Aggregate Search Engine Query Trend Reveal Trade-off in Interest and Awareness on HIV/AIDS in Nigeria and South Africa.

Babatunde Okunoye

Working Draft, v.1

The Acquired Immune Deficiency Virus (AIDS), caused by the Human Immunodeficiency Virus (HIV) is a leading cause of mortality on the African continent and world. HIV and AIDS are among the greatest public health challenges confronting health authorities around the world. The greatest burden of HIV and AIDS is felt in sub-Saharan countries, with Nigeria and South Africa having the greatest incidence of the disease in the world. AIDS is the final stage of HIV infection, but with effective treatment and healthy living, HIV infection will not develop into AIDS. With increased awareness campaigns by national, regional and international health institutions, there has been an increased awareness on HIV and AIDS across the world. Using data from web searches on HIV/AIDS submitted online on the most widely used search engine Google from 2004-2020, a trade-off pattern is seen in the web searches on HIV and AIDS in Nigeria and South Africa showing the decline in search interest for AIDS relative to HIV. This trend mirrors progress on the ground in both countries, with a decline in AIDS-related deaths and people living longer with the HIV virus. This observed trend might be an indicator that multilateral efforts at combating HIV/AIDS, particularly through awareness raising and behavioural interventions in these two most endemic countries are bearing fruit, not just on the ground, but also reflected in the online information seeking on the HIV/AIDS pandemic.

Keywords: HIV/AIDS, Web search, online behaviour, Big Data.

1. Introduction

HIV/AIDS remains a major public health concern for nations of the world, despite concerted efforts aimed at combating the disease. There are an estimated 37.9 (32.7- 44.0) million people living with HIV (Human Immunodeficiency Virus) – the virus which causes AIDS in the world¹. The greatest burden of the HIV/AIDS scourge is felt in sub-Saharan Africa, with South Africa and Nigeria bearing the greatest burdens of the disease in the world. An estimated 7.7

---

¹ UNAIDS Data 2019, https://www.unaids.org/sites/default/files/media_asset/2019-UNAIDS-data_en.pdf

million people were living with HIV and AIDS in South Africa in 2018, with an adult prevalence rate of 20.4%². In 2018, 71,000 South Africans died of AIDS related causes. Nigeria has the second highest number of new HIV infections globally after South Africa, and an estimated 1.5% of the population living with HIV³. Having the largest population on the African continent, this translates to about 1.9 million people living with HIV as at 2018.

Considerable progress has been made in the fight against HIV/AIDS globally. Globally, an estimated 37.9 (32.7–44.0) million people were living with HIV in 2018⁴, an increase from previous years (for example 35.3 million in 2012) as more people are receiving the life-saving antiretroviral therapy. At the same time the number of AIDS deaths is also declining with 770,000 deaths in 2018⁵, from 1.6 (1.4–1.9) million deaths in 2012, down from 2.3 (2.1–2.6) million in 2005⁶. Gains have also been made towards many of the 2020 Sustainable Development Goal targets and elimination commitments, although significant challenges remain⁷.

A key component of the fight against HIV/AIDS has been behavioural interventions⁸. A global meta-analysis of studies determined that behavioural interventions reduce sexual risk behaviour and avert sexually transmitted infections and HIV⁹. Behavioural interventions typically involve harmonizing messages and the dissemination of information about HIV transmission and various prevention approaches using television, radio, outdoor advertising and information and communications technology. The outcome of all these behavioural interventions is that in sub-Saharan Africa, the percentage of young people (15–24 years) demonstrating comprehensive and accurate understanding of HIV has generally increased over the years. For example, this understanding rose by five percentage points for men and by three for women from 2002 to 2011, although knowledge levels remain low¹⁰ (36% for young men and 28% for young women).

