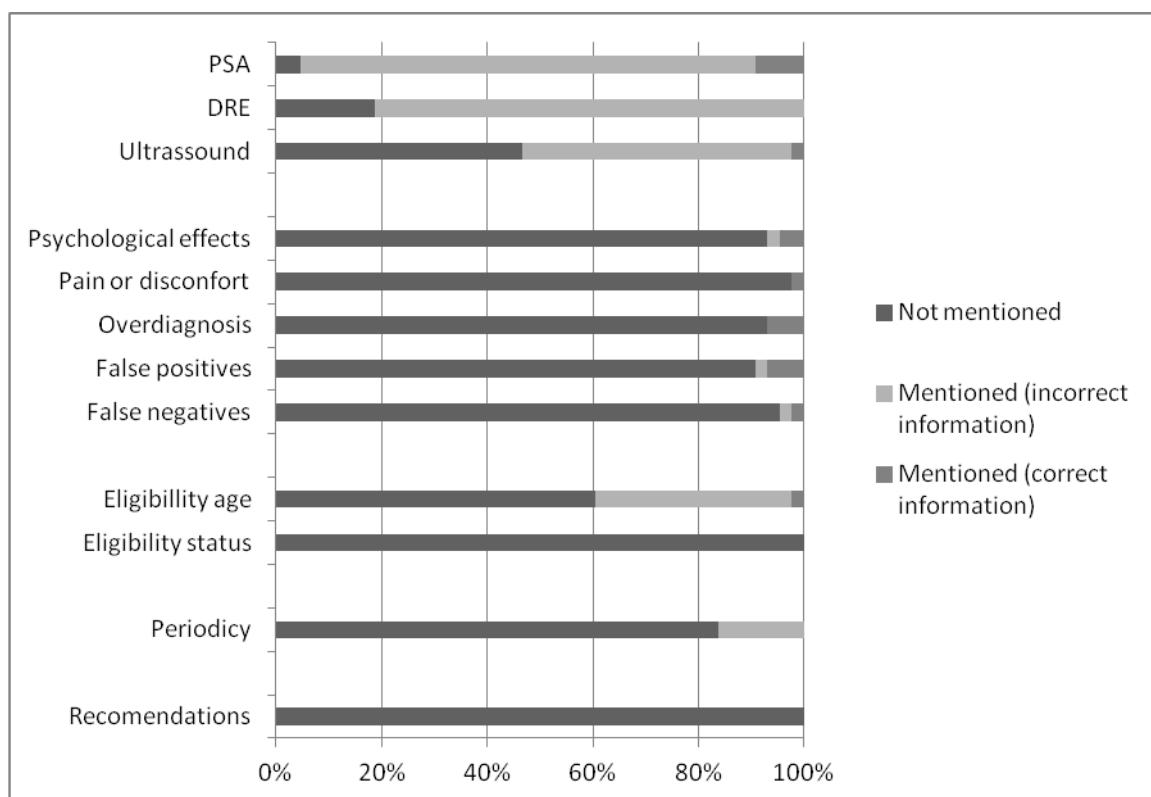
A low percentage of the websites addressed issues related to the harms of an execution of the screenings. The most referred harm with the correct information was the potential for overdiagnosis and false positive (both 6.9%).

None of the websites mentioned that screening was to be conducted by asymptomatic subjects, and the age-groups potentially eligible for screening were addressed by 39.5%, most of the times incorrectly.

The periodicity of the screening was mentioned in less than a fifth of the websites and never with the correct information.

None of the websites provided information on how to be screened.

Figure 18 – Quality of the contents on screening in prostate cancer websites (n=43).



The comparison of websites' contents on breast and prostate cancer screening is described in table 16.

The information on mammography screening is provided correctly more often than information referring to PSA screening (48.6% vs. 9.3%, P<0.001).

About the eligibility for screening and periodicity of testing, approximately one-quarter of the breast cancer websites provided correct information while nearly all prostate cancer websites did not mention or gave incorrect information on these topics.

No statistically significant differences between the breast and prostate cancer websites were observed in relation to the information on harms potentially caused by screening.

No significant differences were observed between the breast and prostate cancer websites regarding readability, with the median index values being approximately 70.

Table 16 – Quality and readability of the contents on breast cancer screening and prostate cancer screening.

	Websites providing information on screening			
	breast cancer n (%)	prostate cancer n (%)	P	
Screening methods*			<0.001	
not mentioned	5 (14.3)	2 (4.65)		
mentioned/incorrect	13 (37.1)	37 (86.0)		
mentioned/correct	17 (48.6)	4 (9.3)		
Effectiveness†			---	
not mentioned	12 (34.3)	---	---	
mentioned/incorrect	18 (51.4)	---		
mentioned/correct	5 (14.3)	---		
Harms‡			0.755	
not mentioned	28 (80.0)	35 (81.4)		
mentioned/incorrect	3 (8.6)	2 (4.6)		
mentioned/correct	4 (11.4)	6 (14.0)		
Eligibility			0.007	
not mentioned	14 (40.0)	26 (60.5)		
mentioned/incorrect	12 (34.29)	16 (37.2)		
mentioned/correct	9 (25.7)	1 (2.3)		
Periodicity			0.000	
not mentioned	16 (45.7)	36 (83.7)		
mentioned/incorrect	9 (25.7)	7 (16.3)		
mentioned/correct	10 (28.6)	0 (0.0)		
Access on screening			0.000	
not mentioned	21 (60.0)	43 (100.0)		
mentioned/incorrect	3 (8.6)	0 (0.0)		
mentioned/correct	11 (31.4)	0 (0.0)		
Readability			0.144	
	median (P25-P75)	73.1 (68.9-76.0)	69.7 (63.5-77.8)	

* Only mammography and PSA screening are being considered in this variable.

† There is no evidence about the effectiveness of screening of the prostate cancer.

‡† The information on potential harms associated with radiation exposure are not being considered to ensure comparability with prostate cancer

not mentioned – none of the items that are included in the topic was mentioned; mentioned/incorrect – at least one item that is included in the topic was mentioned incorrectly, and none was mentioned correctly; mentioned/correct – at least one item that is included in the topic was mentioned correctly.

Table 17 presents the results of the analysis on breast cancer screening contents and corresponding readability.

The websites appearing on the first 30 results tend to have better information about screening harms (30.0% vs. 4.0%, $P=0.014$), and about periodicity of screening (70.0% vs. 12.0%, $P=0.004$).

Cancer-specific websites present better information about the potential harms of screening (42.9% vs. 3.6%, $P=0.033$), and correct instruction on how to be screened (71.4% vs. 21.4%, $P=0.038$). Readability, however, tended to be lower in the cancer-specific websites (median: 68.9 vs. 73.7, $P=0.058$).

The websites which had an origin not Portuguese also have better readability (median: 75.7 vs. 70.2, $P=0.036$).

No significant differences were found in the analysis of contents on screening of the prostate cancer (table 18).

