SOCIAL MEDIA BIG DATA: A DIARY STUDY OF TEN PHARMACEUTICAL FIRMS

MASTER OF BUSINESS ADMINISTRATION

In the

RHODES BUSINESS SCHOOL

By

NADIA SAMANTHA BAKER

Student number: G08B5969

March 2019

Supervisor: Professor Deon Nel

Abstract

Purpose: The goal of the research was to demonstrate how firms can use social media big data, to make strategic business decisions, through the lens of Resource Based Theory (RBT) and Dynamic Capability Theory (DCT), that could lead to a sustained competitive advantage. In and of its own, big data, does not constitute a competitive advantage. It may hold value for the firm, but lacks rarity, inimitability, and is not substitutable (Braganza, et al. 2017; Mata, Fuerst and Barney, 1995; Delmonte, 2003). It is in the analysis of this data, through RBT and DCT, that will turn the information into useful business intelligence (Amit and Schoemaker, 1993; Barney, 1991; 1995; Marr, 2015; Gupta and George, 2016; Kurtmollaiev, et al., 2018). Most importantly, firms must constantly reconfigure their resources in line with the dynamic business environment to ensure superior performance (Teece, Pisano and Shuen, 1997; Helfat, et al., 2007; Teece, 2014; 2018).

Method: In this study, a qualitative approach was used to examine the RBT (Value, Rarity, Inimitability and Non-Substitutable - VRIN Framework) and DCT, to describe and understand the relevant theories and to build upon the quantitative results. While a quantitative approach was used to analyse the social media sentiment as depicted by Social Mention metrics. A novel technique, Chernoff Faces, was used to analyse and visualize the data (de Vos, Strydom, Fouche and Delport, 2011).

Results and Findings: The research results show that, while the 10 firms in the study all have a presence on social media, it is on selective platforms. The content that is posted, is on very specific topics (Narayan, 2017; Cornejo, 2018). The Chernoff Faces indicate that the firms' Social Mention metrics, over the 30 day period, was at low values. Since strength of social mention is depicted by the face line, the thin, long, generally sad looking faces implies that more than 70 percent of the firms' social media strength over the study period, was weak.

Conclusion: The literature indicates that the true value of big data and big data analytics can only be realised if firms make sound business decisions and act upon it swiftly.

Acknowledgements

I would like to extend my heartfelt gratitude to my amazing husband, Rudy Baker. Thank you for keeping our home going and making time to proofread and edit my paper. To my daughter, Keziah who is wise beyond her years, for her constant encouragement and my son Aidan for always keeping me grounded.

✋👍❤️

Thank you to Rhodes University for second chances and Professor Skae, who's words will always ring true "… it's not how you START, it's how you END …".

To my supervisor Professor Nel, thank you for motivating me to always do more.

Table of Contents

List of Tables

List of Figures

1 Chapter 1: Introduction

1.1 Introduction

Social media is developing as the most expeditious source of big data and healthcare content is the fastest growing category, with 1 in 20 Google searches, being for health-related matters (Ray, 2017). The PEW Research Centre reported that 80% of Internet users, seek out health data online. 25% have watched videos concerning a medical or health matter and 24% looked up specific medications and treatment reviews (Fox, 2011; Cornejo, 2018).

The field of social media big data and big data analytics have come to be more and more critical to academics and practitioners, across the various industries. Big data and big data analytics proficiency have become the basis of industry competition and the early adopters are reaping the rewards. Big data and analytics can be exploited to improve firm performance and generate new business prospects. Many firms are under pressure to develop the capability and effective and efficient operational processes, to realize this value. According to the McKinsey Global Institute the US will confront a dearth of 140 000 to 190 000 people with natural analytical abilities and 1,5 million supervisors with the aptitude to analyse big data, to make sound business decisions (Manyika, Chui, Brown, Bughin, Dobbs, Roxburgh and Hung Byers, 2011; Cattell, Chilikuri and Levy, 2013 and Henke, Bughin, Chui, Manyika, Saleh, Wiseman and Sethupathy, 2016).

Even though social media has been identified as the most expeditious source of big data and healthcare content, as the fastest growing category, the health-care sector, particularly Pharmaceutical companies, have been labelled laggards in adopting social media big data strategies (Aitken, 2014; Cornejo, 2018).

1.2 Problem Statement

How can Pharmaceutical firms use social media big data and big data analytics, through the application of the VRIN framework of RBT and the practical application of DCT, to gain a sustained competitive advantage in the healthcare sector?

The Pharmaceutical industry has been described as being slow in taking up the social media big data initiatives, due to several barriers. The most obvious barrier being industry regulation. For the healthcare practioner (HCP), barriers such as patient confidentiality, malpractise lawsuits, etc. represents risk (Aitken 2014, Ray 2017). The literature and this study, aims to indicate ways in which the industry can circumvent these impediments and realise substantial business value from this unprecedented, global phenomenon of big data and big data analytics.

1.3 Purpose of the Research

The purpose of the researcher was to demonstrate how Pharmaceutical firms, through the lens of Resource-Based Theory (RBT) and Dynamic Capabilities Theory (DCT), can make strategic business decisions, that could lead to a sustained competitive advantage, by using social media big data.

1.3.1 Research Goals

The goal was to illustrate to Pharma, the benefits of utilizing RBT and DCT in the assimilation and analysis of big data in formulating business strategy, in order to achieve sustained competitive advantage.

1.3.2 Research Aims

1. To examine how Pharmaceutical firms, through the lens of RBT and DCT, can make strategic business decisions, that could lead to a sustained competitive advantage, by identifying:

a. The key aspects of RBT, that will assist firms in adopting internal business strategies, that can lead to sustainable competitive advantage and;
b. The Dynamic Capabilities required to effectively leverage big data to achieve a sustained competitive advantage.

2. To illustrate how Pharmaceutical firms can use existing social media platforms, as an external source, to gather big data, to analyse their social media sentiment, by utilizing:

   a. A social media platform namely: Social Mention. Social Mention is a social media platform that collects various variables of social media sentiment, from more than 80 social media sites.
   b. A graphical technique, *'Chernoff Faces*' to analyse the various social media sentiment variables.

1.4 Conceptual Framework

During the 1980's, the focus of strategic management was mainly based on the external environment of a firm, to achieve a competitive advantage (Porter, 1980). Conversely, in the same epoch, Wernerfelt (1984; 1989) argued that firms should be looking at internal resources, to determine profitability. This resource-based view moved the attention of strategic analysis, from the industry, to the firm. It emphasized that firms' resources and internal capabilities are primarily diverse and should be inimitable. It suggests that resources and capabilities are the foundation of strategies (Barney, 1991; Ray, Barney and Muhanna, 2004; Barney and Clark, 2007; Lockett, O'Shea and Wright, 2008). The literature emphatically states that it's a firm's ability to uniquely combine their various resources and capabilities in such a way that the rivals are unable to copy or match it, that ultimately will allow them to achieve sustained competitive advantage (Barney, 1991; Grant, 1991; Teece, Pisano and Shuen, 1997; Helfat, 2011; Pisano, 2015; Teece, 2018).

Figure 1.1 Conceptual Framework

1.4.1 Theoretical Underpinnings

This notion of moving from external to internal strategy analysis gave rise to Resource Based Theory (RBT). Following the proliferation of RBT literature, several sub-fields arose. This has seen the emergence of the Dynamic Capability Theory (DCT).

1.4.1.1 Resource Based Theory

RBT gained traction when Barney (1986a; 1986b; 1986c; 1991) identified four empirical indicators of resources, namely: value, rarity, inimitability and non-substitutability (VRIN Framework). The VRIN framework analyses, under which situations, resources will give rise to greater firm performance. Further to this, Grant (1991), created a five-phase procedure, which is a pragmatic way to resource-based strategy analysis, i.e. assessing resources, evaluating capabilities, analysing income potential of resources and capabilities, selecting a plan and the augmenting and replenishing of resources and capabilities. The study will look at how Barney (1991) VRIN framework and Grant (1991) practical approach to strategy analysis, can be applied by Pharmaceutical firms, to adopt social media big data as a business strategy, to gain competitive advantage.

1.4.1.2 Dynamic Capability Theory

Teece, Pisano and Shuen (1997), describes dynamic capability as a firm-specific know-how, to develop, recreate and fit internal and external capabilities, to deal with the rapidly transforming business environment. They go on to explain how it's in the grouping of various capabilities and resources, that firms will set themselves apart from their rivals (Teece, 1992; 1996). Fundamentals of this approach are evident in Teece's earlier works and that of other researchers as cited in (Teece, Pisano and Shuen, 1997, p. 510). In a more recent study, Teece (2018) says that the Internet has created prospects for radical, innovative business models, to which firms' strategies must react. This study will analyse how the firms can use DCT to adapt their capabilities, to enhance big data analytics as to exploit opportunities offered by this constant changing business environment, specifically focusing on the social media landscape.

1.5 Social Media Landscape

One of the marvels of today, changing the globe as we are acquainted with it, is the global accessibility to the Internet. According to Kemp (2018), as at January 2018, of the total population of 7,593 billion people, 4,021 billion people are Internet users and 3,196 billion are active on social media. The annual (2017 – 2018) digital growth for Internet users has increased by 7% (248 million users) and social media active users by 13% (362 million users). The recent Digital 2019 report reveals that social media users have increased by 9% from 2018. That is 3,484 billion users as at January 2019 (Kemp, 2019). To put this in context, an additional 288 million social media users are now contributing to the big data pool and healthcare content category. This user-generated content is created on various social media platforms. In the study, the researcher will explore which of the social media platforms are frequently engaged by the actors in the healthcare sector.