In South Africa, there have been a number of nationwide communication campaigns related to raising awareness of HIV and AIDS. Soul City and MTV Shuga are two examples of such multi-media campaigns with effective outcomes and promote good sexual health and well-being¹¹. Research into the impact of the Soul City and other communication campaigns found

---

² Ibid

³ Ibid

⁴ Ibid

⁵ Ibid

⁶ UNAIDS data, 2012, https://www.unaids.org/en/resources/documents/2012/20121120_UNAIDS_Global_Report_2012

⁷ UNAIDS, “Sustainable Development Goals and the HIV response”, https://www.unaids.org/sites/default/files/media_asset/SDGsandHIV_en.pdf

⁸ UNAIDS 2012, 2012 UNAIDS Report on the Global AIDS Epidemic, 20 November 2012, https://www.unaids.org/en/resources/documents/2012/20121120_UNAIDS_Global_Report_2012

⁹ Ibid

¹⁰ UNAIDS 2012, 2012 UNAIDS Report on the Global AIDS Epidemic, 20 November 2012, https://www.unaids.org/en/resources/documents/2012/20121120_UNAIDS_Global_Report_2012

¹¹ HIV and AIDS in South Africa, <https://www.avert.org/professionals/hiv-around-world/sub-saharan-africa/south-africa>

that they had a positive effect on the sexual behaviour of adults that had been exposed to the campaign message, for instance by bringing about positive changes in condom use and HIV testing¹².

In Nigeria, educational campaigns in schools, where Family Life and HIV Education lessons are part of the Nigerian school curriculum¹³ and nationwide multimedia campaigns like the one created by the Society for Family Health¹⁴ are thought to have been successful in increasing knowledge and changing behaviour.

The combined effects of these awareness campaigns have been an increase in knowledge and understanding of HIV/AIDS in these two countries. An outcome of an increase in awareness of HIV/AIDS in South Africa and Nigeria is probably that a percentage of the population will turn to the Internet via search engines for information on HIV/AIDS. Internet penetration in South Africa and Nigeria is put at 56% and 42% respectively¹⁵.

An increasing number of people are using the Internet to support their healthcare¹⁶ and search engines such as Google, Yahoo! and Bing are often their first port of call. One of the early studies on search engine use found that one in every 28 (3.5%) pages viewed on the Web is a search results page, making the use of a search engine the second most popular Internet task next to email¹⁷. Another early study found that worldwide there are over 6.75 million health-related internet searches carried out every day, representing approximately 5% of all internet searches¹⁸. Search engine use is the most important approach to online information seeking¹⁹, and a study by the Pew research centre found that half of all Internet users now use search engines on a typical day²⁰.

This paper examines trends in web searches on HIV and AIDS on Google search which is the dominant search engine worldwide²¹ and is the top search engine in most countries with a few exemptions such as Russia (Yandex) and China (Baidu). In the countries of this study, Nigeria

---

¹² Peltzer *et al* (2012), Impact of National HIV and AIDS Communication Campaigns in South Africa to Reduce HIV Risk Behaviour, ScientificWorld Journal, Nov 8 2012, doi: 10.1100/2012/384608, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3504395/>

¹³ National Agency for Control of AIDS (2015). Global AIDS Response, Country Progress Report, https://www.unaids.org/sites/default/files/country/documents/NGA_narrative_report_2015.pdf

¹⁴ Society for Family Health, "Enhancing Nigeria's Response to HIV and AIDS", <https://www.sfhnigeria.org/enhancing-nigerias-response-to-hiv-and-aids/>

¹⁵ ITU Statistics, "Individuals using the Internet 2000 - 2018", [https://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2019/Individuals Internet 2000-2018 Dec2019.xls](https://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2019/Individuals%20Internet%202000-2018%20Dec2019.xls)

¹⁶ Powell J, Clarke A (2002). The WWW of the World Wide Web: Who, What and Why? Journal of Medical Internet Research 2002; 4(4).

¹⁷ Janseen BJ and Pooch U (2001). A Review of Web Searching Studies and a Framework for Future Research Journal of the American society for information science and technology—52(3):235–246 February 1, 2001.

¹⁸ Eysenbach G and Kohler C (2003). What is the Prevalence of Health-related Searches on the World Wide Web? Qualitative and Quantitative Analysis of Search Engine Queries on the Internet. Proc AMIA Symp 2003:225-229)

¹⁹ Morris *et al* (2010). A comparison of Information seeking on search engines and social networks. Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media.

