

## Web-based Medical Information

R H Baud, L E Cousineau, S Roessingh\*

Medical Knowledge Discovery Project  
Cousineau Consulting LLC  
2806 Fennel Rd, Edgewater MD 21037  
E-mail: [steve.roessingh@gmail.com](mailto:steve.roessingh@gmail.com)

\*Corresponding author

**Abstract:** The study analyzed web-based information on approximately 700 diseases in 4 classes. The objective of the study is to quantify and analyze the amount of web based disease information in order to determine the best ways to search for and present medical knowledge and information on the internet to consumers, patients, physicians, health care professionals, academics, and students. The over 700 diseases researched have an average of 1 Million web page references, 99% have images, 44% have videos, and 27% have sponsored links. The study concluded that web based health and medical information must be presented within a defined patient and medical context, the source of the information must be clearly identified, and the web sites must be ranked according to their knowledge and depth of information about a disease class and disease.

**Keywords:** web based medical information, disease information profile

### 1. Introduction

An enormous amount of health and medical information and knowledge is available on the Internet and added to on a daily basis. This ranges from health advice and information for patients, complete information on medications and drugs, to the latest research for medical specialists. Large libraries of medical images and videos are now available on the Internet. This radically changes how health and medical knowledge is unlocked and shared around the world. The Web Medical Information study quantified per disease for approximately 700 diseases and disorders the scope of web based medical information.

### 2. Research Methodology

For research purposes, *The Merck Manual of Diagnosis and Therapy* was used as a base reference for the "Medical Knowledge Discovery" Study. Approximately 700 diseases and disorders in 5 classes were selected for the research. Emphasis is given to 42 cardiovascular, 206 skin diseases (68 melanoma and 138 skin cancer disorders), 65 diabetes mellitus (endocrine and metabolic), 220 hematology and oncology (including 37 cancer and 23 leukemia disorders), and 154 musculoskeletal disorders.

A "disease information profile" database and statistical analysis was created for 700 diseases in 5 major disease classes (cardiovascular, diseases of the skin, endocrine & metabolic, hematology & oncology, musculoskeletal). For each of these 700 diseases 6 search engine statistics were recorded to create a disease information profile (web page references, average references per disease class, images, videos, sponsored links, rare diseases).

### 3. Scope of web based disease information

The following table summarizes the results of the 687 cardiovascular, disease of the skin, diabetes, hematology and oncology, and musculoskeletal diseases and disorders researched.

Classification	Disorders	References	Average per Disorder	Image	Video	Spons. Link	Rare
Cardiovascular	42	69 307 100	1 650 169	490 718	507	25	0
Disease of the Skin	206	78 952 128	383 263	3 615 566	6 187	9	20
Endocrine and metabolic	65	77 899 300	1 198 451	480 948	763	59	0
Hematology and Oncology	220	359 234 710	1 632 885	3 930 967	4 487	105	5
Musculoskeletal	154	136 845 770	888 609	1 588 669	2 665	183	5
<b>Research Total</b>	<b>687</b>	<b>722 239 008</b>	<b>1 051 294</b>	<b>10 106 868</b>	<b>14 609</b>	<b>381</b>	<b>30</b>
% of disorders				99%	39%	20%	4%

The over 700 diseases researched have an average of 1 Million web page references ranging from over 33 Million reference for hypertension to 22 references for rare cancer disorders. 22 diseases could be identified as “rare information diseases” with relatively few web page references.

Various types of cancer and leukemia have over 30 million web page references. The most common medical disorders such as Hypercholesterolemia, Hypertension, Diabetes Mellitus, Neuropathy, Pancreatitis, Retinopathy, Hyperglycemia, Fibromyalgia, Gout, Neuralgia, and Osteoarthritis each have over 3 million web page references. There is an average of 1 million web references per disorder for the over 700 medical disorders researched. If medical context descriptors are added to the search request this is reduced on the average to 340,000 for medical process descriptors, 155,000 by adding a patient descriptor and 74,000 by adding a weight descriptor.

### 3.1 Cardiovascular and heart disease disorders researched

Over 40 cardiovascular and heart disease disorders are used from 65 sub-classes in the research as per the below table.