1.5.1 Social Media Platforms

There is a plethora of social media platforms, such as blogs, forums, business networks, photo sharing, social gaming, microblogs, chat apps and social networks (Kaplan and Haenlein, 2010). Social media platforms enable the establishment of networks, instantaneous dissemination of information and generation of user content (Gupta, Tyagi and Sharma, 2013). Monthly, people are active on the following key social media platforms: Facebook (2,167 million), You Tube (1,500 million), What's App (1,500 million), Facebook Messenger (1,300 million), We Chat (900 million) and Twitter (330 million) and to a lesser extent, platforms such as Reddit (250 million), WordPress, Photobucket and Topix are used (Kemp, 2019; Statista, 2018). For the purposes of this study, the researcher will collect source-data from Social Mention, over a 30-day period, to analyse which social media platforms Internet users are active on, in the context of the study.

1.5.2 Social Media Sentiment

"What other people think" has always been a vital part of information in any decision-making process. Before social media, we asked family members, friends and colleagues to recommend a doctor, a mechanic or even a restaurant. We requested reference letters from employers and turned to consumer reports when deciding to buy electrical equipment. The Internet and the various social media platforms have now made it possible to access the opinions, perspectives and experiences of a cosmic pool of people. Conversely, people are also able to express their opinions, share their experiences (good and bad) and articulate their perspectives (Pang and Lee, 2008; Gupta, Tyagi and Sharma, 2013). While this study will not analyse the sentiments of individuals on social media, it will assess the specific social media sentiments as per the Social Mention metrics.

1.5.3 Social Media Sentiment Analysis

According to Pang and Lee (2008), sentiment analysis deals with the computational management of what social media users are saying. The computational techniques vary, e.g. machine learning, data mining, natural language processing, information retrieval and database technique. These diverse disciplines are used to categorize the various sentiments. The polarities of sentiment are generally negative, positive or neutral (Yang, Huang and Wang, 2017). Understanding and the continuous analysis of social media sentiment is vital for firms, as this user-generated content can assist firms with making important strategic business decisions (Mukhopadhyay, 2018). In this study, the Social Mention sentiment analyses of 10 Pharmaceutical firms is followed and collected over a period of 30 days. This is to gauge how the individual firms have fared and to see how they stack up against each other. The next section looks at the various actors in the healthcare sector and how they impact on Pharmaceutical firms' social media strategies.

1.6 The Healthcare Industry

Digital technology has been on an exponential upsurge year on year and has had a profound impact universally, on everyday life and business, with healthcare being no exception (Aitken, 2014; Gupta, Tyagi and Sharma, 2013). This global sector is valued at 1,100 billion USD (Livinec, 2018; Statista, 2018). Social media has radically transformed the landscape of the healthcare industry. The way the actors in this sector, namely: Pharmaceutical firms, health care regulatory authorities, health care practitioners and patients, source and disseminate healthcare information, has been affected significantly (Hawkins, DeLaO and Hung, 2016; Unmetric, 2017).

1.6.1 Pharmaceutical Firms (Pharma)

Pharma has been dubbed the laggards in adopting social media, compared to other business sectors. Some of the reasons cited for this lag is the heavily regulated environment, loss of content control, privacy concerns and the measuring of social media return on investment (ROI) (Hunter, Gough, O'Kane, McKeown, Fitzpatrick, Walker, McKinley, Lee and Kee, 2018; Aitken, 2014). According to Cornejo (2018), of the 6 major social media platforms, namely: Facebook, Twitter, YouTube, LinkedIn, Instagram and Pinterest, as identified by Spitz and Einarsen (2017), Pharma is only present on 5.

With time, Pharma has embraced social media and even though their engagements are highly regulated, firms are increasingly adopting social media, to reach and interact with customers, potential new employees and Health Care Practitioners (HCPs). Given the very restricted environment, a recent report showed that Pharma is not just cautious about the social media platforms they are active on, they are also very selective with the information disseminated on social media. Pharma has divided their social media presence into four specific areas, i.e. corporate profiles, careers, over-the-counter brands and community pages (Narayan, 2017). While Pharma has been very selective regarding their social media presence, customers are constantly seeking, using, generating and disseminating information and data amongst each other online. Firms should be vigilant in scrutinizing this content, directing the dialogues and analyzing the data, to make informed business decisions (Fischbach and Zarzosa, 2018).

1.6.2 Regulatory Authorities

Some of the reasons cited for Pharma's lag in adopting social media, is the heavily regulated environment, loss of content control, privacy concerns and the measuring of social media return on investment (ROI). The borderless nature of the Internet and heterogeneous market regulations call for healthcare information that is consistent and that ensures that there is a stable environment for content contributors (Aitken, 2014; Ray, 2017). Table 1.1 lists the major regulatory authorities in various countries. Many of the Pharma firms are multinationals and therefore must comply with the various regulatory authorities in the different countries.

Table 1.1 List of Regulatory Authorities

No.	***Regulatory Authority***	***Country***
1	Food and Drug Administration (FDA) The Office of Prescription Drug Promotion (OPDP)	United States of America (USA)
2	International Federation of Pharmaceutical Manufacturers and Associations (IFPMA)	European Union (EU)
3	ABPI’s Code of Practice for the promotion of Prescription-Only Medicines (ABPI Code)	United Kingdom (UK)
4	National Medical Products Administration (NMPA)	China
5	Pharmaceuticals and Medical Devices Agency (PMDA)	Japan
6	South African Health Product Regulatory Authority (SAHPRA)	South Africa

In an Instagram post Kim Kardashian (reality TV celebrity), shared her experience of a Pharmaceutical product and mentioned that she was now partnering with the Pharmaceutical firm to raise awareness for the condition. The social media post had major repercussions for the firm. A letter of warning was issued to the firm to comply with regulatory advertising standards. The Food and Drug Administration (FDA) found her post to be misleading. The post was picked up by the FDA’s watchdog, The Office of Prescription Drug Promotion (OPDP) (Sullivan, 2018). Whilst this post was detected by the regulatory authorities in the US, its reach on social media was global.

1.6.3 Health Care Practitioners (HCPs)

Like Pharma, HCPs have also been dubbed the late adopters of social media (Aitken, 2014). HCPs have been slow in engaging on social media, mainly due to privacy issues, lack of trust, time constraints, information disorder and inadequate regulation (Ray, 2017; Hawkins, DeLaO and Hung, 2016; Panahi, Watson and Partridge, 2014). While several challenges have been highlighted, the benefits of being active on social media for HCPs is invaluable. HCPs can provide credible peer-reviewed, accurate, valid information that patients are looking for (Hawkins, DeLaO and Hung, 2016). Of the HCPs that have embraced social media, 90% use it for personal reasons and 65% use it for professional networking. They see social media as a key source to gain new knowledge, connect with peers, stay abreast with the latest research developments and integrate evidence-based material into their medical practice (Tutelman et al., 2018).

1.6.4 Patients

The various social media platforms have provided a venue for patients to, not only seek and find information, but has also provided a place for peer interaction. (Benetoli, Chen and Aslani, 2018). According to Aitken (2014), between 70 and 75 percent of online users are searching for over-all and specific health information, either for themselves, or their loved ones. The information that patients generally seek, ranges from chronic conditions, medical procedures or treatments, to lifestyle issues, medical test results and many other health related or medical problems. More than 40% of the patients share their personal health experiences online and this has provided a space where patients with similar health concerns are able to connect, support and help one another cope with their medical conditions (Li, Wang, Lin and Hajli, 2018). Koumpouros, Toulias and Koumpouros (2015) research revealed that patients are more proactive, participatory and mature in their engagement on social media, unlike Health Care Practitioners (HCPs).

1.7 Big Data

The greatest advantage of social media for Pharma, is the mega trend of big data. Big data is described as the colossal volumes of unstructured, semi-structured and structured data, available on social media and the Internet. Big data is characterized by the volume (amount), variety (formats), velocity (rate of data inflow), variability (dynamic nature), veracity (complexity), value (measure) and visualization (imaging). More than 40 trillion gigabytes of data would have been created, replicated and used by the year 2020 (Shah, Rabhi and Ray, 2015; Sivarajah, Kamal, Irani and Weerakkody, 2017; Günther, Mehrizi, Huysman and Feldberg, 2017). According to Mgudlwa and Iyamu (2018), social media platforms are the most rapidly growing sources of big data. Ray (2017) indicates that healthcare information is the fastest growing content and 1 in 20 Google searches is for health-related matters. While Aitken (2014) has identified big data as a challenge for Pharma, the McKinsey Global Institute (cited in Ray, 2017) has projected that the adoption of big data strategies, for sound decision-making, could generate 300 billion dollars per annum for the US healthcare economy (Manyika et al., 2011; Cattell, Chilikuri and Levy, 2013; Henke et al., 2016).

1.8 Big Data Analytics

The significance of big data is only realized when leveraged to implement sound business decision-making. The process of unearthing insights from big data, is big data analytics. It is the methods used to analyse and assimilate intelligence from big data (Gandomi and Haider, 2015). Recent research by Thirathon, Wieder, Matolcsy and Ossimitz (2017), has found that big data, or the ability to decipher big data, are not just hypes and that firms who have this capability can achieve superior performance, directly and indirectly. Wang and Hajli (2017) says that big data prompts firms into committing resources to big data analytics. This allows them to acquire valuable insights to facilitate timely decision-making, minimize risks, reduce costs and grow revenue. A McKinsey Global Institute report by Cattell, Chilikuri and Levy (2013), revealed that big data and big data analytics, is not only valuable to the sales and marketing functions, it has also been extended to other areas of business, such as research and development throughout industries. In the case of Pharma, big data and big data analytics allows them to innovate new treatment options rapidly.

Ray, Barney and Muhanna (2004) says that most of the RBT empirical literature examines the impact of firm resources on the overall performance of the firm.

They argue that adopting the effectiveness of a business process, as a dependent variable, may be more appropriate than adopting overall firm performance. Big data and big data analytics, as a business process, requires the different business units in a firm to collaborate and partner, to acquire the relevant resources and capabilities to derive value from the process (Janssen, van der Voort and Wahyudi, 2017). This study will explore how social media big data and analytics, as business processes, can be used to improve firm performance.