Table 17 – Quality of the contents on breast cancer screening according to websites' order of appearance, country of origin, specific for cancer and ISPUP score.

	Contents on breast cancer screening										ISPUP Score	
	Order of appearance in search			Country of origin			Cancer-specific website					
	<=30 n (%)	>=30 n (%)	P	Portuguese n (%)	Other n (%)	P	No n (%)	Yes n (%)	P	median (P25-P75)	p	
Screening methods												
not mentioned	0 (0.0)	5 (20.0)	0.458	4 (16.7)	1 (9.1)	0.788	5 (17.9)	0 (0.0)	0.393	6 (7-7)	0.236	
mentioned/incorrect	4 (40.0)	9 (36.0)		8 (33.3)	5 (45.4)		11 (39.3)	2 (28.6)		5 (3-6)		
mentioned/correct	6 (60.0)	11 (44.0)		12 (50.0)	5 (45.4)		12 (42.9)	5 (71.4)		6 (4-7)		
Effectiveness												
not mentioned	2 (20.0)	10 (40.0)	0.588	8 (33.3)	4 (36.4)	>0.999	10 (35.7)	2 (28.6)	>0.999	6 (3-7)	0.800	
mentioned/incorrect	6 (60.0)	12 (48.0)		12 (50.0)	6 (54.6)		14 (50.0)	4 (57.1)		6 (5-7)		
mentioned/correct	2 (20.0)	3 (12.0)		4 (16.7)	1 (9.1)		4 (14.3)	1 (14.3)		4 (4-8)		
Harms												
not mentioned	5 (50.0)	23 (92.0)	0.014	18 (75.0)	10 (90.9)	0.442	24 (85.7)	4 (57.1)	0.033	5 (3-7)	0.287	
mentioned/incorrect	2 (20.0)	1 (4.0)		2 (8.3)	1 (9.1)		3 (10.7)	0 (0.0)		8 (4-9)		
mentioned/ correct	3 (30.0)	1 (4.0)		4 (16.7)	0 (0.0)		1 (3.6)	3 (42.9)		6.5 (6-7)		
Eligibility												
not mentioned	2 (20.0)	12 (48.0)	0.224	10 (41.7)	4 (36.4)	<0.001	11 (39.3)	3 (42.9)	0.369	5.5 (2-7)	0.969	
mentioned/incorrect	4 (40.0)	8 (32.0)		8 (33.3)	4 (36.4)		11 (39.3)	1 (14.3)		5.5 (4-7.5)		
mentioned/ correct	4 (40.0)	5 (20.0)		6 (25.0)	3 (27.3)		6 (21.4)	3 (42.9)		6 (4-7)		
Periodicity												
not mentioned	2 (20.0)	14 (56.0)	0.004	13 (54.2)	3 (27.3)	0.148	13 (46.4)	3 (42.9)	0.089	5.5 (4-7.5)	0.440	
mentioned/incorrect	1 (10.0)	8 (32.0)		4 (16.7)	5 (45.4)		9 (32.1)	0 (0.0)		3 (2-7)		
mentioned/correct	7 (70.0)	3 (12.0)		7 (29.2)	3 (27.3)		6 (21.4)	4 (57.1)		6 (5-7)		
Access on screening												
not mentioned	5 (50.0)	16 (64.0)	0.322	12 (50.0)	9 (81.8)	0.155	19 (67.9)	2 (28.6)	0.038	6 (5-8)	0.203	
mentioned/incorrect	0 (0.0)	3 (12.0)		2 (8.3)	1 (9.1)		3 (10.7)	0 (0.0)		3 (1-6)		
mentioned/correct	5 (50.0)	6 (24.0)		10 (41.7)	1 (9.1)		6 (21.4)	5 (71.4)		5 (4-6)		
Readability	median (P25-P75)	69.5 (68.5-73.5)	73.7 (69.8-76.8)	0.177	70.2 (68.2-74.6)	75.7 (73.7-77.4)	0.036	73.7 (69.5-76.9)	68.9 (67.8-73.5)	0.058	-0.093	0.203

* The information on potential harms associated with radiation exposure are not being considered to ensure comparability with the analysis referring to prostate cancer
 not mentioned – none of the items that are included in the topic was mentioned; mentioned/incorrect – at least one items that is included in the topic was mentioned incorrectly, and none was mentioned correctly;
 mentioned/correct – at least one items that is included in the topic was mentioned correctly

Table 18 – Quality of the contents on prostate cancer screening according to websites' order of appearance, country of origin, specific for cancer and ISPUP score.

	Contents on prostate cancer screening										ISPUP score median (P25-P75)	P	
	Order of appearance in search			Country of origin			Cancer-specific website						
	<=30 n (%)	>=30 n (%)	P	Portuguese n (%)	Other n (%)	P	No n (%)	Yes n (%)	P				
Screening methods													
not mentioned	0 (0.0)	2 (6.9)	0.811	1 (5.0)	1 (4.4)	0.801	2 (5.9)	0 (0.0)	0.223	2.5 (0-5)	0.021		
mentioned/incorrect	13 (92.9)	24 (82.8)		18 (90.0)	19 (82.6)		30 (88.2)	7 (77.8)		7 (5-8)			
mentioned/correct	1 (7.1)	3 (10.3)		1 (5.0)	3 (13.0)		2 (5.9)	2 (22.2)		3.5 (1.5-4)			
Harms													
not mentioned	11 (78.6)	24 (82.8)	<0.001	15 (75.0)	20 (87.0)	0.869	28 (82.4)	7 (77.8)	0.571	6 (4-8)	0.312		
mentioned/incorrect	1 (7.1)	1 (3.4)		1 (5.0)	1 (4.4)		1 (2.9)	1 (11.1)		2.5 (0-5)			
mentioned/correct	2 (14.3)	4 (13.8)		4 (20.0)	2 (8.7)		5 (14.7)	1 (11.1)		6 (1-7)			
Eligibility													
not mentioned	10 (71.4)	16 (55.2)	0.666	13 (65.0)	13 (56.5)	0.871	19 (55.9)	7 (77.8)	0.559	5 (4-8)	0.191		
mentioned/incorrect	4 (28.6)	12 (41.4)		7 (35.0)	9 (39.1)		14 (41.2)	2 (22.2)		7 (5-7.5)			
mentioned/correct	0 (0.0)	1 (3.4)		0 (0.0)	1 (4.4)		1 (2.9)	0 (0.0)		3 (3-3)			
Periodicity													
not mentioned	13 (92.9)	23 (79.3)	0.396	17 (85.0)	19 (82.6)	>0.999	27 (79.4)	9 (100.0)	0.314	6.5 (4-8)	0.148		
mentioned/incorrect	1 (7.1)	6 (20.7)		3 (15.0)	4 (17.4)		7 (20.6)	0 (0.0)		4 (3-7)			
mentioned/correct	0 (0.0)	0 (0.0)		0 (0.0)	0 (0.0)		0 (0.0)	0 (0.0)		---			
Access on screening													
not mentioned	14 (100.0)	29 (100.0)	---	20 (100.0)	23 (100.0)	---	34 (100.0)	9 (100.0)	---	6 (4-8)	---		
mentioned/incorrect	0 (0.0)	0 (0.0)		0 (0.0)	0 (0.0)		0 (0.0)	0 (0.0)		---			
mentioned/correct	0 (0.0)	0 (0.0)		0 (0.0)	0 (0.0)		0 (0.0)	0 (0.0)		---			
Readability													
	median (P25-P75)	71.8 (69.4-72.7)	67.1 (63.5-77.8)	0.551	70.4 (64.2-75.6)	67.1 (63.5-77.8)	0.559	68.3 (63.3-77.8)	71.7 (69.4-72.7)	0.698	-0.234	0.131	