²⁰ Deborah Fallows, "Search Engine Use", Pew Research Centre Internet and Technology, August 6 2008, <https://www.pewresearch.org/internet/2008/08/06/search-engine-use/>

²¹ StatCounter Global Stats, <https://gs.statcounter.com/search-engine-market-share>

and South Africa, Google is also the dominant search engine²². Google data have been used severally in the aid of understanding health trends and in health forecasting²³²⁴²⁵²⁶²⁷²⁸²⁹. And although the use of Google search engine data in the aid of health and development in general have been criticized in recent years for an unbalanced and opaque approach which did not satisfy scientific transparency, in recent years organizations such as the World Bank³⁰³¹ and several government departments around the world have begun revisiting the behavioural BigData methods of search engines data in the aid of international development. Drawing on feedback from the scientific community on the best practices of BigData research³²³³³⁴³⁵³⁶³⁷, particularly paying attention to integrating both BigData and ‘small data’ (e.g. survey data and other traditional data collection mechanisms) to create more complete and accurate datasets, they have made significant advances in this field.

This paper shows that aggregate search term data for the search terms ‘hiv’ and ‘aids’ in Nigeria and South Africa for the years 2004 – 2014 both show a pattern of the search interest for ‘aids’ declining relative to ‘hiv’ perhaps revealing increasing general understanding on the difference between HIV and AIDS, particularly the awareness that with proper treatment HIV infection may not lead to AIDS. This observed trend might be an indicator that multilateral efforts at combating HIV/AIDS, particularly through awareness raising and behavioural interventions in these two most endemic countries are bearing fruit. This relative shift in search interest mirrors the reality of the disease on the ground, with declining numbers of deaths from AIDS and with more people living with HIV in both countries.

---

²² Ibid

²³ Carneiro HA and Mylonakis E (2009). Google trends: a web-based tool for real-time surveillance of disease outbreaks. *Clinical infectious diseases* 49 (10): 1557-1564.

²⁴ Polgreen PM, Chen Y, Pennock DM, Nelson FD (2008) Using internet searches for influenza surveillance. *Clin Infect Dis* 47: 1443–1448

²⁵ Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS and Larry Brilliant L (2009). Detecting Influenza epidemics using search engine query data. *Nature* 457: 1012-1014.

²⁶ Hulth A, Rydevik G, Linde A (2009) Web queries as a source for syndromic surveillance. *PLoS One* 4: e4378–e4378.

²⁷ Pelat C, Turbelin Cm, Bar-Hen A, Flahault A, Valleron A-J (2009) More diseases tracked by using Google Trends. *Emerg Infect Dis* 15: 1327–1328.

²⁸ Chan EH, Sahai V, Conrad C, Brownstein JS (2011) Using Web Search Query Data to Monitor Dengue Epidemics: A New Model for Neglected Tropical Disease Surveillance. *PLoS Negl Trop Dis* 5(5): e1206. doi:10.1371/journal.pntd.0001206.

²⁹ Ocampo AJ, Chunara R, and Brownstein JS (2013). Using search queries for malaria surveillance, Thailand. *Malaria Journal* 12:390.

³⁰ Reyes G, “Leveraging behavioral insights in the age of big data”, World Bank blogs, May 31 2017, <http://bit.ly/34dezWR>

³¹ Zeina A et al (2018), “Behavioral Science around the world: Profiles of 10 countries”, The World Bank, <http://bit.ly/2XBCxZv>

³² Butler D, “When Google got flu wrong”, *Nature* 494, 155–156, 14 February 2013, doi:10.1038/494155

³³ Lazer D et al (2014), Google Flu Trends Still Appears Sick: An Evaluation of the 2013-2014 Flu Season, March 13, 2014, available at <http://dx.doi.org/10.2139/ssrn.2408560>

³⁴ Fung K, “Google Flu Trends’ Failure Shows Good Data > Big Data”, *Harvard Business Review*, March 25 2014, <http://bit.ly/2XPiORb>

³⁵ Lazer et al (2014), “The Parable of Google Flu: Traps in Big Data Analysis”, *Science* 343, 1205, 10.1126/science.1248506

³⁶ Ruths D and Pfeffer J (2014), “Social media for large studies of behavior”, *Science* 346 (6213) 1062 – 1063, <https://science.sciencemag.org/content/346/6213/1063>

³⁷ Lazer et al (2014), “The Parable of Google Flu: Traps in Big Data Analysis”, *Science* 343, 1205, 10.1126/science.1248506

2. Method

Aggregate, anonymized Data on web searches on the search terms ‘hiv’ and ‘aids’ were obtained on Google Trends in order to visualize how web searches were conducted on both search terms. Google trends data was obtained on the two search terms for 2004-2020. Data on regional interests in both search terms for both South Africa and Nigeria were downloaded, together with related queries bearing the search terms.