1.9 Conclusion

The key actors in the healthcare industry are Pharmaceutical firms (Pharma), regulatory authorities, HCPs and patients. While these entities form part of the Pharma's

external environment, they play a pivotal role in the generation and creation of big data on social media platforms. Equally, social media platforms and social media big data, forms an integral part of this external environment. Conversely, big data analytics is described in the literature, as an internal capability and therefore vital for the internal business operations environment. For firms to make sound business decisions, big data sourced externally, cannot be used in isolation of the firm's internal data. More importantly, firms should be collaborating across its various business units, to exploit the value of big data and big data analytics. This study will investigate how the amalgamation of external and internal resources and capabilities can lead to sustained competitive advantage.

In chapter two the RBT and DCT literature will be explored to investigate the concepts, both from an academic and business perspective. Chapter three will define the methodology. Chapter four will document the results and discuss the findings against the backdrop of the literature. Chapter five concludes the research and outlines the study's limitations and recommendations.

2 Chapter 2: Literature Review

2.1 Introduction

This chapter discusses RBT and DCT. According to Grant (1991), resources and capabilities, as the basis for a firms long term strategy, is based on two grounds: 1) internal capabilities and resources provide direction and 2) it is the prime sources of revenue. He says that in an environment where client profiles are changing, their preferences are volatile and technologies serving their needs, are constantly advancing and an external business focus is not ideal for long-term strategy formulation. A firm's internal capabilities and resources however, provide a better source to develop its strategy, when the external environment is volatile. Companies like Honda and 3M Corporation has been very successful in applying this strategy (Barney, 1986a; 1986b; 1986c; 1991; Lockett, O'Shea and Wright, 2008; Wernerfelt, 1984, 1989, 2013).

The aim of this study is to investigate how social media big data, as an external resource, can be utilized through internal firm capabilities, namely, big data analytics, to make better business decisions, to gain a sustained competitive advantage. The literature is examined from both an academic and a practical perspective.

2.2 Theoretical Framework

2.2.1 Background

During the 1980s, the focus of strategic management was mainly based on the external environment of a firm, to achieve a competitive advantage (Porter, 1980). Conversely, in the same epoch, Wernerfelt (1984) argued that firms should be looking at internal resources, to determine profitability. The balancing act of developing new resources and exploiting existing ones, leads to optimal growth for firms (Wernerfelt, 1984, 1989; Barney, 1991; Grant, 1991; Dierick and Cool, 1989). The resource-based view moved the attention of strategic analysis, from the industry, to the firm. It emphasized that firms' resources and internal capabilities are primarily diverse and should be inimitable. It suggests that resources and capabilities are the foundation of corporate strategies. This notion gave rise to RBT (Wernerfelt, 1984, 1989, 2013; Barney, 1991, 1995; Grant, 1991; Lockett, O'Shea and Wright, 2008). Wernerfelt goes on to say that a resource is any weakness or strength of an organization. He defines resources "as those tangible and intangible assets which are tied semi-permanently to the firm e.g. brand names, in-house knowledge of technology, employment of skilled personnel, trade contracts, machinery, efficient procedures, capital, etc." (Wernerfelt, 1984, p.172). Barney (1991, p.101) has defined firm resources as three categories namely:

i. "Physical capital resources … physical technology, plant and equipment, geographical location and access to raw materials."
ii. "Human capital resources … training, experience, judgement, intelligence, relationships, insight of individual managers and workers."
iii. "Organizational capital resources … firm's formal reporting structure, formal and informal planning, controlling and coordinating systems, informal relationships among groups within the firm and between a firm and those in its environment."

There are several definitions of resources in the literature. The most current definition is "***Resources are the tangible and intangible assets firms use to conceive of and implement their strategies*** (Barney and Arikan, 2008, p.138).

2.2.2 Impact of RBT

In a recent study Braganza, Brooks, Nepelski, Ali and Moro (2017) argues that RBT in the context of big data is challenged. The study presumes that big data is a tangible, homogeneous resource that is generally sourced externally. This data is available to any firm to acquire and use (free or for a fee). Therefore big data provides limited competitive advantage (Mata, Fuerst and Barney, 1995) . Gupta and George (2016) concurs that big data, as a resource on its own, is not likely to be a source of competitive advantage. As all firms will be aggregating masses of data from numerous sources and need an inimitable combination of its tangible and intangible resources to generate a capability that is hard for their competitors to match (Barney, 1991, 1995; Ray, Barney and Muhanna, 2004; Barney and Clark, 2007; Amit and Schoemaker, 1993). Wang and Hajli (2017) acknowledge that some of the literature contests that not all resources, as assumed by RBT, add value to a firm's operations (Kim, Shin, Kim and Lee, 2011; Grover, Chiang, Liang and Zhang, 2018). They rebut by saying that the DCT can be applied to supplement the gap in the RBT (Bharadwaj, 2000; Doherty and Terry, 2009; Lin and Wu, 2014).

2.2.3 RBT Framework

Figure 1.1 is a reproduction of the RBT Framework as depicted by (Barney, 1991, p.112) . Barney intended for firms to use the framework as a tool to analyse a wide variety of resources that are likely to be a source of repeated competitive advantage. He also articulates that the framework, not only be utilised to analyse resources from a theoretical perspective, but that particular empirical questions needs to be interrogated.

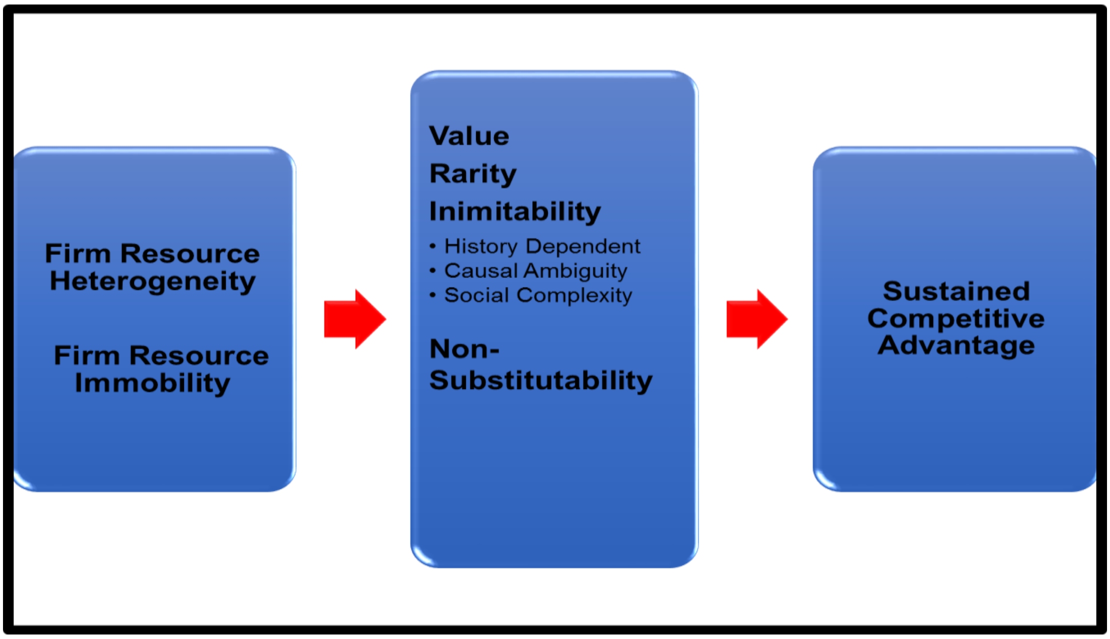

Figure 2.1 The relationship between resource heterogeneity and immobility; value, rarity, inimitability, non-substitutability and sustained competitive advantage.

2.2.3.1 Heterogeneity

Birger Wernerfelt is considered as the one of the founders of RBT. The field of strategic management has adopted several of his ideas, since the publication of his article in 1984. One of his fundamental ideas, embraced by the fraternity, is that firms are heterogeneous in nature and their resources are not and should not be identical (Lockett, O'Shea and Wright, 2008; Wernerfelt, 1984, 1989, 2013). Barney (1991) built on Wernerfelt's ideology by saying that if firms are homogeneous in the resources they manage, then it is not possible for any of them to earn a competitive advantage. This will merely give the firms competitive parity (Peteraf, 1993; Helfat and Peteraf, 2003; Barney and Arikan, 2008).

2.2.3.2 Immobility

Further to this, Barney (1991) reiterates that firms resources cannot be perfectly mobile across companies, for its resources to be a source of sustained competitive advantage. Several studies concur that resources cannot be traded, if it is imperfectly mobile (Dierick and Cool, 1989; Peteraf, 1993; Helfat, Finkelstein, Mitchell, Peteraf, Singh, Teece, Winter and Maritan, 2007; Barney and Arikan, 2008).

2.2.3.3 Value, Rarity, Inimitability and Non-Substitutability (VRIN Framework)

RBT gained traction when Barney (1986a; 1986b; 1986c; 1991) identified four empirical indicators of resources, namely: value, rarity, inimitability and non-substitutability (VRIN Framework). The VRIN framework analyses under which conditions resources will give rise to superior firm performance.

1. **Value**

According to LaValle, Lesser, Shockley, Hopkins and Kruschwitz (2011) in all business sectors, executive leaders are questioning whether they are getting full value from the resources available in their firms. Many of the firms are trying to find ways to acquire value from their resources, to compete in the marketplace. RBT says resources can only be valuable when it allows firms to generate or execute strategies that improves its competences and effectiveness by taking advantage of opportunities and neutralizing threats. Valuable resources only give firms competitive parity (Barney, 1991; Barney and Clark, 2007; Wang and Hajli, 2017).

2. Rarity

Valuable resources, owned by several competing or potentially competing firms is not a source of competitive advantage or sustained competitive advantage. In this situation, firms can exploit the resources, in the same way, hence merely creating temporary competitive advantage. Thus only increasing their chance of economic survival. When the number of firms that possess a particular resource is less than what is required to implement business strategies to create perfect competition the resource is rare (Barney, 1991; Barney and Arikan, 2008).

