

An Audit of Indian Clinical Laboratories' websites: Analysis of Optimization Suitability for Search Engines and Digital Marketing

P. Prakash¹, S.K. Sharma², A. Joshi³, P. Kishor⁴, A. Jahan⁵, R. Gupta⁶, N. Sarin⁷, S. Singh^{8*}

^{1,2,3,4,5,7,8}Department of Pathology, Hindu Rao Hospital, Delhi, India

⁶Division of Cytopathology, ICMR-National Institute of Cancer Prevention and Research, Noida, India

*Corresponding Author: sompal151074@gmail.com, Tel.: +91-11-23905861

Received: 11/Sept/2021, Accepted: 20/Oct/2021, Published: 31/Oct/2021

Abstract— Search engine optimization is an important aspect of the digital world we live in today. An extensive literature review showed only an occasional report of analysis of search engine optimization of medical services' websites. We analyzed 16 websites of clinical laboratories in India for their search engine optimization status through features like secured socket layer, robots.txt and sitemap.xml file, page quality-related tags, page loading speed, link analysis, structured data, and "viewport" tag for mobile responsiveness. Few shortcomings were identified such as thin and duplicate page content in 50% of the websites, lack of structured data in 44.75%, and absence of robots.txt, sitemap.xml files and 'alt' and 'srcset' tags from 18.75%, 25%, 56.25% and 87.5% websites, respectively. Avoiding these short comings may improve organic traffic to the websites beneficial to both the marketer and the consumers.

Keywords—Component, Formatting, Style, Styling.

I. INTRODUCTION

In current times, internet and web resources have become one of the most important primary sources of information for various purposes. The most frequent utilization of internet for finding information is through search engines and Google is the most used search engines [1].

The internet is currently inundated with numerous websites related to each of the variety of contents. For a website to be listed in the first few searches in the result list of search engines, optimization is required for the algorithms used by these search engines (Search Engine Optimization, SEO). This concept of SEO, a digital marketing tool, is rapidly gaining importance in designing of the websites to improve search rankings and connect the marketer with prospective consumers [2,3].

Search engines employ algorithms that analyze the incoming links to (off-page) and the page characteristics (on-page) of each of the websites [4]. The terms used for searching are known as keywords. SEO is a procedure to improve the websites' ranking on the search engines. Some authors have described two main dimensions of SEO viz. internal and external website optimization. Website design, meta tags and keywords were shown to affect the internal website optimization; whereas public domain, social media, and linkage affected the external website optimization [5]. Among the dimensions of internal website optimization, optimal designing of the webpage leads to high rankings of websites for specific search terms [6]. Meta tags or metadata of the webpage (available only in its source code) provide data to search engines and visitors to the website as

well as help in association of output of multiple search engines (meta search engines or search aggregators) to improve the search effectiveness and traffic to the website [6,7]. Usage of appropriate keywords in the title and/or page text also improves webpage visibility [8]. However, keywords stuffing is likely to result in penalty by search engines [9].

It has been found that links from and presence in social media of a website may have important roles to play in the results of search engine [10]. Webpages having more to and from links can be more significant than other webpages for search engines [8].

An extensive search of the available peer-reviewed English literature revealed only an occasional publication analyzing the SEO of medical services' websites [1]. In view of the importance of SEO and paucity of literature, the present study was undertaken to analyze the SEO of clinical laboratories in India to guide the laboratory service providers in optimal utilization of web resources in digital marketing of their respective services

II. RELATED WORK

"Search Engine Optimization: An Analysis of Rhinoplasty Web sites" analyzed the SEO of rhinoplasty web sites to assess the impact of search engine optimized websites on the service utilization of facial surgery. No earlier study of SEO status of clinical laboratories' websites was found.

III. MATERIALS AND METHODS

For the present study, a search was conducted for websites of clinical diagnostic laboratories. Those laboratories which had fully functioning laboratory in at least two Indian cities were included in the study. The websites of 16 Indian clinical laboratories (Table 1) were analysed to assess their SEO status using open-source online tools as well as manually (Table 2). For those features which were applicable to the entire website, all the pages of websites were analysed. On the other hand, for page-related features and image optimisation, the first landing page of the website was assessed.