* The information on potential harms associated with radiation exposure are not being considered to ensure comparability with the analysis referring to prostate cancer not mentioned – none of the items that are included in the topic was mentioned; mentioned/incorrect – at least one items that is included in the topic was mentioned incorrectly, and none was mentioned correctly;

7. Discussion

The websites analyzed had unsatisfactory formal quality, regardless the instrument used for the assessment. Although most of them addressed topics related with cancer screening, the information available was often incomplete or inaccurate.

Most websites identified the screening methods but the possible harms of the screening were frequently overlooked. Despite the poor quality on the contents presented, the websites obtained good scores on readability. The websites addressed for breast or prostate cancers were relatively homogeneous regarding formal quality but differed on the quality of the contents on screening. There were no structural significant or meaningful associations between formal quality and the quality of the contents on screening.

The present dissertation was an attempt to describe the assessment of the websites' quality with the necessary detail to ensure the transparency of the process and its replication for others and with the context of monitoring the information made available in the Internet. It provides a framework for analysis that can be used by other researchers, to study the quality of the websites that provide information on cancer and cancer screening and, with the necessary adaptation, may be used to address the information on other health topics.

This study, however, presents limitations that need to be addressed. The number of websites selected for analysis is relatively small, which contributes to imprecise estimates of the quality indicators as well as limited power for the assessment of the potential determinants of the website quality. However, we were attempting to replicate searches conducted by a layperson looking for general information on cancer. Under this scenario, the small number of studies is probably an unavoidable limitation, given the need to use relatively simple and unspecific search expressions and the expectation that most people are no willing to screen a large number of references to obtain the needed information. Although the number of websites could have been increased with broader inclusion criteria or the use of more specific search expressions, this would increase the heterogeneity of the characteristics of the websites or the results would not correspond to those obtained by a layperson that is expected to use simple search expressions, at least in the initial approaches to the topic. Words such as "cancro", "mama", or "prostata", could not easily be replaced by laypersons, and as consequence, there would not exist many different unspecific and simpler expressions. The analysis of a more than 200 search results for each cancer would result in a more comprehensive assessment of the available information in the internet, but most laypersons are not expected to screen a large number of URLs in their searches. It has been shown that people do not usually go beyond the first three pages when they search for information on the web.^(106, 135) Jansen and Pinks (2002), in an analysis of the

process of searching for information on the web, concluded that 54% of search users only look for information in the first web page results, 19% go the second web page of results and 27% go beyond, until the third web page of results.⁽¹³⁵⁾ Our results demonstrated that in the first 30 results, the probability to find a cancer specific website, as well as a specific site for breast or prostate cancer, is higher. This means that people tend to find websites which information is directed towards cancer or the specific type of cancer searched, meaning best quality of the information and less unnecessary contents.

Another limitation of our study was that the collection of information from the websites was made by only one investigator. However, the impact of this limitation was probably minor, as we made an important effort to standardize the evaluation of the websites, to make the assessment of their quality as replicable as possible. For the analysis of the contents, we created the quality criteria "a priori", according to credited institutions, which is also expected to minimize interobserver variation. Furthermore, the researcher that evaluated the websites was a health professional with experience in oncology and doubts were discussed until consensus with a co-investigator with experience in cancer epidemiology.

We used the Index of Fernandez-Huerta to evaluate readability. It has been used in previous works on Portuguese language, to assess the quality of Brazilian governmental websites,⁽¹⁰⁹⁾ but it is not validated to the Portuguese language. Although the comparison across websites is still possible, further work is needed to establish the correspondence between the score attributed to the website and the education level needed to understand the information (according to the Portuguese curricula).⁽¹¹³⁾ The results suggest that the websites present good levels of readability. We considered 70 as a good value of the readability, once it is near the easiest level of readability (100), and it has already been accepted as a good level of readability, in Portuguese texts, when using the Index Fernandez-Huerta.⁽¹⁰⁹⁾ As far as we know, there are not similar works on health related aspects that tried to assess readability in websites in Portuguese language, which unable a deeper discussion of our results.

Besides the quality of the contents provided, namely on screening, our study also pretended to analyze the formal quality of the websites, according to instruments specially created for the purpose.

We chose the Discern tool, the Health On Net code and the AMA guidelines for medical and health information sites because they had been referred in the majority of the studies related to the assessment of the quality of websites. Our choice on the eEurope guidelines for health related websites was made regarding its specificity for the European context, and its concern in making the guidelines similar between the different countries of the European Union.

In the majority of the websites, the results on the Discern tool were low for most criteria, like in other works.⁽¹³⁶⁾ A previous study that used the Discern tool to evaluate the quality of information on the Internet in pediatric neuro-oncology, reached the same low scores, with exception to the criteria related with the date of publication of the information, where the results tended to be slightly better.⁽¹⁰⁶⁾

Less than 10% of the sites analyzed display the HON seal, a value slightly lower than the one found in a previous similar research of the quality of websites related to breast cancer,⁽⁶⁶⁾ but higher in relation to another conducted in information on the Internet related to acute myocardial infarction and stroke in Portuguese Language.⁽¹³⁶⁾ The reasons for this low percentage of websites that display the HON seal, could be not only a poor quality of some websites that would not accomplish the principles required by the organization, but also the fact that some websites which could possibly have the necessary quality to display the HON seal, did not apply for that certification by the HON Foundation.

Regarding the assessment with the eEurope guidelines for health related websites, less than half accomplish each of the items, with exception for accountability, where the majority did accomplish. Like in a previous assessment of Swedish breast cancer websites (2004), the overall results showed low compliance with the guidelines. Nevertheless, in that study the websites tend to perform better in the criteria related to currency, which is not in accordance with our results.

