



HAL
open science

Assessing the internet prevalence of cancer

James Matthew Lloyd Williamson, David Hocken

► **To cite this version:**

James Matthew Lloyd Williamson, David Hocken. Assessing the internet prevalence of cancer. *International Journal of Clinical Practice*, 2011, 65 (2), pp.178. 10.1111/j.1742-1241.2010.02617.x . hal-00608950

HAL Id: hal-00608950

<https://hal.science/hal-00608950>

Submitted on 16 Jul 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Assessing the internet prevalence of cancer

Journal:	<i>International Journal of Clinical Practice</i>
Manuscript ID:	IJCP-09-10-0508
Wiley - Manuscript type:	Original Paper
Date Submitted by the Author:	21-Sep-2010
Complete List of Authors:	Williamson, James; The Great Western Hospital, General Surgery Hocken, David; The Great Western Hospital, General Surgery
Specialty area:	

SCHOLARONE™
Manuscripts

Wiley Only

Title Page

Title: Assessing the internet prevalence of cancer

Authors: James ML Williamson and David B Hocken

Institution: The Great Western Hospital, Marlborough road, Swindon, SN3
6BB, England, UK. Tel: 01793 604020, Fax: 01793 646204.

Corresponding author: James Williamson, c/o Department of Surgery, The
Great Western Hospital. Email: jmw@doctors.org.uk

Author contributions: Both authors contributed equally to the design of this
study, data collection and editing of the paper.

Funding and Acknowledgements: Nil

Conflicts of Interest: Nil

Abstract

Introduction

The internet is one of the most frequently accessed sources of information by patients. There is a variety of cancer related information provided by the internet, aimed at both the general public (from official sites and non-regulated sites) and for health care professionals. Little is known about whether the information provided reflects the prevalence of disease.

Methods

Searches for the 10 most common UK cancers were performed using 5 internet search engines. The number of relevant webpages was recorded and compared to the prevalence of each cancer according to Cancer Research UK.

Results

Of the 985687623 webpages identified, the majority were related to breast cancer (37.2%), followed by lung (16.2%) and prostate cancer (12.5%). Colorectal cancer, Oesophageal cancer and Non-Hodgkin's lymphoma had the least number of websites (4.1%, 0.3% and 0.9% respectively). There was over-representation of breast, kidney and stomach cancer with ratios of prevalence to number of websites of 1.7, 2.6 and 2.5 to 1 respectively. There was under-representation of colorectal cancer, non-Hodgkin's lymphoma and oesophageal cancer (ratios 0.2, 0.2 and 0.1 respectively).

Conclusion

This data highlights the enormity of information available on the internet. However, there is over-representation of certain cancers (and under-representation of others) which may influence how patients perceive their

1
2
3 illness. Healthcare professionals should be aware of the nature of information
4 available on the internet and known when and how to direct patients to
5 reputable sites that provide high quality information.
6
7
8
9

10 11 12 13 14 15 **What's known about this subject?**

16
17 The Internet has a large number of webpages related to cancer, from both
18 official and unofficial sources. Little is known if the number of webpages is
19 representative to the overall incidence of cancer
20
21
22
23
24

25 26 27 **What's new?**

28
29 This paper highlights the fact that certain cancers are over- or under-
30 represented on the internet. This may have an influence on how the general
31 population perceives cancer.
32
33
34
35
36
37

38 39 **Take home message for the clinic**

40
41 Patients are increasingly resourceful in retrieving information. Clinicians must
42 be aware of the limitations of internet derived information and should be able
43 to direct patients to official sources of information.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

The use of the internet to seek health information is common and increasing, particularly among patients with life-threatening or chronic diseases, young adults (18-29) and the elderly (65 years and above) [1-3]. In 2003 more than 90 million people had sought health information online and 20% of the population of America performed health searches in 2006 [1,4]. Of these online health information seekers, 70% report that the internet influenced decisions about their own or a loved one's health care [5] and more than 1/3 of patients ask clinicians about health websites during consultations [6]. These patients believe that internet use enhances the doctor-patient relationship [6-7]. Most health information searches are performed via a online search engine (e.g. Google, Bing, Yahoo, Ask and Lycros)[8].