Table 1. Websites of the clinical laboratories in the study

Name	URL used in this study
Thyrocare Technologies	https://www.thyrocare.com/
Dr LalPathLabs	https://www.lalpathlabs.com/
SRL Limited	https://www.srlworld.com/
Metropolis Healthcare Ltd.	https://www.metropolisindia.com/
Dr Lalchandani lab	https://lalchandanipathlab.com/
Oncquest Labs Ltd	https://www.oncquest.net/
Pathcare Labs Private Ltd	https://www.pathcarelabs.com/
Suraksha Diagnostic	https://www.surakshanet.com/index-m.php
Ecotown Diagnostics	https://www.ecotowndiagnostics.com/
360 Diagnostic & Health Services	https://www.360healthservices.com/
Siemens Healthineers India	https://www.siemens-healthineers.com/en-in/laboratory-diagnostics
Quest Diagnostics	https://www.questdiagnostics.com/home/
Suburban Diagnostics	https://www.suburbandiagnosics.com/
Vijaya Diagnostic Centre Pvt Ltd	https://www.vijayadiagnostic.com/
Lucid Medical Diagnostics	https://www.luciddiagnostics.in/
Medall Healthcare Pvt. Ltd	https://www.medall.in/

Domain check-up: The presence of secure socket layer (SSL) certificate, the ability of website to be served from secure connection, redirection of http to https, and ability to load with and without “www” prefix were assessed manually without the use of any software tool.

Files for search engine: The presence and location of robots.txt and sitemap.xml file was assessed by SEO audit tool of seomator (URL: <https://seomator.com/free-seo-audit-tool>). The robots.txt file was manually accessed, and

its content analysed including the presence of location information of sitemap file.

Page quality check-up: Page quality check-up was also done using the SEO audit tool of seomator. The presence of meta description tag was assessed by reading the source code (html text) of the first landing page of the website being evaluated. Other page quality features included: the length of page title, presence of header tags (H1 through H6), content quantity in terms of word count and content quality in terms of uniqueness. The uniqueness of page content was graded as “Almost Duplicate”, “Duplicate”, “Similar” and “Unique”.

Image optimisation: Format of image file, presence of “alt” tag and “srcset” tag was assessed by reading the source code (html text) of the first landing page.

Page speed: The page speed-related features such as Speed Index, Largest Contentful Paint (LCP) and Lighthouse performance scoring were evaluated using the page speed insight tool from google (URL: <https://developers.google.com/speed/pagespeed/insights/>).

Link analysis: The presence of broken link(s) on the website and to the website, and Back links were analysed using the tool provided by ahrefs (URL: <https://ahrefs.com/broken-link-checker> and <https://ahrefs.com/backlink-checker> respectively). The presence of link to social media was assessed using SEO audit tool of seomator.

Miscellaneous features: The presence of schema.org code was assessed using structured data tool from google (URL: <https://search.google.com/structured-data/testing-tool/>). The presence of favicon was assessed using SEO audit tool of seomator. The presence of “viewport” tag was assessed by reading the source code (html text) of the first landing page.

IV. RESULTS

Domain check-up: All the websites (100%) included in the study had a valid SSL certificate and all could be served from https with ability to redirect http to https (Table 3). Twelve (75%) out of 16 websites were successfully loaded with or without the use of “www” prefix.