- There are on the average 1.6 million web page references per cardiovascular disorder. Web references from the National Library of Medicine, American Heart Association and Medical Associations, Medical Publications and Universities and Medical Schools are the most predominant and relevant.
- 100% of cardiovascular disorders have images; many of them are scans and x-rays that illustrate the disorder.
- 74% of cardiovascular disorders have videos including a large number of videos of live surgical interventions.
- 26% of cardiovascular disorders have Sponsored Links

Cardiovascular Disorders	Sections	Dis-orders	Image	Video	Spons. Link	Web Page References	Rare	Average per Disorder
Aorta disorders and diseases	4	2	7'400	17	1	1'580'000	0	790'000
Arrythias and Conducting Disorders	11	1	7	0	0	1'410'000	0	1'410'000
Arterial Hypertension	2	0	0	0	0	0	0	0
Arteriosclerosis	2		0	0	0	0	0	0
Cardiac disorders	9	10	354'240	336	20	49'289'000	0	4'928'900
Cardiac tumors	1	0	0	0	0	0	0	0
Coronary artery disease	2	2	15'710	8	1	2'780'000	0	1'390'000
Endocarditis	2	0	0	0	0	0	0	0
General cardiovascular disorders	4	4	10'163	28	0	1'731'700	0	432'925
Heart failure and Cardiomyopathies	5	2	19'470	8	1	2'530'000	0	1'265'000
Orthostatic (postural) hypotension	1	17	18'245	29	2	4'442'400	0	261'318
Palpitations	1	0	0	0	0	0	0	0
Pericarditis	1	1	4'220	6	0	1'220'000	0	1'220'000
Peripheral arterial disorders	7	0	0	0	0	0	0	0
Peripheral Venous and lymphatic disorders	7	1	4'560	3	0	1'200'000	0	1'200'000
Syncope	1	1	9'940	42	0	2'270'000	0	2'270'000
Valvular Disorders	9	1	2'970	30	0	854'000	0	854'000
<b>17 Cardiovascular Classes</b>	<b>65</b>	<b>42</b>	<b>446'925</b>	<b>507</b>	<b>25</b>	<b>69'307'100</b>	<b>0</b>	<b>1'650'169</b>
% of disorders			100%	74%	26%		0%	

### 3.2 Diabetes and metabolism disorders researched

65 out of over 90 diabetes and metabolism disorders from 5 classes are researched as per the below table.

- There are on the average 1.2 million web page references per diabetes disorder. The most relevant are from the National Library of Health, Diabetes Associations, Medical Publications, and Medical Schools and academic institutions.
- 100% of diabetes disorders have images, the majority being illustrations.
- 45% of diabetes disorders have videos including a large number of videos on life style, diet, and proactive healthcare.
- 28% of diabetes disorders have Sponsored Links

Diabetes mellitus	Sections	Dis-orders	Image	Video	Spons. Link	Web Page References	Rare	Average per Disorder
Alcoholic ketoacidosis	1	6	6'536	4	0	2'520'800	0	420'133
Diabetes Mellitus	2	43	433'444	741	54	63'197'500	0	1'469'709
Diabetic ketoacidosis (DKA)	3	11	34'696	18	4	10'819'000	0	983'545
Hypoglycemia	4	2	4'646	0	0	1'200'000	0	600'000
Nonketotic hyperosmolar syndrome	5	3	1'626	0	1	162'000	0	54'000
<b>5 Diabetes Classes</b>	<b>5</b>	<b>65</b>	<b>480'948</b>	<b>763</b>	<b>59</b>	<b>77'899'300</b>	<b>0</b>	<b>1'198'451</b>
% of disorders			100%	45%	28%		0%	

### 3.3 Hematology and Oncology disorders researched

220 hematology and oncology disorders from 15 classes are researched. This includes 37 cancer and 23 leukemia disorders.

- There are on the average 1.6 million web page references per disorder. For cancer, the average is over 5 million web references per disorder.
- 99% of hematology and oncology disorders have images with almost 2 million images for cancer.
- 37% of disorders have videos with 2,260 cancer videos out 4,500.
- 23% of disorders have Sponsored Links.