The evaluation using the AMA guidelines for medical and health information sites tended to demonstrate that in general, the websites analyzes did not agree with the principles of the instrument, like it was found in previous studies.^(66, 136, 137) However, the assessment suggest some agreement in structural aspects like intrasite navigation, functional intersite links, funding and sponsorship, as found in other researches.⁽¹³⁷⁾

As we were conducting the formal evaluation of the websites we monitored the time that we spent on the formal evaluation of the website, according to each instrument, like it was done by Hsu and Bath (2008) in a previous work.⁽¹³⁸⁾ In the referred research, data is not available related to the time spent according to the order of use of the instrument, or even if that aspect was taken into account. In the study, HON code took on average 8.24 minutes to be applied and the Discern tool 7.36 minutes.⁽¹³⁸⁾ These values are higher than the results presented in our research, once the time spent in the application of the HON code or the Discern tool when they were used in the first place were approximately 5 minutes, taking more time to apply the Discern tool, but with a very small difference. This can be explained with some training that investigator presented in our study, before monitoring the time spent on the application of the instrument. This can also be explained if we take into account that results tend to demonstrate that the first websites needed more time to be evaluated, which is coincident with those assessed with less training with the use of the instruments. We

monitored the time spent when each instrument was used at first, second, third or fourth place, because the knowledge of the website would become easier its evaluation. The instrument used in fourth place would benefit from information that the researcher already found in previous evaluations using different instruments, which would decrease the time needed to its application. Our results tend to confirm this hypothesis, once each instrument needed less time to be applied, as they were used as the last. The instruments applied in the assessment share some quality criteria, which means that information collected for some criterion with the application of an instrument, will be readily found when using a different instrument, on the same website, decreasing the time needed to be applied. The AMA guidelines for medical and health related sites was the instrument that took more time to apply, while the eEurope guidelines for health related websites was the one that took less time. This can be explained by the extension of each instrument, once the number of criteria, 12 vs. 15, is different, and a longer instrument would need more time for its application. We hypothesized that this could be a reason to explain why the AMA instrument took more time in the assessment of the websites, when compared to the eEurope guidelines.

We found some similarities between the instruments. These similarities were higher between the HON code and the eEurope guidelines for health related websites, according to the evaluation criteria and structure of the instrument. We elaborated a different instrument, the ISPUP-score based on the HON code and the eEurope guidelines for health related websites that would contain different aspects of each one, and that have the advantage of being easier to use. Although we did not validate the ISPUP-score, and this is an objective to be pursued in future research, the websites having the HON Seal had higher scores in the ISPUP-score, and had a correlation with the other two scores developed as part of our study. The latter, however, is not surprising given the fact that the ISPUP score was based in the other two instruments and there is a partial overlap with each of them. The need for validation applies also to the other instruments, once few studies have tested their reliability and validity.^(102, 138)

As conducted in previous researches our evaluation of the quality of the websites did not rely only in the formal quality.^(106, 136, 139) We also considered important to assess the accuracy or correctness of information, as done in other researches that evaluated the quality of websites.^(49, 83, 104-106, 136, 139-141) Some websites could have good performances in relation to their formal or structural characteristics, but not present good and accurate information.^(138, 142) Regarding the quality of the contents, our study evaluated specifically the screening. Other works, which aimed to assess the overall quality of the contents of websites, relied their assessment in a wider range of aspects, according to the specific subject.^(57, 83, 104, 105, 136, 139-141) The correctness of information was based on the best evidence that we access (available in the US Preventive Services Task Force, 2009) and according to

the actual European and Portuguese guidelines. Once the websites selected were in Portuguese language, it was expected that the contents should be in accordance with the guidelines of their countries of origin, namely Portugal or Brazil. Results tend to demonstrate that websites which country of origin was Portugal tend to provide better information on screening. This can be explained because the Portuguese setting is different from the Brazilian, and we considered the correctness of information according to the Portuguese setting.

The results suggest that the information on screening tend to be better in the female breast cancer websites than in prostate related ones, namely in the screening methods. In the setting of our study, we hypothesized that this fact must be explained with the existence of organized screening for breast cancer,⁽¹³³⁾ while the effectiveness of the prostate cancer screening remains controversial.⁽⁷⁸⁾ In relation to the quality of the contents, and similarly with other works, the analysis tend to demonstrate that websites appearing in the first 30 results (with lower page rank) tend to provide better information,⁽⁶⁶⁾ which can be explained with a more specificity of the website for breast or prostate cancer issues (as shown by our results). Also, better websites, tend be more linked or referred by other websites, which increase their importance and as a consequence decreased their page rank, placing them in the first places of the search results.

Related to aspects of external validity, these results can only be generalized to the Portuguese setting, taking into account the moment of the selection of the websites.

We selected the Google® search engine because it was our intention to simulate the search that a layperson would make on information related with female breast and prostate cancers, when selecting the Internet as a source of information. We also choose this search engine because Google® is the most used search engine among Portuguese population⁽⁶⁴⁾ (93.6% of Portuguese individuals who access the Internet used it on September 2009, and on September 2011, 3585000 Portuguese unique users, above 4 years of age, living in the Portugal, islands excluded, accessed the Internet in their own households, using the Google® search engine) We don't have reasons to believe that other important websites would be out of the scope of Google®. We believed that the most important results or a major part of all the websites would be available by a Google® search.^(66, 84, 106, 136, 143) Nevertheless, other similar searches were conducted with different search engines, which could retrieved different results,^(66, 144, 145) but we think that at the time Google® did not exist as the most used search engine, collecting most part of the existing websites.^(57, 83, 104-106) Although, other similar works just selected 50,⁽⁸⁴⁾ 100,⁽⁸³⁾ or 30 results, since in the latter it is their opinion that besides 30 results the information is repeated,^(105, 106) we selected the first 200 results for each cancer analysis because we wanted to make a wide search on the information

available, and we it was our opinion that results beyond the first 200 would not contribute with significant variation in the analysis.

We conducted our search in one day for each type of cancer. The results of that search were saved, allowing further analysis. The results can vary frequently, because the page rank is dynamic and the relative “position” or order of appearance could change. Some owners of websites inclusively pay to other individuals or institutions to turn better their page rank, in order to be presented in the first results and be more accessed, which does not reflect their quality at all.⁽¹⁴⁶⁾ As consequence, our evaluation regards the search conducted at that time. Nevertheless, we do not have reasons to believe that significant changes could happen that would interfere in the final results presented.

This work demonstrates that further investigation must be done in the field of web information. In general, there is not an instrument that could assess the overall quality of the websites by itself. Further efforts must be conducted towards a tool that could be easily used not only by laypersons, but also by health professionals. Regarding the aspects connected with the websites quality in Portuguese language, it must be pointed that as the quality is not good, institutions must conduct research to raise the overall quality of their websites. Patients and general population are very “exposed” to the possible harms of misleading information on the Internet, and health professionals have a crucial importance to educate patients towards the process of searching for online information. The quality of the contents available on the Internet, that is easily accessed by people in general and patients in particular, must be a concern for all health professionals, once it can empower patients with inaccurate or misleading information that interferes in major and important decisions.⁽¹⁴⁴⁾

8. Conclusion

The websites analyzed had unsatisfactory formal quality, regardless the instrument used for the assessment. Although most of them addressed topics related with cancer screening, the information available was often incomplete or inaccurate.