3. Inimitability

Even though resources may be valuable and rare, it can only be a source of sustained competitive advantage, if other firms that do not have it, cannot acquire it (Barney, 1986a; 1986b; 1986c; 1991; 1995; Barney and Arikan, 2008). Firms' resources can be inimitable for three reasons:

i. Historical Uniqueness e.g. organizational culture
ii. Causal Ambiguity e.g. when the link between firm resources and sustained competitive advantage cannot be pinpointed
iii. Social Complexity e.g. interpersonal relationships among staff, with firm's culture, suppliers and customers

4. Non-Substitutability

There should be no similar resources that rival firms could use to create and apply equivalent strategies (Barney and Clark, 2007; Barney, 1991; Peteraf, 1993).

2.2.3.4 Sustained Competitive Advantage

The RBT has been used as a viewpoint to comprehend the correlation between resources and/or capabilities and sustained competitive advantage (Yang, 2008). Table 2.1 depicts what performance outcomes leaders can expect from resources, relative to the resource attributes (Barney, 1991; Barney and Arikan, 2008).

Table 2.1 Resource Attributes relative to performance

	Resource Attributes	Resource Performance
Resources	Valuable	Competitive Parity
	Valuable and Rare	Temporary Competitive Advantage
	Valuable, Rare and Inimitable	Temporary Competitive Advantage
	Valuable, Rare, Inimitable and Non-Substitutable Firm organised to capture and exploit resources with the relevant attributes	Sustained Competitive Advantage

Sustained competitive advantage is not generated by simply analysing environmental opportunities and threats. Generating sustained competitive advantage hinges on the unique internal resources and capabilities a firm brings to the marketplace. To identify these resources and competencies, leaders must apply the VRIN framework internally and exploit their resources (Barney, 1995; Mata, Fuerst and Barney, 1995; Barney and Arikan, 2008).

2.2.4 A Resource-Based Strategy Analysis

Further to this, Grant (1991, p.114) created a five-phased procedure, which is a practical approach to resource-based strategy analysis, i.e. assessing resources, evaluating capabilities, analysing income potential of resources and capabilities, choosing a plan and the enhancing and changing of resources and capabilities. He considers strategy as “the match an organization makes between its internal resources and skills … and the opportunities and risks created by its external environment.” Figure 1.2 is a flowchart of : A Resource-

Based Approach to Strategy Analysis: A Practical Framework (Grant, 1991, p.115).

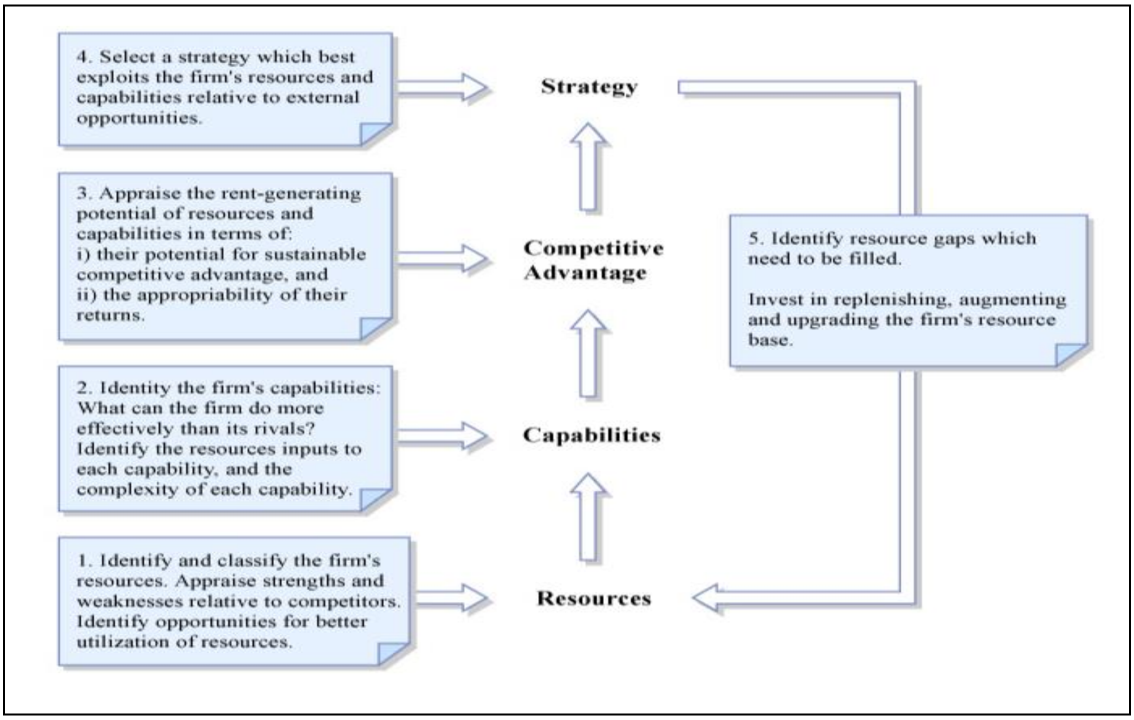

Figure 2.2 A Resource-Based Approach to Strategy Analysis: A Practical Framework

2.3 Dynamic Capability Theory (DCT)

Following the proliferation of RBT literature, several sub-fields arose. These sub-fields differ slightly in characterizing firm attributes, but underpin the same theoretical rationale of RBT. This has seen the emergence of the DCT (Teece, Pisano and Shuen, 1997; Barney and Arikan, 2008; Barney, Ketchen and Wright, 2011; Helfat, 2011; Gupta and George, 2016). Hill and Jones and Hitt, Ireland and Hoskisson (1992,1999 cited in Barney and Arikan, 2008, p.139) defines capabilities as "…***those attributes of a firm that enable it to exploit its resources in implementing strategies."*** Teece, et al. (1997) describes it as a firm's capacity to develop, reconstruct and adapt internal and external capabilities, to address the rapidly changing business environment. In a more recent study, Teece (2018), in line with Grant (1991)'s definition on strategy, says that the Internet has created prospects for radical, innovative business models, to which firms' strategies must retort. He goes on to say that the

Internet is a source of big data and it creates an innovative type of capability, that can be used for the internal and external business environment.

2.3.1 DCT Framework

Teece (2018) mentions that RBT focuses on bringing together resources that meet the VRIN attributes to derive sustainable competitive advantage. Teece argues that this will not suffice, as these resources must support a sound strategy and an all-encompassing business model (Teece, 2010; Casadesus-Masanell and Ricart, 2011). In his latest study Teece (2018, p.44) developed a simplified version of the dynamic capabilities framework. Emphasizing that dynamic capabilities and strategy together, generate and improve business models that enable firms to maintain and develop capabilities and resources that leads to sustained competitive advantage. Leaders who can develop their sense, seize and transform capabilities, in other words leaders with superior dynamic capabilities, can adjust and transform quicker to a dynamic business environment and can therefore stimulate innovation in firms (Helfat and Martin, 2015; Kurtmollaiev, Pedersen, Fjuk and Kvale, 2018). Figure 1.3 is an illustration how synergy between superior dynamic capabilities, business models and strategy can lead to sustained competitive advantage.

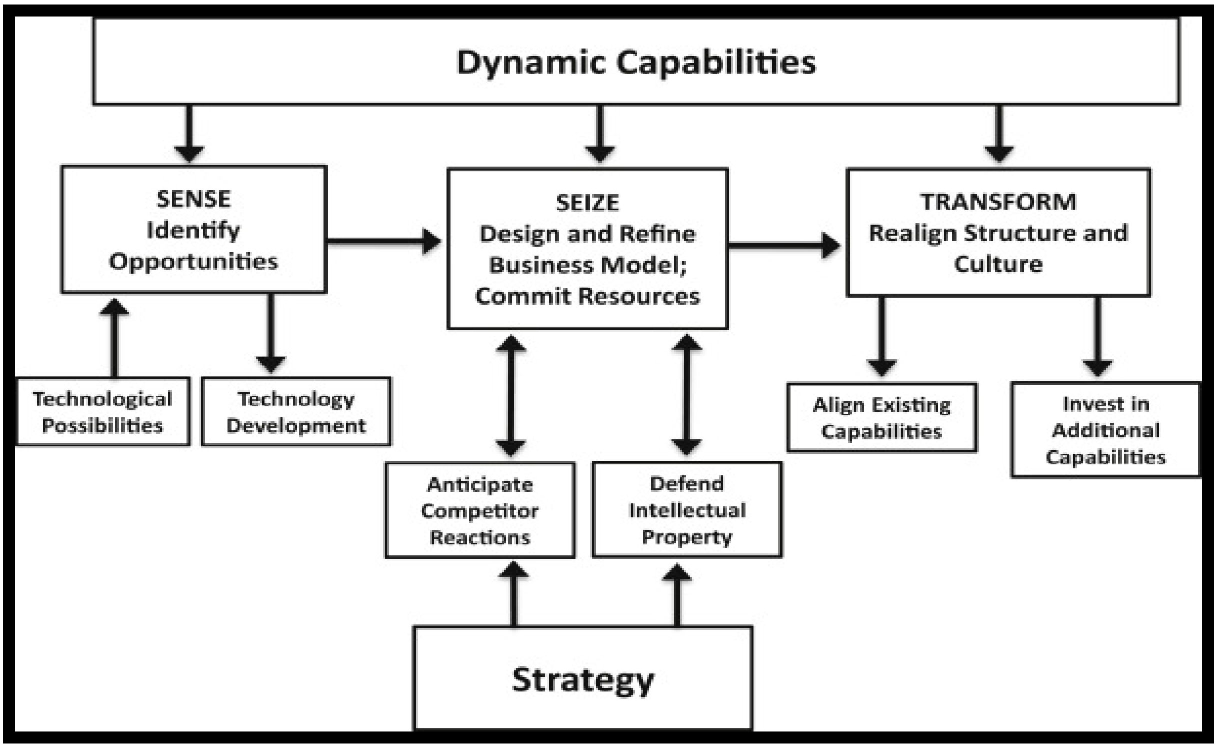

Figure 2.3 Dynamic Capabilities, Business Models and Strategy

2.4 Big Data and RBT

2.4.1 Big Data Heterogeneity and Immobility

According to Mgudlwa and Iyamu (2018) social media is the fastest growing sources of big data. Ray (2017) indicates that healthcare is the fastest growing content category and that 1 in 20 Google searches is for health-related content. The greatest advantage of social media for Pharma, is the mega trend of big data. The McKinsey Global Institute eluded that big data is also now considered as a factor of production across all business sectors (Manyika, et al., 2011; Henke, et al., 2016). Big data is described as the massive volumes of unstructured, semi-structured and structured data, available on social media and the Internet (Schroeder, 2014; Janssen, van der Voort and Wahyudi, 2017). While the sources of big data are considered heterogeneous and perfectly mobile (Gandomi and Haider, 2015). Data as a resource is homogeneous and perfectly mobile therefore in the context of RBT may not be considered as resource for superior performance on its own (Grant, 1991; Barney, 1991; Gupta and George, 2016; Wang and Hajli, 2017; Braganza, et al., 2017).