Files for search engine: Three (18.75%) websites did not have “robots.txt” file. Thirteen websites had robots.txt file with the location of the file in the root folder and the file

Table 2. SEO features and method of assessment used in the study

Parameters Studied	Method of Assessment	URL of Tool Used
Domain check-up		
Secure socket layer certificate	Manually	NA
Ability of website to be served from secure connection	Manually	NA
Ability to redirect http to https	Manually	NA
Ability to load with and without www prefix	Manually	NA

Files for search engine		
Presence and location of robots.txt	SEO audit tool of seomator	https://seomator.com/free-seo-audit-tool
Sitemap.XML file		
Presence of location information of sitemap file	By manually opening of robots.txt file	
Page quality check-up		
Page quality check-up	SEO audit tool of seomator	https://seomator.com/free-seo-audit-tool
Presence of Meta description tag	Reading the source code (html text) of the desired page	NA
Image optimisation		
Format of image file	Reading the source code (html text) of the desired page	NA
Presence of “alt” tag		
Presence of “srcset” tag		
Page speed		
Speed index	page speed insight tool from google	https://developers.google.com/speed/pagespeed/insights/
Largest Contentful Paint (LCP)		
Lighthouse performance scoring		
Link analysis		
Broken link on the website	tool provided by ahrefs	https://ahrefs.com/broken-link-checker and
Broken link to the website		
Back links	tool provided by ahrefs	https://ahrefs.com/backlink-checker
Presence of link to social media	SEO audit tool of seomator	https://seomator.com/free-seo-audit-tool
Miscellaneous features		
Presence of schema.org tag	structed data tool from google	https://search.google.com/structured-data/testing-tool/
Presence of favicon	SEO audit tool of seomator	https://seomator.com/free-seo-audit-tool
Presence of viewport tag	reading the source code (html text) of the desired page	NA

NA = Not applicable

Table 3. Summary of analysis of SEO features in the present study

Parameters Studied	Summary Statistics (% of sites in which a feature was present or median statistics as applicable)
Domain check-up	
Secure socket layer certificate	100
Ability of website to be served from secure Connection	100
Ability to redirect http to https	100
Ability to load with and without www prefix	75
Files for search engine	
Presence of robots.txt	81.25
Presence of sitemap.XML file	75
Presence of location information of sitemap file in robots.txt	50
Page quality check-up	
Page quality check-up	Appropriate Page Title length: 50% Presence of H1 tag: 43.75% Content Quantity: 115 Words Almost Duplicate Content: 50%
Presence of Meta description tag	93.75
Image optimisation	
Format of image file	jpeg: png: svg:
Presence of "alt" tag	43.75
Presence of "srcset" tag	12.5
Page speed	
Speed index	9.5 Sec
Largest Contentful Paint (LCP)	10.6 Sec
Lighthouse performance scoring	29.5 Score
Link analysis	
Broken link on the website	396 links *
Broken link to the website	626 links *

Back links	1514 links
Presence of link to social media	31.25%
Miscellaneous features	
Presence of schema.org tag	56.25
Presence of favicon	87.5
Presence of viewport tag	81.25

* "Mean" Statistics

having valid instructions for search engines. However, the location information of sitemap file within the robots.txt file was missing in 8 (50%) out of 16 websites. Four (25%) websites did not have the sitemap xml file while 12 websites had the file located in root folder.

Page quality check-up: In 8 (50%) websites, the pixel length of page title exceeded the recommended length of 285 to 575 pixels. Meta description tag was absent in 1 (6.25%) out of 16 websites. H1 heading tag was absent in 9 (56.25%) websites. Similarly, H2, H3, H4, H5 and H6 heading tags were absent in 10, 8, 9, 12 and 13 websites, respectively.

Page content quantity ranged from 23-1924 with a median of 115 words. Eight (50%) websites had presence of "almost duplicate" content. In these eight websites, the extent of duplicate content varied from 0-40% with a median of 6.5%. One (6.25%) website did not have any unique content whereas eight (50%) websites had 100% unique content. Median percentage of unique content was 90%.

Image optimization: The landing page of eight, six and two out of 16 websites used "jpeg", "png" and "svg" image formats, respectively. The alternate tag ("alt" tag) was missing in 9 (56.25%) out of 16 websites. Only 2 (12.5%) websites used "srcset" tag for various device optimization.

Page speed: The speed index of the landing page of websites ranged from 4.2 to 27.1 seconds with a median index of 9.55 seconds. Time to display largest content (Largest Contentful Paint - LCP) ranged from as early as 2.3 seconds to as delayed as 44.5 seconds. The median LCP was 10.6 seconds. The Lighthouse performance score of the websites studied ranged from 5 to 60 with a median of 29.5.