The websites addressed for breast or prostate cancers were relatively homogeneous regarding formal quality but differed on the quality of the contents on screening, once breast cancer related websites tend to provide more accurate information than prostate cancer related websites. Most websites identified the screening methods but the possible harms of the screening were frequently neglected.

There were no structural significant or meaningful associations between formal quality and the quality of the contents on screening. Despite the poor quality on the contents presented, the websites obtained good scores on readability.

In conclusion, there is a large margin for improving the quality of the websites providing information on breast and prostate cancer in Portuguese language.

There were no significant or meaningful associations between formal quality and the quality of the contents on screening, and the present work provides a framework for the standardized assessment of different dimensions for the quality of websites providing information on breast or prostate cancer, that may be used for the monitoring of the quality of the health information provided in the Internet.

9. References

1. Kickbusch I. Health literacy: addressing the health and education divide. *Health Promotion International* 2001;16(3):289-97.
2. Kwan B, Frankish J, Rootman I. The development and validation of measures of "health literacy" in different populations. Vancouver: University of British Columbia Institute of Health Promotion Research & University of Victoria Centre for Community Health Promotion Research 2006.
3. Kirsch I, Jungeblut A, Jenkins L, Kolstad A. *Adult Literacy in America: A First Look at the Findings of the National Adult Literacy Survey*. In: Education USD, editor. Washington DC 2002.
4. Rudd R. Health Literacy Skills of U.S. Adults. *American Journal of Health Behaviour* 2007;31:8-18.
5. Speros C. Health literacy: concept analysis. *Journal of Advanced Nursing* 2004;50(6):633-40.
6. Baker D, Wolf M, Feinglass J, Thompson J, Gazmamarian J, Huang J. Health Literacy and Mortality Among Elderly Persons. *Archives of Internal Medicine* 2007;167(14):1503-9.
7. Berkman N, DeWalt D, Pignone M, Sheridan S, Lohr K, Lux L, et al. *Literacy and Health Outcomes*. Rockville: Agency for Healthcare Research and Quality 2004.
8. Kanj M, Mitic W. *Health Literacy and Health Promotion: Definitions, Concepts and Examples in the Eastern Mediterranean Region - Individual Empowerment Conference Working Document*. World Health Organization 2009.
9. Golbeck A, Ahlers-Schmidt C, Paschal A, Dismuke S. A definition and operational framework for health numeracy. *American Journal of Preventive Medicine* 2005;29:375-6.
10. Norman C, Skinner H. eHEALS: The eHealth Literacy Scale. *Journal of Medical Internet Research* 2006;8.
11. Nelson W, Reyna V, Fagerlin A, Lipkus I, Peters E. Clinical Implications of Numeracy: Theory and Practice. *Annals of Behavioral Medicine* 2007.
12. Nutbeam D. Defining and measuring health literacy: what can we learn from literacy studies? *International Journal of Public Health* 2009;54:303-5.
13. Williams M, Baker D, Honig E, Lee T, Nowlan A. Inadequate Literacy is a Barrier to Asthma Knowledge and Self-Care. *CHEST Journal* 1998;114:1008-15.
14. Williams M, Baker D, Parker R, Nurss J. Relationship of Functional Health Literacy to Patients' Knowledge of Their Chronic Disease: A Study of Patients With Hypertension and Diabetes. *Archives of Internal Medicine* 1998;158:166-72.
15. Shieh C, Halstead J. Understanding the Impact of Health Literacy on Women's Health. *JOGNN* 2009;38:601-12.
16. Adams R, Appleton S, Hill C, Dodd M, Findlay C, Wilson D. Risk associated with low functional health literacy in Australian population. *MJA* 2009;191(10):530-4.
17. Pearlman D, Clark M, Rakowski W, Beverly E. Screening for breast and cervical cancers: The importance of knowledge and perceived cancer survivability. *Women & Health* 1999;28(4):93-112.
18. Hahn E, Cella D. Health Outcomes Assessment in Vulnerable Populations: Measurements Challenges and Recommendations Archives Physical Medical Rehabilitation 2003;84:35-42.
19. Ngoh L. Health literacy: A barrier to pharmacist-patient communication and medication adherence. *Pharmacy Today* 2009;15(8):45-57.
20. Cavanaugh K, Wingard R, Hakim R, Eden S, Shintani A, Wallston K, et al. Low health literacy associates with increased mortality in ESRD. *Journal of American Society of Nephrology* 2010;21(11):1979-85.
21. King A. Heart failure: Poor health literacy increases mortality. *Nature Reviews Cardiology* 2011;8(363).
22. Davis T, Williams M, E. M, Parker R, Glass J. Health literacy and cancer communication. *CA Cancer Journal Clin* 2002;52:134-49.

23. Bennett C, Ferreira M, Davis T, Kaplan J, Weinberger M, Kuzel T, et al. Relation between literacy, race, and stage of prevention among low-income patients with prostate cancer. *Journal of Clinical Oncology* 1998;16:3101-4.
24. Canadian Public Health Association. *The Impact of Low Health Literacy on Chronic Disease Prevention and Control*: Canadian Public Health Association2006.
25. Guerra G, M. K, Shea J. Literacy and knowledge, attitudes and behaviour about mammography in Latinas. *Journal of Health Care for the Poor and the Underserved* 2005;16:152-66.
26. Garbers S, Chiasson M. Inadequate functional health literacy in Spanish as a barrier to cervical cancer screenings among immigrants Latinas in New York City. *Preventing Chronic Disease* 2004;1.
27. Shohet L, Renaud L. Critical Analysis on Best Practices in Health Literacy. *Canadian Journal of Public Health*2006;97:10-4.
28. Guerra G, Dominguez F, Shea J. Literacy and Knowledge, Attitudes and Behaviour About Colorectal Cancer Screening. *Health Communication* 2007:651-63.
29. Michielutte R, Alciati M, El-Arculli R. Cancer control research and literacy. *Journal of Health Care for the Poor and the Underserved* 1999;10:281-97.
30. Tessler L, Basu A, Leitsch S. Health literacy as a predictor of follow-up after an abnormal pap smear: a prospective study. *Jounal of General Internal Medicine* 2006;52 134-49.
31. Health Promotion: Health Literacy and Health Behaviour. 2011 [cited 2011 October, 2011]; Available from: http://www.who.int/topics/health_education/en/.
32. Rubinelli S, Schulz P, Nakamoto K. Health literacy beyond knowledge and behaviour: letting the patient be a patient. *International Journal of Public Health* 2009;54:307-11.
33. World Health Organization. *Health Promotion Glossary*. Geneva: World Health Organization 1998.
34. Bass P, Wilson J, Griggith C. A Shortened Instrument for Literacy Screening. *Journal of General Internal Medicine* 2003;18:1036-8.
35. Diviani N, Schulz P. What should laypersons know about cancer? Towards an operational definition on cancer literacy. *Patient Education and Counseling* 2010.
36. Galloway S. Measuring information needs of women with breast cancer. *Canadian Oncology Nursing* 1994:4-29.
37. Jenkins V, Fallowfield L, Saul J. Information needs of patients with cancer: results from a large study in UK cancer centers. *British Journal of Cancer* 2000;84(1):48-51.
38. Fallowfield L, Ford S, Lewis S. No news is not good news: Information preferences of patients with cancer. *Psycho-oncology* 1994;4:197-202.
39. Degner L, Kristjanson L, Bowman D, Sloan J, Carriere K, O'neil J, et al. Information Needs and Decisional Preference in Woman With Breast Cancer. *JAMA*1997;277(18):1485-92.
40. Cassileth B, Zupkis R, Sutton-Smith K. Information and participation preferences among cancer patients. *Annals of Internal Medicine*1980;92:832-6.
41. Meredith C, Symonds P, Webster L, al. e. Information needs of cancer patients in West Scotland; cross sectional survey of patients' views. *BMJ*1996;313:724-6.
42. Smith-McAllen., Aaron., Fishbein., Martin., Hornik., Robert C. Psychological Determinants of Cancer-Related Information Seeking Among Cancer Patients. *Journal of Health Communication* 2011;16(2):212-25.
43. Cegala D, Bahnsen R, Clinton S, David P, Gong M, Monk J, et al. Information Seeking and Satisfaction With Physician-Patient Communication Among Prostate Cancer Survivors. *Health Communication* 2008;23(1):62-9.
44. Rees C, Sheard C, Echlin K. The relationship between the information-seeking behaviours and information needs of partners of men with prostate cancer: a pilot study. *Patient Education and Counseling* 2001;49:257-61.
45. Leydon G, Boulton M, Moynihan C, Jones A, Mossman J, Boudioni M, et al. Cancer patients' information needs and information seeking behaviour: in depth interview study. *BMJ* 2000;320:909-13.