Internet Search engines provide the user with a list of results based on the search, from what they feel is the most to least relevant. They also provide a total number of webpages pertinent to the search term. With over 29.7 billion webpages in existence (estimate from 2007)[9] and an estimated 82 million hits in 2005 on a search for 'cancer' (from Google) there is bound to be some overlap on cancer information available on the internet [8]. To date no papers have assessed if the number of websites related to cancer is representative to the overall prevalence. However, cancers that are more prevalent, have a high mortality and have a high media profile are associated with an increased level of internet search activity [8].

Method

The five most frequently used internet search engines (Google, Bing, Yahoo, Ask and Lycros) were assessed for the 10 most common UK cancers (table 1)[8,10-11]. The number of websites related to each cancer was recorded from each search engine and combined. The individual numbers of search engines were not considered as important to the overall proportion of cancer results.

Cancer	Prevalence (% population)
Breast	15.6
Lung	13.3
Colorectal	12.8
Prostate	12.1
NHL	3.6
Melanoma	3.6
Bladder	3.5
Kidney	2.7
Oesophagus	2.7
Stomach	2.6

Table 1: The prevalence of each of the most common UK cancers (NHL = Non-Hodgkin's Lymphoma)[10]

Results

A total of 985,687,623 results were retrieved from all of the 5 search engines on all of the malignancies. Breast cancer received the greatest number of hits (367 million, 37.2%), followed by lung (159 million, 16.2%) and Prostate cancer (123 million, 12.5%)(table 2). The least number of websites related to Colorectal 40.7 million (4.1%), NHL (9 million, 0.9%) and Oesophageal cancer (2.5 million, 0.3%).

Cancer	Total Internet Search	% cancer websites on internet
Breast	367,125,752	37.2
Lung	159,475,935	16.2
Colorectal	40,733,143	4.1
Prostate	123,693,147	12.5
NHL	9,337,248	0.9
Melanoma	42,849,155	4.3
Bladder	55,282,822	5.6
Kidney	95,086,343	9.6
Oesophagus	2,506,726	0.3
Stomach	89,597,352	9.1

Table 2: The number of webpages about each type of cancer (combined from the major 5 search engines)

A ratio was calculated between cancer prevalence and number of webpages to give a relative internet profile of each malignancy (figure 1). This revealed that there was over-representation of breast, kidney and stomach cancer with ratios of prevalence to number of websites of 1.7, 2.6 and 2.5 to 1 respectively. There was under-representation of colorectal cancer, non-Hodgkin's lymphoma and oesophageal cancer (ratios 0.2, 0.2 and 0.1 respectively).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