2.4.2 Big Data Value

While Aitken (2014) has identified big data as a challenge for Pharma, the McKinsey Global Institute (cited in Ray, 2017) has projected that the adoption of big data strategies, for sound decision-making, could generate 100 billion dollars per annum for the US healthcare economy. Côrte-Real, Oliveira and Ruivo (2017) agrees that big data can provide vital insights and competitive advantage with the appropriate technological and firm resources. Günther, et al. (2017) states that big data is valuable socially if utilised by organisations to assist people with better healthcare (Cech, Spaulding and Cazier, 2015; Raghupathi and Raghupathi, 2014) and economically when firms can determine profit growth, business expansion and competitive advantage (LaValle, et al., 2011; McAfee, Brynjolfsson and Dearstyne, 2012). Koscielniak and Puto, (2015) assert that quality decision making processes is becoming a crucial compoment of leadership and that big data is not just about amassing information, the greateast value lies in the processing and visualisation of it.

2.4.3 Big Data Rarity

Braganza, et al. (2017) posits that big data is a challenge to the RBT and its VRIN attributes. The data in their study was sourced from several external providers therefore making it assessible to any firm. Concluding that when one of the VRIN attributes is not present, a resource like big data as a source of competitive advantage is constrained. Citing the research of (Mata, Fuerst and Barney, 1995; Delmonte, 2003) to corroborate their claim. Gupta and George (2016) and Gandomi and Haider (2015) concur that big data in isolation, is questionable as a source of competitive advantage, as firms of akin scope would be aggregating data from various sources (for free or a fee). Therefore, the fundamental resource, data, is not rare.

2.4.4 Big Data Inimitability

The literature is clear that big data in isolation is not inimitable and that it is big data analytics that enable firms to derive value from the big data. Gupta and

George (2016) says the crucial question is how do firms generate big data capabilities to achieve sustained competitive advantage. In the research done by Teece posits that it is a firm's dynamic capability that will enable them to generate capabilities that their rivals will find hard to copy (Teece, et al., 1997; Helfat, 2011; Teece, 2014; 2015; 2018). In section 2.5 this concept is explored further.

2.4.5 Is Big Data Substitutable?

The literature indicates that big data in the current era is not substitible however, to make critical business decisions big data should be used in conjuction with sound business experience and intuition (Davenport, 2013; Mallinger, 2015; Gaudiano, 2017) . LaValle, et al. (2011, p.23) research showed that high performing firms applied analytics to all business operations within the company, compared to low performing companies who applied mainly intuition see figure 1.4:

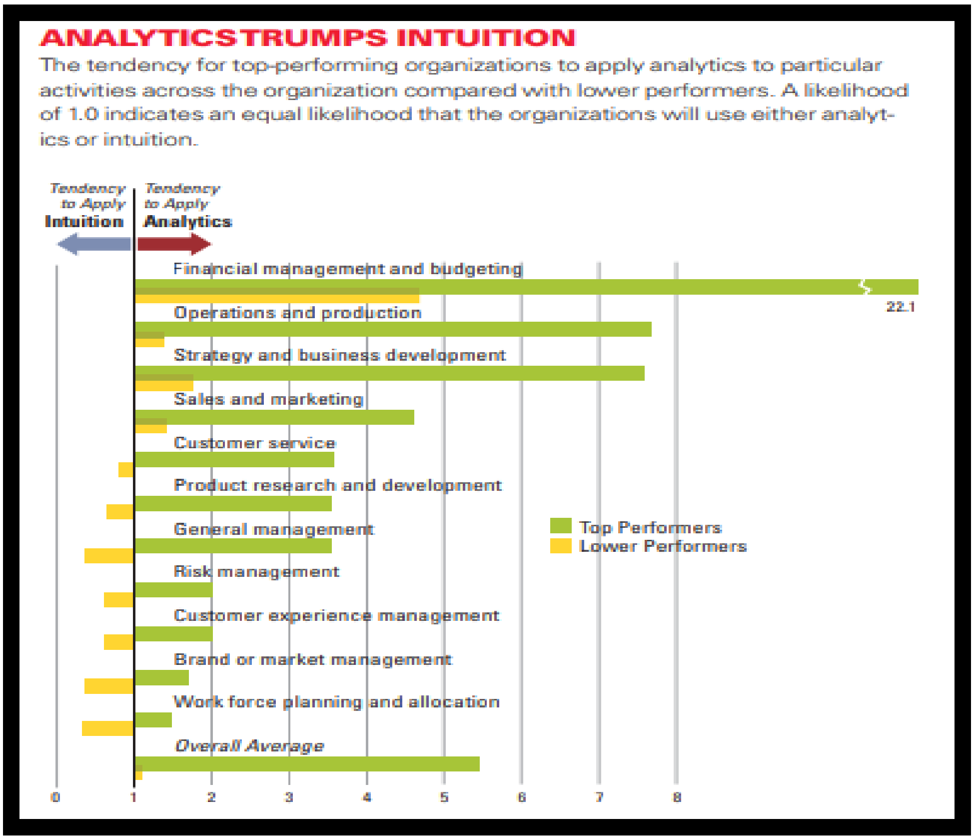

Figure 2.4 Anaytics Trumps Intuition

2.4.6 External Big Data

Data aggregated from external sources like the Internet, social media, mobile devices and sensors are deemed external data. One could argue that although external data is not rare, it has an element of value for firms (Gandomi and Haider, 2015). Even though the data may not be directly linked to the firm's internal business processes and procedures, it does offer an innovative and more flexible viewpoint than internal data (Zhao, Fan and Hu, 2014; Grant, 1991; Gandomi and Haider, 2015; Ram, Zhang and Koronios, 2016) .

2.4.7 Internal Big Data

Data generated as a result of a firm's internal business processes and procedures like stock records, sales information, people management, financial management and accounting transactions, is referred to as internal data (firm-specific). Firms who rely solely on their internal data to make business decisions, is not likely to achieve a competitive advantage. The balance a firm crafts between its internal resources and skills and the opportunities and risks produced by the external environment, is more likely to achieve sustained competitive advantage (Grant, 1991; Zhao, Fan and Hu, 2014; Ram, Zhang and Koronios, 2016).

Marr (2015) asserts that the big challenge for business leaders does not pertain to big data attributes, but relates to how these leaders will integrate their internal and external data to make sound business decisions. According to Gupta and George (2016) there has been an overemphasis on technical aspects of big data in the literature and not sufficient on other resources, such as human skills and organizational culture. LaValle, et al. (2011) emphasize the big data adoption challenges practitioners face, is not related to data and technology, but managerial and cultural. Mazzei and Noble (2017) agrees by saying that the literature concentrates on how big data will influence management research, instead of investigating how big data is revolutionising the critical thinking processes of business leaders.

Big data is categorized as a tangible asset, typified by sources, technology, data specialists, statisticians and internal firm structures (Braganza, et al., 2017). RBT does not make a clear distinction between resources and capabilities. Resources are assets owned and controlled by the firm, while capabilities are unique resources that allow firms to combine and utilize their resources jointly to make sound business decisions (Amit and Schoemaker, 1993; Makadok, 2001; Mazzei and Noble, 2017). It is apparent in the literature that big data on its own is not a source of competitive advantage for firms. It is the firm's ability in utilising big data analytics, that will lead to superior performance (Gupta and George, 2016; Wang and Hajli, 2017; Braganza, et al., 2017; Teece, 2018).

2.5 Big Data Analytics (BDA) and DCT

Internal and external data is essential, but not adequate for firms to derive a sustained competitive advantage over their rivals. It is therefore imperative that firms assess tangible and intangible resources to create a capability which is difficult for their rivals to match (Amit and Schoemaker, 1993; Barney, 1991; 1995; Marr, 2015; Gupta and George, 2016; Kurtmollaiev, et al., 2018). Most importantly, firms must constantly reconfigure their resources in line with the dynamic business environment to ensure superior performance (Teece, Pisano and Shuen, 1997; Helfat, et al., 2007; Teece, 2014; 2018).

2.5.1 Big Data Analytics (BDA) Capability

The value of big data is only realized when leveraged to drive decision-making. Skowronek-Mielczarek (2004 cited in Koscielniak, 2015, p.1054) says decision making is strengthened by four attributes namely:

- Focus (on areas the firm is aquainted with hence generating distinctive capabilities)
- Fast (react swiftly to signals from the external and internal environment)
- First (satisfying the customer need)
- Flexibility (in adapting and augmenting firm resources and operations)

Miller and Waller (2003) concur that the Skowronek-Mielczarek rule of 4F for quality decision making allows for firms to develop effective operating procedures. This enables internal future decisions to be made easily and quickly and minimise the margin of decision errors. Most profoundly decisions no longer have to be made based on instinct or in the dark. It can now be based on evidence, more precise forecasts and experiments (Henke, et al., 2016).