Hematology & Oncology Disorders	Dis-orders	Web Page References	Average per Disorder	Image	Video	Sponsored Link	Rare
Anemias Caused by Deficient Erythropoiesis	25	12'247'500	489'900	74'070	34	3	1
Anemias Caused by Hemolysis	36	17'247'460	479'096	85'464	106	8	2
Cancer	37	196'231'000	5'303'541	1'926'794	2'258	39	0
Coagulation Disorders	11	11'761'300	1'069'209	36'099	127	2	0
Eosinophilic Disorders	5	1'889'000	377'800	5'252	2	3	0
Hemostasis	11	11'914'000	1'083'091	100'258	23	7	0
Histiocytic Syndromes	4	638'500	159'625	3'100	2	2	0
Iron Overload	11	3'438'202	312'564	17'021	18	2	1
Leukemias	23	43'883'800	1'907'991	235'041	928	18	0
Lymphomas	14	34'341'300	2'452'950	179'423	932	10	0
Myeloproliferative Disorders	7	1'677'000	239'571	5'507	2	3	0
Neutropenia and Lymphocytopenia	8	5'191'348	648'919	12'766	3	3	1
Plasma Cell Disorders	10	5'334'800	533'480	1'216'329	34	4	0
Spleen Disorders	5	3'318'000	663'600	10'184	11	1	0
Thrombocytopenia and Platelet Dysfunction	13	10'121'500	778'577	23'659	7	0	0
<b>15 Hematology &amp; Oncology Disorders</b>	<b>220</b>	<b>359'234'710</b>	<b>1'632'885</b>	<b>3'930'967</b>	<b>4'487</b>	<b>105</b>	<b>5</b>
% of disorders				99%	37%	23%	2%

### 3.4 Musculoskeletal disorders researched

Over 150 musculoskeletal disorders in 14 classes are included in the research study. The following observations are made.

- There are on the average 900 thousand web page references per musculoskeletal disorder. The National Library of Medicine, Medical Associations, Arthritis and Rheumatism Associations, Medical Publications, Medical Schools, and specialized Arthritis and Rheumatism Web sites primarily publish musculoskeletal disorder references on the web.
- 100% of musculoskeletal disorders have images; the majority x-rays, photographs of the symptoms, and illustrations.
- 46% of musculoskeletal disorders have videos.
- 32% of musculoskeletal disorders have Sponsored Links
- 5 „rare“ musculoskeletal disorders have less then 200 web references. (Calcific peri-arthritis, Jaccoud's arthritis, Adventitial bursitis, Epiphysitis of the calcaneus, Chondromyxofibroma).

Musculoskeletal Disorders	Section	Dis-orders	Image	Video	Sponsored Link	Rare	Average per Disorder
Joint Disease disorders	2	20	69'468	197	44	1	693'935
Autoimmune rheumatic disorders	7	20	46'178	36	8	1	583'775
Avascular necrosis (AVN)	1	1	3'020	11	0	0	600'000
Bursitis, Tendinitis and Fibromyalgia	3	8	59'605	205	27	0	1'350'500
Crystal-induced Arthritis	3	14	895'240	1'410	20	0	2'533'000
Foot and ankle disorders	14	26	48'850	182	17	2	265'833
Hand disorders	6	16	67'769	121	8	0	539'625
Infection of joints and bones	4	9	39'333	10	2	0	586'889
Joint disorders	5	11	96'221	180	20	0	1'116'727
Neck and back pain	4	7	33'294	103	19	0	756'714
Osteoporosis	1	6	193'590	153	11	0	3'378'333
Paget's disease of bone	1	2	3'506	1	0	0	370'000
Tumors of bones and joints	4	10	30'277	43	7	1	414'150
Vasculitis	1	4	2'318	13	0	0	216'355
<b>14 Musculoskeletal Disorders</b>	<b>56</b>	<b>154</b>	<b>1'588'669</b>	<b>2'665</b>	<b>183</b>	<b>5</b>	<b>888'609</b>
% of disorders			100%	46%	32%	3%	

#### **4. Web based medical information conclusions**

The study found that for most diseases there are millions of web pages, thousands of images, and hundreds of videos. To make sense of this enormous volume of health and medical information and make it useful, the following has to be applied.

- 1) Web based health and medical information must be consulted within a patient and medical context. Separating medical context from content, as is done by current internet search engines, will result in very large quantities of information that are not relevant to the information request of the user.
- 2) The source of web based medical information must be clearly identified. Does the information come from medical libraries, medical associations, health care associations, universities, the health care industry, medical journals, general health and medical web sites?
- 3) Within a source category (e.g. universities, medical associations, medical web sites), the web sites must be ranked as to their knowledge and depth of information on a disease or class of diseases.