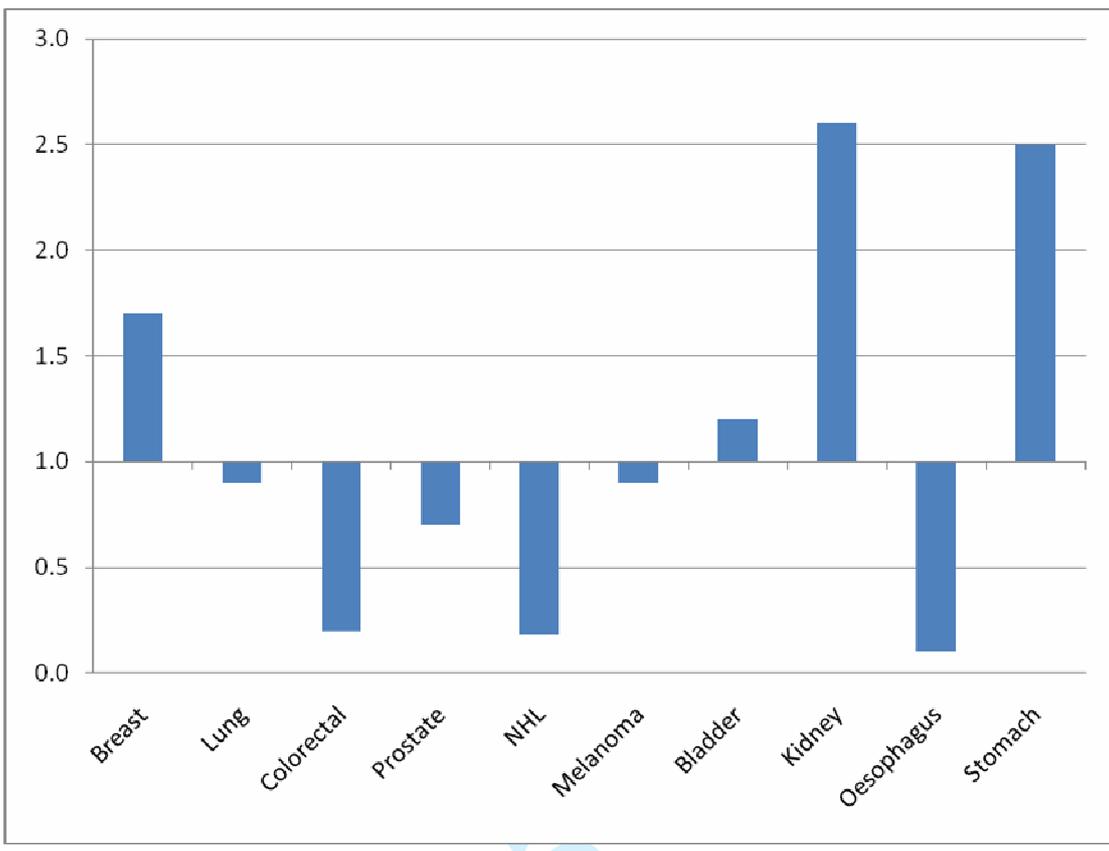


Figure 1: The relative internet profile of cancer

Pre-View Only

Discussion

There is a vast amount of health related material on the world wide web, and the proportion of the population that has access to this is increasing [1-2,4,8]. While the media may be the most important source of public health information, in terms of exposure and opportunistic health education [12-14], the internet enables people to directly search for subjects that are important for them at that time. Searches on the internet are likely to be stimulated by a specific event (e.g. after consultation with a clinician, or when a friend/relative has been diagnosed with an illness)[2,4,6,8]. Media coverage of cancer prompts online searching behaviour (especially in cancer awareness months)[8]. Thus specific search terms are likely to be used by an individual to access information. Patients who use the world wide web for health information are likely to be highly motivated people who are likely to be better educated and informed about their own health when they present to their health care provider [1-2,4,7].

Our data shows that the internet profile of breast, kidney and stomach cancer is over represented. Given the scope of the number of webpages related to each search, the exact content could not be fully assessed. While some of these 'hits' may not be accurate, it is likely that these cancers do have an overall dominance of internet traffic. A possible reason for the dominance of breast cancer is the increased use internet health sites by women [1]. Colorectal cancer, NHL and oesophageal cancer seem to have a lower than to be expected number of webpages associated with them. It is possible that the search term 'colorectal' may prove too specific a term and (some websites

1
2
3 may use the terms colon or rectal cancer) thus there may be an under-
4 recording of websites. However, this term is likely to be used by clinicians in
5 discussion with patients and thus they will perform an internet search with the
6 information given. Oesophagus can be spelt two ways, and thus we may have
7 not accurately recorded all sites that use the American 'esophagus', which
8 may be more common. Two searches were not performed as there may be
9 cross-linkage on the same site and thus duplication of results (link into cancer
10 in media paper and state results similar however).
11
12
13
14
15
16
17
18
19
20
21
22
23
24