The process of unearthing insights from big data, is big data analytics. It is the methods used to analyse and assimilate intelligence from big data (Gandomi and Haider, 2015). LaValle, et al. (2011) reiterates that insights and intelligence derived from big data is not sufficient to achieve sustained competive advantage. They assert that insights and intelligence, instituted to drive action, will bring value. Ray, Barney and Muhanna (2004) says that most of the RBT empirical literature examines the impact of firm resources, on the overall performance of the firm. They argue that adopting the effectiveness of a business process, as a dependent variable, may be more appropriate than adopting overall firm performance. Big data and big data analytics, as a business process, requires the different business units in a firm to collaborate and partner, to acquire the relevant resources and capabilities to improve operations and derive value from the process (LaValle, et al., 2011 ; Janssen, van der Voort and Wahyudi, 2017).

Drawing on the original and later RBT research (Wernerfelt, 1984; Barney, 1991; Grant, 1991; Teece, et al., 1997; Teece, 2015) and research that accentuated the constraints of big data (McAfee, Brynjolfsson and Dearstyne, 2012; Chen, Chen, Du, Li, Ly, Zhao and Zhou, 2013; Kaisler, Armour, Espinasa and Money, 2013; Ross, Beath and Quaadgras, 2013; Davenport, 2014; Zhao, Fan and Hu, 2014), Gupta and George (2016, p.1051) identified seven resources that will enable firms to generate a BDA capability. They believe that separately or even per classification (as per the three classifications), resources cannot produce a BDA capability. It is in the inimitable blend of these resources, that will lead to a firm's specific BDA capability. Highlighting two intangible resources namely: data driven culture and intensity of organizational learning, as the most critical resources . Figure 2.4 is a diagram of the classification of the big data resources required to build a BDA capability.

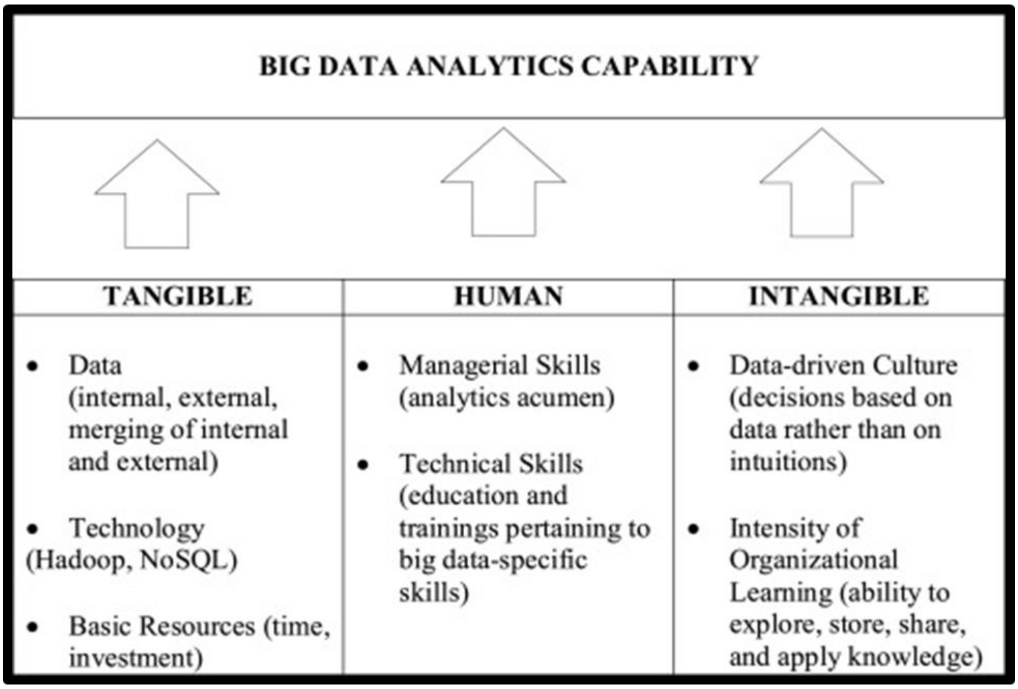

Figure 2.5 Classification of Big Data Resources

Wamba, Gunasekaran, Akter Ren, Dubey and Childe (2017) revealed in their research that BDA capability can be leveraged as a source of sustained competitive advantage and that the underlying RBT resources are generally the same as cited by Gupta and George (2016).

2.5.2 BDA Capability as a business model

RBT focuses on bringing together resources that meet the VRIN attributes to derive sustainable competitive advantage. However, this will not suffice, as these resources have to back a sound corporate strategy and an all-encompassing internal business model to derive long term value (Teece, 2010; Casadesus-Masanell and Ricart, 2011; Teece, 2018). A business model is the premise from which customers are served and how revenue is generated. It directs technological innovation and knowhow in conjunction with the application of tangible and intangible resources, into revenue. The following list relects the relevant components of a business model developed by Teece (2018, p.41), namely:

- Value Proposition: products and service; customer needs; geography
- Revenue Model: pricing logic; channels; customer interaction
- Cost Model: core assets and capabilities; core activities; partner network

The most critical issue when developing a business model is how the pertinent factors establish distinction from rivals. Firms with superior dynamic capabilities are inimitable because their dynamic capabilities are built on distinctive charateristics e.g. entrepreneural leaders, history-honed firm-specific routines, organisational culture and a degree of complexity. While business models can eventually be imitated, VRIN resources and superior dynamic capabilties (entrenched deep in the firm and not only with the executive leaders) is the foundation for sustainable competitive advantage. The Internet has made innovative business models possible. Furthermore the accessibility of the enormous volumes of data from social media platforms, gives rise to a new form of intellectual capital. This can be exploited or transacted with, for internal innovation or external collaboration (Helfat and Martin, 2015; Teece, 2018).

For firms to develop BDA capability as a business model, the BDA capabilities as defined by (Gupta and George, 2016) and superior dynamic capabilities as defined by (Teece, Pisano and Shuen, 1997; Helfat, 2011; Teece, 2018) is imperative.

2.6 Social Media and Pharma

According Gandomi and Haider (2015) social media analytics is an emerging field that has advanced in the 2000s after the dawn of the Web2.0. While social media analytics span across several fields, marketing has been the prime application. In a recent study Kapoor, Tamilmani, Rana, Patil, Dwevedi and Nerur (2018) found that of the 132 social media journal articles they reviewed, many of the authors have not focused on defining social media and that only 9 of the authors came close to defining and clarifying the concept. Having considered the definitions in the literature and based on Kapoor et al. (2018) understanding of social media, they proposed the following definition: ***"Social media is made up of various user-driven platforms that facilitate diffusion of compelling content, dialogue creation, and communication to a broader audience. It is essentially a digital space created by the people and for the people and provides an environment that is conducive for interactions and networking to occur at different levels (for instance, personal, professional, business, marketing, political, and societal)".***

The field of social media, big data and big data analytics, have become increasingly important to both academics and practitioners. According to a report by the McKinsey Global Institute by 2018 the US will face a dearth of 140 000 to 190 000 people with deep analytical skills and 1,5 million supervisors with the skill to analyse big data to make sound business decisions (Manyika, et al.,). One of the marvels of today, changing the globe as we are acquainted with it, is the global accessibility to the Internet. According to Kemp (2018) as at January 2018, of the total population of 7,593 billion people, 4,021 billion people are Internet users and 3,196 billion are active on social media. The annual (2017 – 2018) digital growth for Internet users has increased by 7% (248 million users) and social media active users by 13% (362 million users). Social media has radically transformed the landscape of the healthcare industry. The way Pharmaceutical firms, healthcare regulatory authorities, healthcare practitioners, and patients' source and disseminate healthcare information, has been affected significantly (Hawkins, DeLaO and Hung, 2016; Unmetric, 2017).

2.7 Conclusion

The RBT and DCT has illuminated that Pharmaceutical firms will require the following resources, to leverage social media big data as a source of sustained competitive advantage namely:

- tangible resources e.g. internal and external data, merging of the data, technology, time and investment

- intangible resources e.g. data driven culture (firms decisions are based on data rather than intuition) and organisational learning (a firms ability to explore,store, share and apply knowledge)

- capabilities such as managerial (data analytics acumen) and technical skills (big data specific training).

What is apparent in the literature, is that social media big data on its own, is not a source of sustained competitive advantage for firms. This advantage lies in the firm's ability to turn the data into useful, business intelligence and then take the relevant action to realise this value.

The next chapter speaks to the methods and procedures employed, to turn social media big data into valuable insights, for decision making and implementation.

3 Chapter 3: Methodology

3.1 Introduction

This chapter speaks to the research methodology that was followed for the study. It commences by stating the goals, aims and method of the research. Followed by a detailed explanation of the sampling technique and size and how the data was collected and analyzed. It also indicates how the Chernoff Faces has been generated in Stata 15. Stata 15 is a general-purpose statistical analysis software package, created in 1985 by StataCorp. It is used by many businesses and academic institutions around the world. Most of its users work in research, especially in the fields of economics, sociology, political science, biomedicine and epidemiology (Longest, 2015).

3.2 Research Goals and Aims

The research goals are to examine how Pharmaceutical firms, through the lens of RBT and DCT, can make strategic business decisions, that could lead to a sustained competitive advantage, by identifying:

a. The key aspects of RBT, that will assist firms in adopting internal business strategies, that can lead to sustainable competitive advantage and
b. The Dynamic Capabilities required to effectively leverage big data to achieve a sustained competitive advantage

The research illustrates how Pharmaceutical firms can use existing social media platforms, as an external source, to gather big data to analyse their social media sentiment, by utilizing:

a. A social media platform namely: Social Mention. Social Mention is a social media platform that collects various variables of social media sentiment, from more than 80 social media sites.
b. A graphical technique, ‘Chernoff Faces’ to analyse by visualizing the various social media sentiment variables.

3.3 Research Method

Considering the research aims and goals, a mixed method will be applied to the study. A quantitative approach will be utilized to analyse the social media sentiment as depicted by Social Mention metrics and interpret through visualization of the Chernoff Faces. A qualitative approach will be used to examine the RBT (VRIN Framework) and DCT to describe and understand the relevant theories and to build upon the quantitative results (de Vos, Strydom, Fouche and Delport, 2011).