46. Feldman-Stewart D, Brundage M, Hayter C, Groome P, Nickel J, Downes H, et al. What Questions Do Patients with Curable Prostate Cancer Want Answered? *Medical Decision Making* 2000;20(7):7-19.
47. Eheman C, Berkowitz Z, Lee J, Mohile S, Purnell J, Rodriguez E, et al. Information-Seeking Styles Among Cancer Patients Before and After Treatment by Demographics and Use on Information Services. *Journal of Health Communication* 2009;14:487-502.
48. Rutten L, Arora N, Bakos A, Aziz N, Rowland J. Information needs and sources of information among cancer patients: a systematic review of search (1980-2003). *Patient Education and Counseling* 2004;57:250-61.
49. Humphris G, Duncalf M, SHolt D, al. e. The experimental evaluation of an oral cancer information leaflet. *Oral Oncology* 1999;35:375-82.
50. Humphris G, Ireland R, Field E. Information on oral cancer encourages primary-care patients to accept oral cancer screenings and reduces associated anxiety *Oral Oncology* 2001;37:548-52.
51. Hack T, Degner L, Dyck D. Relationship between preferences for decisional control and illness information among women with breast cancer: a quantitative and qualitative analysis. *Society Scientific Medical* 1994;39:279-89.
52. Expert Advisory Group. A policy framework for commissioning cancer services. London: Department of Health 1995.
53. Fallows D. How women and men use the Internet. Pew Internet and American Life Project 2005.
54. Santana S. Tendências na Utilização da Internet: Para Questões de Saúde e Doença em Portugal. *Acta Médica Portuguesa* 2008;22:5-14.
55. Eysenbach G, Köhler C. Database of adverse events related to the internet has been set up. *BMJ* 2002;324:239.
56. Eysenbach G, Köhler C, Yihune G, Lampe K, Cross P, Brickley D. A Framework for Improving the Quality of Health Information on the World-Wide-Web and Bettering Public (E-)Health: The Medcertain Approach. *MedInfo* 2001:1450-4.
57. Impicciatore P, Pandolfini C, Casella N, Maurizio B. Reliability of health information for the public on the world wide web: systematic survey of advice and managing fever in children at home. *BMJ* 1997;314:1875-81.
58. Santana S, Pereira A. Da Utilização da Internet para Questões de Saúde e Doença em Portugal. *Acta Médica Portuguesa* 2007;20(47-57).
59. Health Topics: 80% of internet users look for health information online [database on the Internet]. California HealthCare Foundation: Pew Internet & American Life Project. 2011 [cited October 10th 2011]. Available from: <http://pewinternet.org/Reports/2011/HealthTopics.aspx>.
60. Instituto Nacional de Estatística. Indicadores Sociais. Lisboa: INE - Instituto Nacional de Estatística 2008.
61. Ferguson T, Kelly W. The Ferguson Report: E-patients prefer e-groups to doctors for ten of twelve aspects of health care. Austin: Consumer Health Informatics and Online Health 1999.
62. Six Great Ways to Search the Web. Investintech.com; [10th November 2011]; Available from: <http://www.investintech.com/content/searchtheweb>.
63. Lawrence S, Giles C. Accessibility of information on the web. *Nature* 1999;400:107-9.
64. Graphic, Visualization, and Usability Center. GVU's 10th WWW User Survey [database on the Internet] 1998 [cited October 2th 2011]. Available from: http://www.gvu.gatech.edu/user_surveys/survey-1998-10/.
65. U.S. Search Engine Rankings. comScore Releases; 2009 [updated September 15th 2011 October 11th 2011]; Available from: http://www.cumscore.com/Press_Events/Press_Releases/2009/12/cumScore_Releases_November_2009_U.S._Search_Engine_Rankings.
66. Meric F, Bernstam E, Mirza N, Hunt K, Ames F, Ross M, et al. Breast Cancer in the world wide web: cross sectional survey of quality of information and popularity of websites. *BMJ* 2002;324:577-81.