25 Perhaps the most alarming feature from the data shows that the number of
26 webpages related to colorectal cancer is far lower than should be expected,
27 especially as this is the second most common cancer in both men and
28 women. One potential reason for this is the taboo associated with its
29 symptoms – rectal bleeding and altered bowel habit may not be the most
30 appealing of subjects to write about. If harnessed properly the internet could
31 be a powerful tool to help break the taboo's of this disease [15-16]. These
32 patient orientated websites can be invaluable to ensure that patients have a
33 point access to information that they can read at their own leisure. They may
34 improve patient comprehension of their disease and give a realistic
35 expectation of therapy. In addition, it may also have links to local support
36 groups or personal accounts of other cancer sufferers, which can help to break
37 down any feelings of isolation and stigma felt by patients[15-16].
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 The unregulated content of the world wide web may result in a distortion of
4 information available to the public. As the internet will have webpages that
5 mention cancer, but are not primarily providing health information, the number
6 of websites that are listed on a websearch is not necessarily representative to
7 overall cancer incidence. Moreover, some of this data will not be suitable for
8 patients and some of these websites will not be from regulated sources [3-
9 4,6,17-18]. Although the majority of health seekers find on-line health
10 information trustworthy, there is concern about the accuracy and quality of
11 internet-derived information [3,17]. Acting upon faulty advice or information
12 found on the internet undoubtedly causes harm [3,17-18]. As sites sponsored or
13 administered by medical societies and universities are associated with
14 reliability, the reputation and authority of these organisations might be
15 undermined if they provide inaccurate [17-18]. Potentially some cancers which
16 have a large amount of internet coverage may be perceived as being more
17 common than they truly are, with the converse likely to be true, with under-
18 reported malignancies thought to be less common. It can be postulated that
19 the general public will have more understanding of diseases which receive a
20 large amount of internet coverage.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Conclusion

The internet is an ideal medium to disseminate health information to the public as it is easily accessible, provides information in an interactive audio-visual format and contains links to other relevant websites. It would be impossible to regulate all of the information provided on the internet and one of the main appeals of the world wide web is that anyone is free to make their own website. This data does show that there are discrepancies between the prevalence of cancers and the amount of webpages in existence. As clinicians we should be aware of this data as patients may think certain malignancies are more (or less) common than they truly are. We should ensure that patients seeking consultation should be provided with a list of website addresses that are regulated sources of information that can enhance patient understanding and comprehension of their illness.

Legends

Figure 1: The relative internet profile of cancer

Table 1: The prevalence of each of the most common UK cancers (NHL = Non-Hodgkin's Lymphoma)[10]

Table 2: The number of webpages about each type of cancer (combined from the major 5 search engines)

References

- 1: Madden M. Internet penetration and impact. Washington, DC: Pew charitable trust, 2006.
- 2: Basch EM, Thaler HT, Shi W, Yakren S, Schrag D. Use of information resources by patients with cancer and their companions. *Cancer* 2004;**100(11)**:2476-2483
- 3: Ekman, A. Hall P, Litton J-E. Can we trust cancer information on the internet? – A comparison of interactive cancer risk sites. *Cancer causes and control* 2005;**16**:765-772
- 4: Fox S, Fallows D. Internet health resources: health searches and email have become more commonplace, but there is room for improvements in searches and overall internet access. Washington, DC: Pew charitable trust, 2003
- 5: Fox S, Rainie L. Vital decisions. How internet users decide what information to trust when they or their loved ones are sick. Washington, DC: Pew charitable trust, 2002
- 6: Murray E, Lo B, Pollack L, et al. The impact of health information on the internet on the physician-patient relationship: patient perceptions. *Archives of Interenal Medicine* 2003;**163**:1727-1734

1
2
3 7: Ferguson T. Digital doctoring – opportunities and challenges in electronic
4 patient-physician communication. *JAMA* 1998;**280**:1361-1362
5
6

7
8
9
10 8: Cooper CP, Mallon KP, Leadbetter S, Pollack LA, Peipins LA. Cancer
11 internet search activity on a major search engine, United States 2001-2003.
12 *Journal of medical internet research* 2005;**7(3)**:e36
13
14

15
16
17
18
19
20 9. how many websites are there? Boutell.com.
21
22 <http://www.boutell.com/newfaq/misc/sizeofweb.html> (accessed August 2010)
23
24

25
26
27 10: Cancer Research UK. Cancer Incidence statistics.
28
29 <http://info.cancerresearchuk.org/cancerstats/index.htm> (accessed May 2010)
30
31

32
33
34 11: Search Engine Guide (www.searchengineguide.com, accessed August
35
36 2010)
37
38