Link analysis: Eleven (68.75%) out of 16 websites did not show on-site broken link. On-site broken link ranged from 0-5975 links with a mean of 396 links. Seven (43.75%) out of 16 websites did not show broken link directed to site. Broken link to site ranged from 0 to 7510 links with a mean of 626 links.

Back link ranged from 41 to 3607478 links with a median of 1514 links. Link to social media was present in 5 (31.25%) out of 16 websites.

Miscellaneous: Two (12.50%) out of 16 websites, did not have favicon. Structured data for search engines in the form of "schema.org" tag, was present in 9 (56.25%) out of 16 websites. Thirteen (81.25%) websites used "viewport" tag for mobile device responsiveness.

V. DISCUSSION

In current times, health and medical information is becoming increasingly readily available on the internet [11]. The internet is used by millions of people around the world for seeking relevant health and medical information [12]. The increasing use of mobile has led to ever-expanding access to the internet as well as expansion of online search related to health and medical information [13]. Of the clients seeking health information on internet, about 80% of clients start with a search engine [1]. A few previous studies have demonstrated that clients' viewership is higher for the top 10 websites listed by a search engine [1]. Hence is the need of search engine optimization (SEO).

Although researchers and scientists do not consider themselves as marketers, their research needs marketing in today's world of high internet use [2]. Similarly, SEO is also gaining importance for medical sciences' products such as research work, research publication, medical journal, hospital services and clinical laboratories. The clinical laboratories have emerged as vital components of the modern evidence-based medicine. Although SEO is valuable for small laboratories catering to local geographical patients, it is much more crucial for the larger laboratories having presence in multiple geographical locations. Rayess et al demonstrated that creating a search engine optimized website with good content not only increases the website's traffic but also results in more informed patients [1]. Informed and knowledgeable patients in turn would potentially result in reduction in the number of litigations. An analysis of rhinoplasty websites for SEO showed that having an online presence is crucial for successful facial plastic surgery practice [1]. In the present study, websites of 16 Indian clinical laboratories (Table 1) were analyzed to assess their SEO. In our study, most of the parameters of domain check-up were found to be optimized in the websites of clinical laboratories except for the ability to load with and without "www" prefix.

The robots.txt file is read by web spider (part of search engine) and the file guides the permission aspects for the crawling process of search engine. In the present study, the robots.txt file was found to be missing in 18.75% of websites studied. Although, web spider can crawl the webpages using links, the sitemap.xml file is very helpful to the web spider for this purpose. The sitemap.xml file was absent in 25% of websites studied in our study. Like robots.txt and sitemap.xml file, the presence of structured data on webpage is also helpful for web spiders. Structured data in the form of schema.org was developed with collaborative efforts of Bing, Google, Yahoo and Yandex

[14]. This was developed in consideration of the difficulty of converting HTML text to structured data and the dynamic nature of page HTML. The schema.org" tag was present in only 56.25% of clinical laboratories' websites in the present study.

In the present study, only the first landing page of the websites was studied for assessment of page quality parameters, including the appropriate length of title tag, presence of headers, amount of page content, duplicity of content and presence of meta description tags. Only 50% of websites studied had appropriate length of title tag. Although header tags are important visual cues to readers, search engines can easily understand header tags. Header tags are conventionally labelled from H1 to H6 [15]. H1 level header was present in only 43.75% of websites in our study. The page content of the websites was thin with an average of 115 words. However, presence of "almost duplicate" content was identified 50% of websites of clinical laboratories studied. Aswani et al reported that website administrators sometime use Black Hat techniques including duplicate contents to trick algorithms of search engine [16]. In the presence of google PANDA update, such techniques may result in imposition of penalty by google search engine. Most of the websites (93.75%) in the present study had used meta description tag in the present study.