67. Top Ten Search Engines. seoConsultants; 2010 [October 11th 2011]; September 8th 2010:[Available from: <http://www.seoconsultants.com>
68. Kiley R. Does the internet harm health? BMJ2002;324:238-9.
69. Harrison-Woermke D, Graydon J. Perceived information needs of breast cancer patients receiving radiation therapy after excisional biopsy and axillary node dissection. Cancer Nursing 1993;16:449-55.
70. Mills M, Davidson R. Cancer Patients' Sources of Information: Use and Quality Issues. Psycho-Oncology 2002;11:371-8.
71. Nagler R, Gray S, Romantan A, Kelly B, DeMichele A, Armstrong K, et al. Differences in information seeking among breast, prostate, and colorectal cancer patients. Patient Education and Counseling 2010;81:54-62.
72. Poll-Franse L, EEnbergen M. Internet use by cancer survivors: current use and future wishes. Support Care Cancer 2008;16:1189-95.
73. Helft P, Hlubocky F, Daugherty K. American Oncologists' Views of Internet Use by Cancer Patients: a Mail Survey of American Society of Clinical Oncology Members. Journal of Clinical Oncology 2003;21:942-7.
74. GLOBOCAN 2008 CANCER FACT SHEET: Prostate Cancer Incidence and Mortality Worldwide in 2008 [database on the Internet]. Section of cancer information 2008 [cited October 14th 2011]. Available from: <http://globocan.iarc.fr/factsheets/cancers/prostate.asp>.
75. GLOBOCAN 2008 CANCER FACT SHEET: Breast Cancer Incidence and Mortality Worldwide in 2008 [database on the Internet]. Section of cancer information. 2008 [cited October 14th 2011]. Available from: <http://globocan.iarc.fr/factsheets/cancers/breast.asp>.
76. Sant M, Allemani C, Santaquilani M, Knijn A, Marchesi F, Capocaccia R, et al. EUROCARE-4. Survival of cancer patients diagnosed in 1995-1999. Results and commentary. European Journal of Cancer 2009;45:931-91.
77. Rutten L, Squiers L, Hesse B. Cancer-Related Information Seeking: Hints from the 2003 Health Information National Trends Survey (HINTS). Journal of Health Communication 2006;11:147-56.
78. U.S. Preventive Services Task Force. Screening for Prostate Cancer. U.S. Preventive Services Task Force Recommendation Statement. . 2008 [October 10th 2011]; Available from: <http://www.uspreventiveservicestaskforce.org/uspstf08/prostate/prostaters.htm>
79. Vickers A, Roobol M, Lija H. Screening for Prostate Cancer: Early Detection or Overdetection? Annual Review of Medicine2011;26.
80. Charnock D. The Discern Handbook: Quality criteria for consumer health information on treatment choices. Oxon: Radcliffe Medical Press Ltd 1998.
81. Charnock D, Shepperd S, Needham G, Grann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. Journal of Epidemiology and Community Health 1999;53:105-11.
82. eEurope 2002: uma sociedade de informação para todos. Bruxelas: European Union2000 June 2000.
83. Biermann J, Golladay G, Greenfield M, Baker L. Evaluation of Cancer Information on the Internet. American Cancer Society1999;2:381-90.
84. Hellawell G, Turner K, Monnier K, Brewster S. Urology and the Internet: an evaluation of Internet use by urology patients and of information available on urological topics. BJU International 2000;86:191-4.
85. HON. The commitment to reliable health and medical information on the Internet: The Health On Net Foundation 2011.
86. Barrett S, Chairman B. What Does the HON code Mean? Allentown2000 [11th November 2011]; Available from: <http://chealth.canoe.ca/columns.asp?columnistid=3&articleid=2868>.
87. Brendler B. Why the HON Code is Worthless. 2007 [11th November 2011]; Available from: http://blog.consumerwebwatch.org/2007/07/why_the_hon_code_is_worthless.html.
88. Guest J. Building Trust on the Web. 2003 [11th November 2011]; Available from: <http://www.consumerwebwatch.org/printable.cfm?ID=10218&at=1>.

89. The commitment to reliable health and medical information on the Internet [database on the Internet]. The Health On Net Foundation 2011 [cited 2011.08.15]. Available from: <http://www.hon.ch/HONcode/Webmasters/Visitor/visitor.html>.
90. eEurope 2002: Quality Criteria for Health Related Websites [database on the Internet]2002 [cited 2002, 08].
91. Margaret A. Winker ea. Guidelines for Medical and Health Information Site on the Internet: Principles Governing AMA Web Sites. JAMA 2000;283:1600-6.
92. Winker M, Flanigin A, Chi-Lum B, White J, Andrews K, Kennett R, et al. Guidelines for Medical and Health Information Sites on the Internet: Principles Governing AMA Web Sites. Journal of American Medical Association 2000;283:1600-6.
93. Lewieki E, Rudolph L, Kiebzak G, Chavez J, Thorpe B. Assessment of osteoporosis-website quality. Osteoporos Int 2006;17:741-52.
94. Frické M, Fallis D, Jones M, Luszko G. Consumer health information on the Internet about carpal tunnel syndrome: Indicators of accuracy. The American Journal of Medicine 2005;118(168-174).
95. Ferreira L. A Internet como fonte de informação sobre saúde: um levantamento de percepções dos médicos portugueses: Instituto Superior de Ciências do Trabalho e da Empresa; 2006.
96. Kunst H, Groot D, Latthe P, Latthe M, Khan K. Accuracy of information on apparently credible websites: survey of five common health topics. BMJ2002;324:581-2.
97. Cohen L, Jacobson T. Evaluating Web Content. University Libraries; 2009 [13th November 2011]; Available from: <http://library.albany.edu/usered/eval/evalweb/>.
98. Evaluation Checklist. Midland Technical College; 2005 [13th November 2011]; Available from: <http://www.lib.midlandstech.edu/?Evaluation%20Checklist>.
99. How to evaluate website content. Edinburgh: University of Edinburgh; 2011 [13th November 2011]; Available from: <http://www.ed.ac.uk/schools-departments/information-services/services/library-museum-gallery/finding-resources/library-databases/databases-overview/evaluating-websites>.
100. The CARS Checklist (Credibility, Accuracy, Reasonableness, Support). [08th November 2011]; Available from: http://novella.mhhe.com/sites/0079876543/student_view0/research_center_999/research_papers30/conducting_web-based_research.html.
101. Hong T. Contributing factors to the Use of Health Related Websites. Journal of Health Communication 2006;11:149-65.
102. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the World Wide Web? Quantitative study using focus group, usability tests, and in-depth interviews. BMJ2002;324:573-7.
103. Currò V, Buonuomo P, De Rose P, Onesimo R, Vituzzi A, D'Atri A. The evolution of Web-based medical information on sore throat: a longitudinal study. Journal of Medical Internet Research2003;5(2).
104. Abbott V. Web page quality: can we measure it and what do we find? A report of exploratory findings. Journal of Public Health Medicine 2000;22(2):191-7.
105. Bichakjian C, Schwartz J, Wang T, Hall J, Johnson T, Biermann J. Melanoma Information on the Internet: Often Incomplete - A Public Health Opportunity? . Journal of Clinical Oncology2002;20(1):134-41.
106. Hargrave D, Hargrave U, Bouffet E. Quality of health information on the Internet in pediatric neuro-oncology. Neuro-Oncology 2006;175-82.
107. Purcell G, Wilson P, Delamothe T. The quality of health information on the internet: as for any other medium it varies widely: regulation is not the answer. . BMJ2002;324:557-8.
108. Jones K, Thibault J. Readability Analysis: Assessing and Enhancing the Readability of Holocaust Historical Documents: Providing Accessibility for High School Students Western Washington University2000.
109. Barboza E, Nunes E. A inteligibilidade dos websites governamentais brasileiros e o acesso para usuários com baixo nível de escolaridade. Inclusão Social 2007;2(2):19-33.