39
40
41 12: Schwitzer G, Mudur G, Henry D, Wilson A, Goozner M, Simbra M, Sweet
42 M, Baverstock KA. What are the roles and responsibilities of the media in
43 disseminating health information. *PLoS Medicine* 2005;**2(7)**:576-582
44
45
46
47

48
49
50
51 13: Williamson JML, Hocken DB. Pancreatic cancer in the media: the Swayze
52 shift. *Annals of the Royal College of Surgeons of England* 2010;**92(6)**:537-8
53
54
55

56
57
58 14: Williamson J, Hocken D. How much vascular disease is reported by the UK
59 media? *International Journal of Clinical Practice* 2010;**64(8)**:1167-8
60

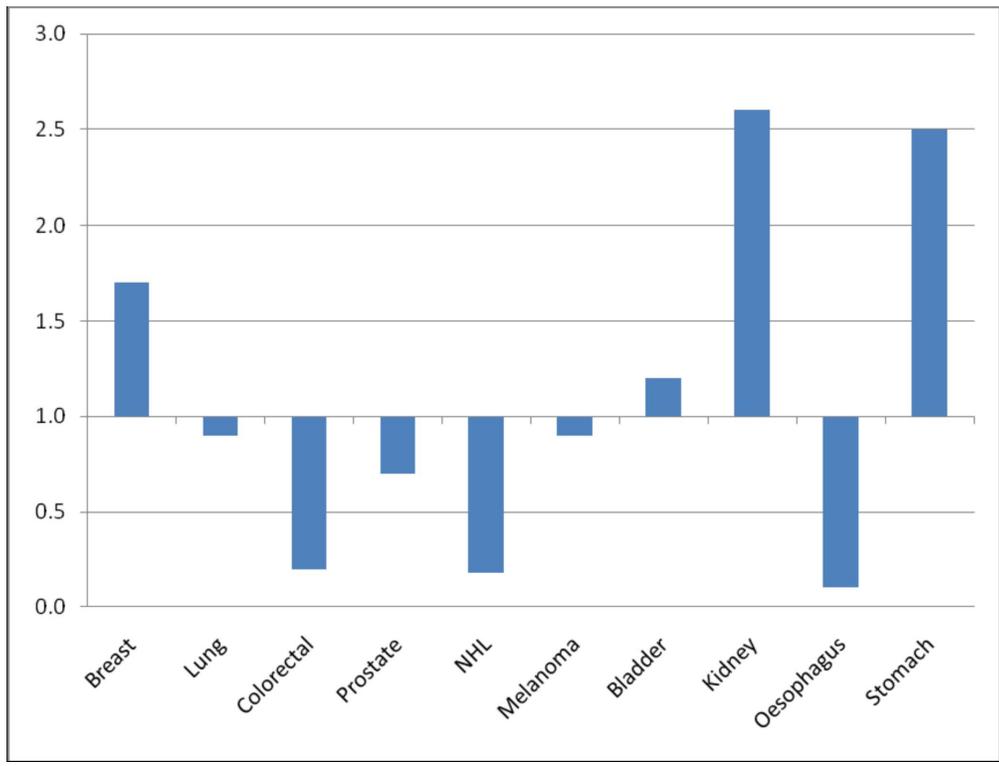
1
2
3
4 15; Muzzin LJ, Anderson NJ, Figueredo AT, Gudelis SO. The Experience of
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Cancer. *Social Science Medicine* 1994;**38(9)**:1201-1208

16: Koller M, Kussman J, Lorenz W, Jenkins M, Voss M, Arens E, Richter E, Rothmund M. Symptom reporting in cancer patients: the role of negative affect and experienced social stigma. *Cancer* 1996;**77(5)**:983-995

17: Berland GK, Elliott MN, Morales, LS et al. Health information on the internet: accessibility, quality, and readability in English and Spanish. *Journal of the American Medical Association* 2001;**285**:2612-2621

18: Biermann JSGGJ, Greenfield MLVH, Baker L. Evaluation of cancer information on the internet. *Cancer* 2000;**86**:381-390

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



The relative internet profile of cancer
175x133mm (300 x 300 DPI)

View Only