3.4 Population and Sample

According to Dezzani (2018), the global diabetes therapeutics market in 2017 was estimated at 33 billion US dollars and projected that it would grow at 7.5% to reach 44.53 billion US dollars by 2021. The top 10 diabetes drugs accounted for revenue of 29.92 billion US dollars in 2017. The researcher will use the top 10 antidiabetic Pharmaceutical firms for 2018, as a sample for the study. Dezzani (2018) has determined the top 10 firms based on their annual revenue, by gathering data from their websites, annual reports and Securities and Exchange Commission (SEC) filings. Figure 3.1 lists the top 10 firms in the sample size (Dezzani, 2018):

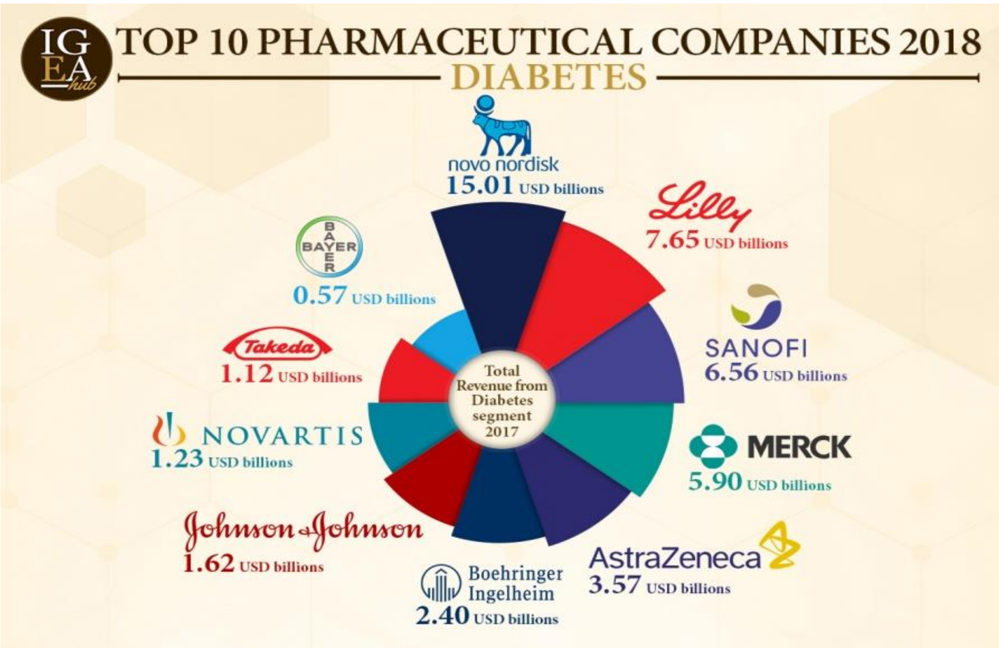

Figure 3.1 Top 10 diabetes Pharmaceutical companies 2018 (Dezzani, 2018)

3.5 Data Collection

Due to the dynamic nature of social media, the researcher has conducted a diary study, to gauge the user's sentiment. It is a method of collecting data daily and offers a mode of scrutinizing the data as it fluctuates (Ohly, Sonnentag, Niessen and Zapf, 2010).

3.5.1 Data Source - Social Mention

The data collection source for the study is 'Social Mention', a real-time social media search and analysis platform. The platform amasses user generated content from more than 80 social media platforms into one source of information. The researcher tracked and measured the social media sentiment (what users are saying about the 10 firms) daily, for thirty days. The following variables, used by Social Mention, namely: Strength, Positive and Negative Mention, Passion, Reach, Unique Authors and Relative Frequency were adopted for the study (Social Mention, 2018). Table 3.1 explains the metrics.

Table 3.1 Social Mention Metrics (Social Mention, 2018)

Social Mention Metrics	***Metric Definitions***	***Metric Calculation***
Strength	The likelihood that the brand is being discussed on social media	Phrase mentions within last 24 hrs divided by the total possible mentions
Sentiment	Ratio of positive mentions compared to negative	Gauged in absolute terms i.e. positive, negative and neutral mentions
Passion	Passion is a measure of the likelihood that individuals talking about your brand will do so repeatedly.	Advocates who talk about the firm all the time will have a higher Passion score. Mentions written by a different author have a lower score. Most frequently used keywords and number of times mentioned. Number of mentions by sentiment.
Reach	Reach is a measure of the range of influence.	It is the number of unique authors referencing your brand divided by the total number of mentions.
Unique Authors	No of authors messaging about the firm	Number of authors messaging within a specific time period
Frequency	The frequency with which mentions appear	Measured in time 1-6 hrs =100, 7-12hrs= 50, 12-24hrs=25 and 25-48hrs= 14,5
Sources	Social media platforms	Number of social media platforms and number of mentions per platform

3.5.2 Data Validity and Reliability

A naming protocol was established for each company and this was consistently applied during the data collection period. The data was downloaded from Social Mention, by inserting the firm's name (as per the established naming protocol) in the website. The website then calculates the metrics for each firm, based on all the social media platforms it is linked to, producing a score for each metric. The daily results were recorded with the Windows screenshot utility namely: Snipping Tool (Fischbach and Zarzosa, 2018; Syrdal and Briggs, 2018). Figure 3.2 is an illustration of the way the result was recorded. It shows the Social Mention metric results for all the variables, the naming convention and the date and time the data was collected.

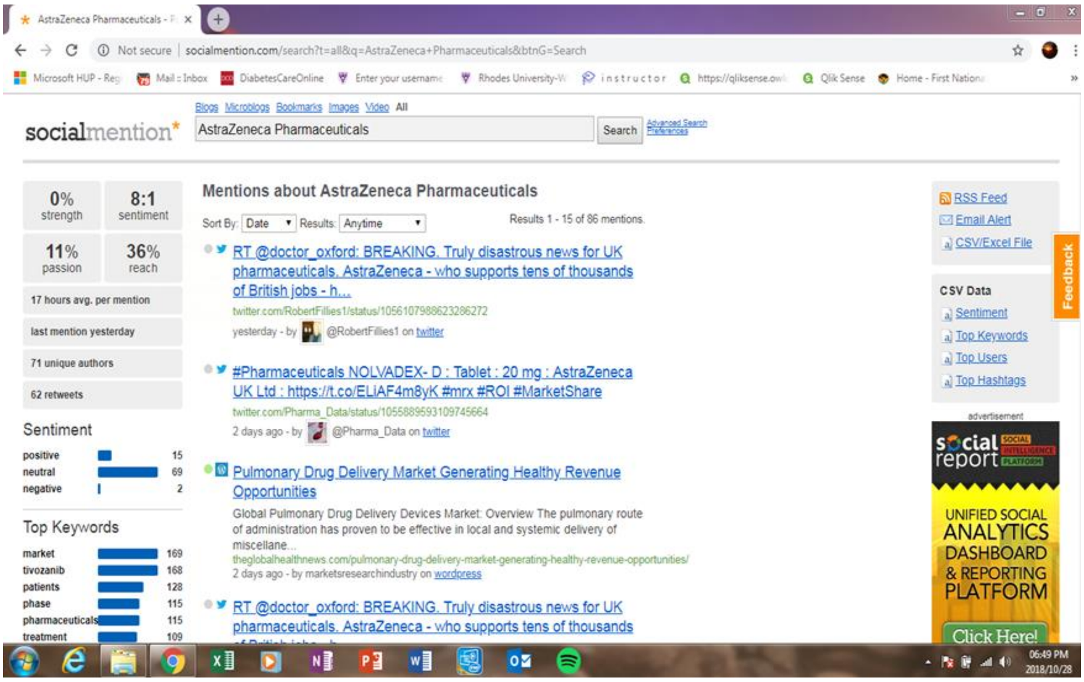

Figure 3.2 Social Mention Variables (Social Mention, 2018)

Figure 3.3 shows the social media platforms that the data was generated from for the day. It also shows the number of mentions per social media platform.

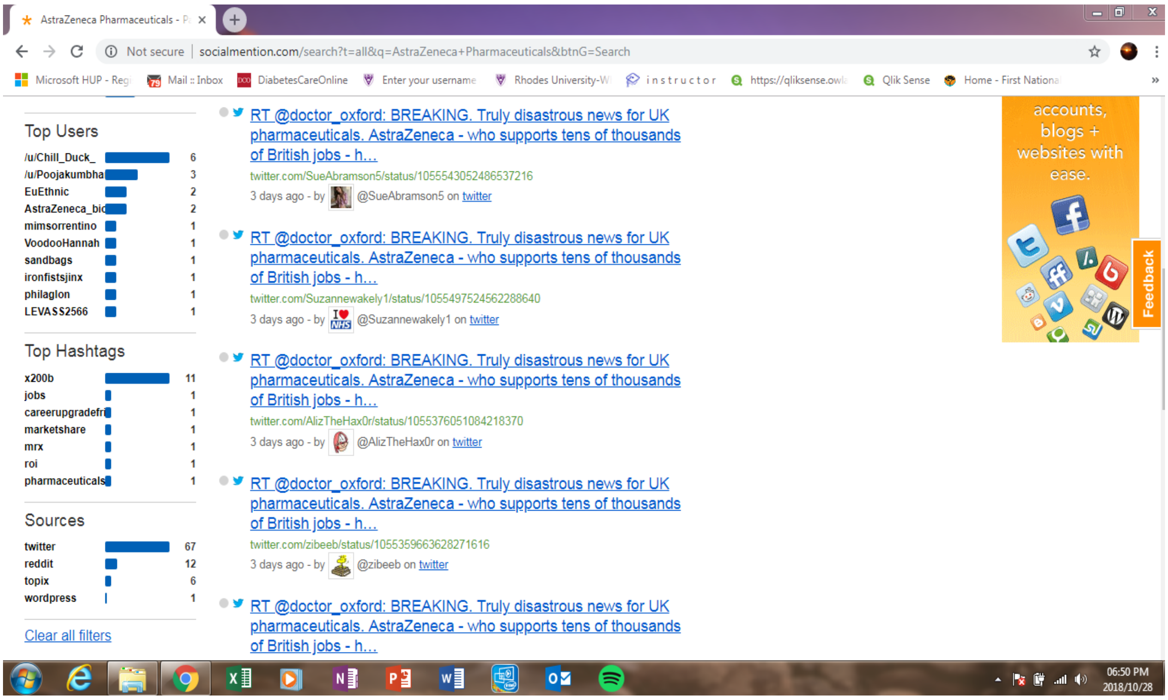

Figure 3.3 Social Mention Sources – Social Media Platforms (Social Mention, 2018)

The data for each firm was downloaded daily, as described above, saved and stored electronically, for purposes of an audit trail. An Excel spreadsheet was created with a worksheet for each firm. The results for each firm was captured onto the specific firm worksheets. With the Social Mention metrics as columns and the days of the month as rows. The scores of each firm was tabulated as per table 3.2.