110. Kobsa A, Stephanidis C. Adaptable and adaptive information access for all users, including disabled and elderly people. Pittsburgh: Workshop on adaptive hypertext and hypermedia: held in conjunction with Hypertext. ACM Conference on Hypertext & Media.1998.
111. Caldwell B, Cooper M, Vanderheiden LGRG. Web Content Accessibility Guidelines (WCAG) 2.0 2008 [2011.08]; Available from: <http://www.w3.org/TR/2008/REC-WCAG20-20081211/>.
112. Berland G, Elliott M, Morales L, Algazy J, Kravitz R, Broder M, et al. Health Information on the Internet: Accessibility, Quality, and Readability in English and Spanish. *JAMA*2001;285(20):2612-21.
113. Flesch R. How to write plain English. Chapter 2: Let's start with the formula. [March 11th 2011]; Available from: <http://pages.stern.nyu.edu/~wstarbuc/Writing/Flesch.htm>.
114. Mira J, Lilinás G, Tomás O, Pérez-Jover V. Quality of Websites in Spanish Public Hospitals. Escola Valenciana de Estudios para la Salud (EVES) 2004.
115. Martins T, Ghiraldello C, Nunes M, Olivera O. Readability formulas applied to textbooks Brazilian Portuguese. Universidade de São Paulo: Instituto de Ciências Matemáticas de São Carlos 1996;28.
116. Barrio-Cantalejo I, Simón-Lorda P, Melguizo M, Escalona I, Marijuán M, Hernando P. Validation of the INFLESZ scale to evaluate readability of texts aimed at the patient. *An Sist Sanit Navar*2008;31(2):135-52.
117. Mueller L, Reid K, Mueller P. Readability of state-sponsored advance directive forms in the United States: a cross sectional study. *BMC Medical Ethics*2010;11(6).
118. Williamson J, Martin A. Analysis of patient information leaflets provided by a district general hospital by the Flesch and the Flesch-Kincaid method. *The International Journal of Clinical Practice*2010;64(13):1824-31.
119. Mc Laughlin G. SMOG Grading - A New Readability Formula. *Journal of Reading* 1969:639-46.
120. Fitzsimmons P, Michael B, Hulley J, Scott G. A readability assessment of online Parkinson's disease information. *J R Coll Physicians Edinb*2010;40:292-6.
121. Burgos F. Entre sílabas e palavras. *Revista Taquigráfica*1949;106.
122. Shepperd S, Charnock D. Against Internet Exceptionalism. *BMJ*2002;324:556-7.
123. Sociedade da Informação e do Conhecimento: Inquérito à Utilização das Tecnologias da Informação e da Comunicação pelas Famílias [database on the Internet]Instituto Nacional de Estatística 2008 [cited October 9th 2011]. Available from: <http://www.ine.pt>.
124. Sociedade da Informação e do Conhecimento: Inquérito à Utilização de Tecnologias da Informação e da Comunicação pelas Famílias: Indivíduos dos 10 aos 15 anos [database on the Internet]Instituto Nacional de Estatística 2009 [cited October 9th 2011]. Available from: <http://www.ine.pt>.
125. Anderson J, McKemmish S, Manaszewicz R. Quality criteria models used to evaluate health websites. *Asia Pacific Special Health and Law Librarians Conference*2003.
126. Jadad A, Gagliardi A. Rating Health Information on the Internet: Navigating to Knowledge or to Babel? *JAMA*1998;279:611-4.
127. Kim P, Eng T, Deering M, Maxfield A. Published criteria for evaluating health related websites: review. *BMJ*1999;318:647-9.
128. Gagliardi A, Jadad A. Examination of instruments used to rate quality of health information on the internet: chronicle of a voyage with an unclear destination. *BMJ*2002;324:569-73.
129. Cortina J. What is coefficient alpha? An Examination of Theory and Applications. *Journal of applied psychology*1993;78:98-.
130. Hu L, Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* 1996;6:1-55.
131. U.S. Preventive Services Task Force. Screening for Breast Cancer. Recommendation Statement. 2009 [updated December 2009October 10th 2011]; Available from: <http://www.uspreventiveservicestaskforce.org/uspstf09/breastcancer/brcanrs.htm>

132. Recommendations on Cancer Screening in the European Union Vienna: European Union: Advisory Committee on Cancer Prevention 1999.
133. Plano Nacional de Prevenção e Controle das Doenças Oncológicas 2007/2010: Ministério da Saúde; Alto Comissariado da Saúde; Coordenação Nacional Para as Doenças Oncológicas2009.
134. Estatísticas de Texto e Legibilidade. LazarusBrasil.org; [March 21st 2011]; Available from: <http://lazarusbrasil.org/textmeter.php>.
135. Jansen B, Spink A. An Analysis of Document Viewing Patterns of Web Search Engine Users. Pittsburgh: School of Information Sciences University of Pittsburgh 2002.
136. Bastos A. Quality of health information on acute myocardial infarction and stroke in the World Wide Web. Porto: Faculdade de Medicina da Universidade do Porto; 2011.
137. Huang J, Discepola F, Al-Fozan H, Tulando T. Quality of fertility clinics websites. *Fertility and Sterility* 2005;83(3):538-44.
138. Hsu W, Bath P. Development of a patient-oriented tool for evaluating the quality of breast cancer information on the Internet: Health Informatics Research Group. Department of Information Studies. University of Sheffield, UK.2008.
139. Nilsson-Ihrfelt E, Fjällskog M, Blomqvist C, Ahlgren J, Edlund P, Hansen J, et al. Breast cancer on the Internet: the quality of Swedish breast cancer websites. *The Breast* 2004;13:376-82.
140. Griffiths K, Christensen H. Quality of web based information on treatment of depression: cross sectional survey. *BMJ*2000;321:1511-5.
141. Black P, Penson D. Prostate Cancer on the Internet - Information or Misinformation? *The Journal of Urology* 2006;175:1836-42.
142. Martin-Facklam M, Kosrzwa M, Schubert F, Gasse C, Haefeli W. Quality markers of drug information on the internet: an evaluation on sites about St. John's Wort. *The American Journal of Internal Medicine* 2002;113(740-745).
143. Lang J, Collen A. Evaluating Personal Health Care and Health Promotion Web Sites. *Methods Inf Med* 2005;44:328-33.
144. Lopez-Jornet P, Camacho-Alonso F. The quality of internet sites providing information relating to oral cancer. *Oral Oncology* 2009;45(9):95-8.
145. Ni Riodain RM. Head and neck cancer information on the internet: type, accuracy and content. *Oral Oncology* 2009;45(8):675-7.
146. Suarez-Almazor MK, Dorgan M. Surfing the Net - information on the World Wide Web for persons with arthritis: patient empowerment or patient deceit? *Journal of Rheumatology*2001;28(19):185-91.