In the present study, almost all the websites used images in their landing page, however the use was poorly optimized. The "alt" and "srcset" tag were present in only 43.75 and 12.5% of websites, respectively.

The time taken by the webpage to load revealed wide variations as evident from LCP range of 2.3 to 44.5 seconds. There was wide variation in the number of back links which ranged from 41 to 3607478 links. Zhang and Dimitroff⁸ concluded that websites with more number of back links were regarded as more influential by the search engines. Our study found that links to social media was present in only 31.25% of the clinical laboratories' websites. Social media has been reported to play a very important role for better page rank by search engine [10].

Due to easy availability and affordability of smart mobile phones, people are using these devices for internet search to find various services. This generates the need for websites to be mobile-responsive and mobile-friendly. Most of the websites (81.25%) in the present study were mobile device responsive as they used "viewport" tag. This indicates that the websites of these clinical laboratories can be easily accessed on mobile phones thus, widening their reach of services.

Even though SEO is vital in medical practice, the use of SEO techniques is fraught with ethical concerns as well. Health care providers with better financial resources can attempt to polarize patients' flux to their advantage, out-bidding competitors even if they themselves are providing

suboptimal or non-standard medical treatment. Search engine providers usually track the usage of keywords to predict user behaviour. Firms with limited medical knowledge may use certain keywords inappropriately which may mislead the patients during web search. All these factors raise the vulnerability of prospective patients to unwarranted laboratory tests and body check-ups [17,18].

The present study is the first attempt till date to analyze the search engine optimization of clinical laboratories' websites. However, a few limitations of our study need mention, such as the small sample size, and provision of only a snapshot of website status as of December 2020. Other limitation of the study is that use and quantity of keywords by web sites was not analyzed.

VI. CONCLUSION

The present study demonstrated that websites of Indian clinical laboratories were mostly optimized for the search engines. Few shortcomings such as thin and duplicate page content, poorly optimized use of images and less frequent use of structured data were identified in certain websites. Hence, it is the need of the hour for the clinical laboratories in India to design better and optimized websites suitable for the search engines and digital marketing needs.

REFERENCES

- [1] H.M. Rayess, A. Gupta, M. Nissan, M.A. Carron, G.F. Zuliani, "Search Engine Optimization: An Analysis of Rhinoplasty Web sites", *Facial Plastic Surgery*, Vol. 33, pp. 665-669, 2017.
- [2] M. Cushman, "Search engine optimization: What is it and why should we care?", *Research and Practice in Thrombosis and Haemostasis*, Vol. 2, pp. 180-181, 2018.
- [3] K. Dye, "Website abuse for search engine optimisation", *Network Security*, Vol. 3, pp. 4-6, 2008.
- [4] R.A. Malaga, "Search Engine Optimization - Black and White Hat Approaches", *Advances in Computers*, Vol. 78, pp.1-39, 2010.
- [5] H.J. Tsuei, W.H. Tsai, F.T. Pan, G.H. Tzeng, "Improving search engine optimization (SEO) by using hybrid modified MCDM models", *Artificial Intelligence Reviews*, Vol. 53, pp.1-16, 2018.
- [6] J. Bar-Ilan, M. Mat-Hassan, M. Levene, "Methods for comparing rankings of search engine results", *Computer Networks*, Vol. 50, pp.1448-1463, 2006.
- [7] G.R. Amin, A. Emrouznejad, "Optimizing search engines results using linear programming", *Expert Systems with Applications*, Vol. 38, pp.11534-11537, 2011.
- [8] J. Zhang, A. Dimitroff, "The impact of webpage content characteristics on webpage visibility in search engine results (part I)", *Information Processing and Management*, Vol. 41, pp.665-690, 2005.
- [9] D. Goyal, D. Goyal, "Comparative Analysis of Various Google Algorithms and their Effects on Performance of a Web Portal", *International Journal of Engineering Research & Technology*, Vol. 2, pp.1005-1008, 2013.
- [10] Z. Xiang, U. Gretzel, "Role of social media in online travel information search", *Tourism Management*, Vol. 31, pp.179-188, 2010.
- [11] K. Viswanath, "Science and society: the communications revolution and cancer control", *Nature Reviews Cancer*, Vol. 5, pp.828-835, 2005.