3. Results

Figure1: Snapshot on searches on the search terms ‘hiv’ and ‘aids’ in South Africa accessed July 12 2020.

Figure 2: Regional search interest for “hiv” in South Africa accessed July 12 2020.

Figure 3: Regional search interest for “aids” in South Africa, accessed July 12 2020.

Figure 4: Some related search queries for “hiv” in South Africa. Accessed July 12 2020.

Figure 5: Some related search queries for “aids” in South Africa. Accessed July 12 2020.

Table 1: Top search queries on ‘hiv’ in South Africa, accessed July 12 2020

Top ‘hiv’ search queries		
aids	hiv cure	signs of hiv
hiv aids	hiv in south africa	what is aids
hiv symptoms	hiv treatment	living with hiv
symptoms	hiv signs	what is hiv aids
hiv and aids	hiv infection	hiv aids in south africa
hiv positive	people with hiv	tb
what is hiv	hiv testing	hiv virus
symptoms of hiv	hiv/aids	hiv news
hiv test	hiv negative	hiv test kit/arv

Table 2: Top search queries on ‘aids’ in South Africa, accessed July 12 2020

Top ‘aids’ search queries		
hiv	meaning of aids	meaning of hiv
hiv aids	symptoms of aids	causes of aids
hiv and aids	aids day	hiv cure
what is aids	what is hiv and aids	hiv aids cure
what is hiv aids	hiv/aids	full meaning of hiv
aids symptoms	hiv symptoms	tb
teaching aids	world aids day	hiv virus
what is hiv	aids cure	aid
aids meaning	symptoms of hiv aids	full meaning of aids

Figure 6: Snapshot on searches on the search terms “hiv” and “aids” in Nigeria accessed July 12 2020.

Figure7: Regional search interest for “hiv” in Nigeria accessed July 12 2020.

Figure 8: Regional search interest for “aids” in Nigeria, accessed July 12 2020.

Figure 9: Some related search queries for “hiv” in Nigeria. Accessed July 12 2020.

Figure 10: Some related search queries for “aids” in Nigeria. Accessed July 12 2020.

Table 3: Top search queries on “hiv” in Nigeria, accessed July 12 2020

Top “hiv” search queries		
symptoms	hiv positive	hiv news
hiv symptoms	signs of hiv	early symptoms of hiv
symptoms of hiv	hiv meaning	hiv/aids
hiv aids	causes of hiv	hiv drug
aids	meaning of hiv	what are the symptoms of hiv
what is hiv	hiv drugs	hiv rash
hiv in nigeria	hiv virus	hiv virus
hiv test	cure for hiv	malaria
hiv and aids	hiv treatment	

Table 4: Top search queries on ‘aids’ in Nigeria, accessed July 12 2020

Top ‘aids’ search queries		
hiv	meaning of aids	aid
hiv aids	symptoms of aids	causes of aids
hiv and aids	aids day	full meaning of aids
what is aids	what is hiv and aids	teaching aids
what is hiv and aids	hiv/aids	aids day
aids symptoms	hiv symptoms	
teaching aids	world aids day	
what is hiv	aids cure	
aids meaning	symptoms of hiv aids	

4. Discussion and conclusion

The data on the search terms ‘HIV’ and ‘AIDS’ in Nigeria and South Africa (2004 – 2020) reveals a decrease in interest in ‘aids’ related search terms relative to ‘hiv’ related search terms. This mirrors the progress on the ground in both countries where AIDS related deaths have reduced over the years and people have lived longer with the HIV virus as life changing medications have prolonged the lives of people living with the disease. Similarly, behavioural interventions through messages delivered through media sources, including the Internet in both countries have increased the numbers of people with accurate understanding of the differences between HIV and AIDS. In response to these behavioural interventions a proportion of the populations in these countries will resort to the Internet via search engines to seek information on HIV and AIDS.

This paper illustrates the potential of aggregate search engine data to provide additional data signals to supplement or corroborate traditional data collection methods such as surveys, in statistically poor contexts of developing countries. This study allowed for lengthy time-series of data (2004 – 2020) to demonstrate that this observed behaviour can be put in the context of a lengthy period of years, thus potentially correcting for platform-specific algorithms relating to Google.