- [12] N.L. Atkinson, S.L. Saperstein, J. Pleis, "Using the internet for health-related activities: findings from a national probability sample", *Journal of Medical Internet Research*, Vol. **11**, pp.e4, 2009.
- [13] S. Koch-Weser, Y.S. Bradshaw, L. Gualtieri, S.S. Gallagher, "The Internet as a health information source: findings from the 2007 Health Information National Trends Survey and implications for health communication", *Journal of Health Communication*, Vol. **15**, Issue **Suppl3**, pp.279-293, 2010.
- [14] R.V. Guha, D. Brickley, S. Macbeth, "Schema.org: Evolution of structured data on the web", *Communications of the ACM*, Vol. **59**, pp.44-51, 2015.
- [15] J.P. Sheffield, "Search engine optimization and business communication instruction: Interviews with experts", *Business and Professional Communication Quarterly*, Vol. **83**, pp.153-183, 2020.
- [16] R. Aswani, A.K. Kar, P.V. Ilavarasan, Y.K. Dwivedi, "Search engine marketing is not all gold: insights from Twitter and SEOClerks", *International Journal of Information Management*, Vol. **38**, pp.107-116, 2018.
- [17] K. Denecke, P. Bamidis, C. Bond, E. Gabarron, M. Househ, A.Y.S. Lau, M.A. Mayer, M. Merolli, M. Hansen "Ethical Issues of Social Media Usage in Healthcare", *Yearbook of Medical Informatics*, Vol. **10**, pp.137-147, 2015.
- [18] W. Yao, C.H. Chu, Z. Li, "The use of RFID in healthcare: Benefits and barriers," 2010 IEEE International Conference on RFID-Technology and Applications, Guangzhou, 2010, pp. 128-134, 2010.

AUTHORS PROFILE

Dr. Praveen Prakash completed MBBS from Calcutta National Medical College & Hospital, Kolkata, West Bengal and MD (Pathology) from North Bengal Medical College and Hospital, Darjeeling, West Bengal. He is working as a Senior resident in the Department of Pathology at North DMC Medical College & Hindu Rao Hospital, New Delhi, India.



Santosh Kumar Sharma, MSc Biotechnology and PhD Scholar, is working as a laboratory technologist in the Department of Pathology, Hindu Rao Hospital, Delhi.



Dr Arpita Joshi, MBBS, MD Pathology, has been working as a Senior Resident at the Department of Pathology, Hindu Rao Hospital since 2019. She pursued her MD in Pathology at Sarojini Naidu Medical College, Agra in 2019. She has few research publications in various national journals. Her main area of interest is haematology. She has a keen interest in teaching and research.



Dr Aarzo Jahan is presently working as medical officer in department of pathology, NDMC medical college and Hindu Rao hospital. She did her MBBS in 2011 from MMIMSR, Kurukshetra University. She completed her MD (Pathology) from PGIMS Rohtak in 2015. After this she did her senior residency from UCMS Delhi. She has a keen interest in cytopathology and hematopathology.



Dr Namrata Sarin is the head and specialist in Department of Pathology, NDMC medical college and Hindu Rao Hospital. She did her MBBS from Government Medical college, university of Jammu followed by MD (Pathology) from AFMC, Pune. She did her senior residency from Lady Hardinge Medical College Delhi. She has keen interest in Dermatopathology.



Dr Sompal Singh, BSc Statistics, MBBS, MD Pathology and MBA (Hospital Administration), has been working as a Senior Specialist at the Department of Pathology, Hindu Rao Hospital since 2005. He pursued his MD in Pathology at Maulana Azad Medical College, New Delhi followed by residency at the same institute. He has more than 120 research publications in various national and international journals. Dr Singh has authored a book on basic biomedical statistics and co-authored two books for paramedical students and one for medical undergraduate students. He has a keen interest in teaching, and research.