Table 3.2 Daily Social Mention Metric Scores

AstraZeneca Pharmaceuticals							
Date	Positive Sentiment	Negative Sentiment	Strength	Passion	Reach	Relative Frequency	Unique Authors
05-Oct-18	18	0	0	54	5	14.5	5
06-Oct-18	19	2	3	33	18	100	26
07-Oct-18	18	2	0	57	6	50	6
08-Oct-18	18	2	0	61	5	25	5
09-Oct-18	19	2	17	40	20	100	30
10-Oct-18	19	2	1	57	4	25	6
11-Oct-18	18	3	2	57	6	25	6
12-Oct-18	18	3	0	53	5	25	7
13-Oct-18	18	3	4	40	34	100	51
14-Oct-18	18	3	0	40	32	25	48
15-Oct-18	19	3	51	6	54	100	107
16-Oct-18	19	4	0	44	7	50	10
17-Oct-18	20	3	50	7	53	100	106
18-Oct-18	18	3	1	50	5	25	7
19-Oct-18	16	3	2	42	6	25	8
20-Oct-18	16	3	7	7	54	100	107
21-Oct-18	16	3	1	42	6	25	8
22-Oct-18							
23-Oct-18	15	3	2	40	6	50	9
24-Oct-18	16	3	1	6	53	50	105
25-Oct-18	17	3	5	8	69	100	103
26-Oct-18	15	2	2	8	68	50	101
27-Oct-18	16	2	1	8	52	50	103
28-Oct-18	15	2	0	11	36	25	71
29-Oct-18	14	2	2	15	23	100	45
30-Oct-18	14	2	0	53	4	14.5	6
31-Oct-18	13	2	0	58	5	14.5	5
01-Nov-18	13	2	2	18	18	100	35
02-Nov-18	13	2	0	63	4	14.5	4
03-Nov-18	13	2	0	25	12	25	24
04-Nov-18	13	2	0	58	4	14.5	5
31	30	30	30	30	30	30	30

Table 3.3 Daily Social Mention Sources – Social Media Platforms and Number of Mentions (Social Mention, 2018)

Date	Sources				
	Twitter	WordPress	Topix	Reddit	Photobucket
05-Oct-18	0	0	9	11	
06-Oct-18	25	0	9	14	
07-Oct-18			13	14	
08-Oct-18			13	13	
09-Oct-18	37		13	13	
10-Oct-18		1	13	13	
11-Oct-18			13	14	
12-Oct-18		1	13	14	
13-Oct-18	71		13	14	
14-Oct-18	67		13	14	
15-Oct-18	100	1	13	14	
16-Oct-18		4	13	14	
17-Oct-18	99	1	13	15	
18-Oct-18		1	12	13	
19-Oct-18		1	8	13	
20-Oct-18	100	3	8	13	
21-Oct-18		1	8	13	
22-Oct-18					
23-Oct-18		3	8	12	
24-Oct-18	98	2	8	12	
25-Oct-18	98		12	14	
26-Oct-18	98		7	12	
27-Oct-18	98	2	7	12	
28-Oct-18	67	1	6	12	
29-Oct-18	40	1	6	12	
30-Oct-18		1	6	12	
31-Oct-18			6	12	
01-Nov-18	29	3	6	11	
02-Nov-18			6	11	
03-Nov-18	20	1	6	11	
04-Nov-18		1	6	11	
	1047	29	287	383	
31	16	20	30	30	0

3.6 Data Analysis and Visualisation

The researcher assessed the individual firms and how they performed in comparison to each other. Analysing the tables proved to be cumbersome, therefore, a statistical software program, known as Stata 15, was used for the data analysis. A tool namely: Chernoff Faces, depicting a more graphical representation of how the data was applied. Table 3.4 shows the Stata syntax applied for the Social Mention variables. The Social Mention sources datasets were analyzed by applying descriptive statistics in Stata 15.

3.6.1 Chernoff Faces

Over the years, statistical data has been displayed in numerous ways, e.g. scatter diagrams, pie charts, histograms and bar charts. Several researchers, as cited in Farshid, Chan and Nel (2012, p.188) studied images as a method of presenting multivariate data. A typical graph generally depicts absolute numerical data, whereas, images are designed to spot clusters, categorize and arrange variables. One of the images studied, is a facial technique by Chernoff (1973). Chernoff Faces is a graphical representation of multivariate data, depicted by animated faces. One of the characteristics of big data is visualization. Data visualization is an effort to assist readers in recognizing the significance of data by placing it in a visual context. Clusters, patterns, trends and correlations that might go unnoticed in text-based data can be exposed and recognized effortless with data visualization (Shah, Rabhi and Ray, 2015; Sivarajah, Kamal, Irani and Weerakkody, 2017; Günther, Mehrizi, Huysman and Feldberg, 2017). In this study Chernoff Faces are utilized to visualize the data. Each Social Mention metric represents different facial features, e.g. Positive Mention = eye size, Negative Mention = pupil size, Strength = facial line, Passion = mouth curve, Reach = eyebrow density, Relative Frequency = hair density, Unique Authors = nose size. Table 3.4 is a summary of how the Social Mention variables were allocated to the various facial features and what Stata syntax were used to generate the images.

Table 3.4 Social Mention list of variables, definitions, Stata syntax and facial feature allocation

No.	Social Mention - Variables	Social Mention – Variable Definition	Facial Feature Allotment	Stata Syntax	Allocation Description
1.	Positive Sentiment	positive mentions	eye size	isize(exp)	Larger eye size = higher positive sentiment
2.	Negative Sentiment	negative mentions	pupil size	psize(exp)	Larger pupil size = higher negative sentiment
3.	Strength	likelihood brand is discussed on social media	face line	fline(exp)	Broader face = higher brand strength
4.	Passion	individuals talking about the brand repeatedly	mouth curve	mcurv(exp)	Higher passion = wider the curve of smile
5.	Reach	range of influence	brow density	bdens(exp)	Wider reach = denser brow
6.	Relative Frequency	frequency of brand mentions	hair density	hdark(exp)	Greater relative frequency = darker hair
7.	Unique Authors	number of authors messaging about the brand	nose line	nose(exp)	Higher unique authors = bigger nose

Chernoff Faces enhances the observer's ability to perceive and grasp significant trends and to remember imperative inferences. It is easy to scrutinize the sensitivity of variables, allowing observers to identify key distinguishing dimensions and recognize longitudinal trends, quickly (Raciborski, 2009; Farshid, Chan and Nel, 2012).

Once the collection and collation of the data and population of the tables as shown in table 3.2, for the specified period is completed, the tables were utilized as input data. The data was imported from Excel into Stata 15. The Stata syntax, as shown in table 3.4, was used to generate the Chernoff Faces. The faces were then scrutinized to detect patterns, clusters, outliers and longitudinal trends (Raciborski, 2009; Longest, 2015).

3.6.2 Sources (Social Mention – Social Media Platforms)

Table 3.3 dataset was utilized to generate the descriptive statistics, using Microsoft Excel. Bar and pie charts were created to interpret the data.

3.7 Ethical Considerations

The researcher used secondary data, which is free and publicly available on the Internet. The researcher will design a naming protocol for the 10 firms. These names will be used consistently for the duration of the data collection period, to ensure the validity of the data. Due to the dynamic nature of social media, the researcher will conduct a longitudinal study to gauge user sentiment. No social media posts from individual respondents would be used. All responses have been anonymized by Social Mention and collated in a score that cannot be traced back to individual social media posts.

3.8 Conclusion

The methodology of the research followed a mixed method approach. The chapter highlights how the data was collected, collated and populated. It explained how the data was analyzed and what computer programmes are utilized. Due to it being a mixed method approach, it also eludes to how the qualitative information was examined.

4 Chapter 4: Results and Findings

4.1 Introduction

This chapter presents the results of the Chernoff Faces from the Social Mention variables, depicted by the various facial features, as shown in Table 4.1, as well as the various social media platforms, as revealed by the data, for the specific companies, over the thirty-day period, presented in Table 4.2. The researcher has applied the actual maximum and minimum values from the datasets in Stata 15 (statistical software programme), to create two extreme images, to use as a reference point for the individual companies and for the companies' combined.

4.2 Chernoff Faces: Maximum and Minimum Values Per Firm

On completion of the specific firm's data collection, collation and population of tables, the researcher analysed the minimum and maximum values of the various companies' datasets, to create what the ideal and least ideal Chernoff Faces would look like. The ideal image would have dark hair, thick eyebrows, big round eyes, round small pupils, broad face, big nose and a wide smiling mouth. The least ideal face would reflect sparse hair, thin eyebrows, narrow eyes with dilated pupils, long thin face, sharp thin nose and a sad mouth curve, as depicted by figure 4.1. Table 4.1 elaborates on seven of the Social Mention variables and how it has been allotted to the different facial features. It also defines the Stata syntax for the facial features and Social Mention variables.

Table 4.1 Social Mention variables definition and facial feature allocation

No.	Social Mention - Variables	Social Mention – Variable Definition	Facial Feature Allotment	Stata Syntax	Allocation Description
1.	Positive Sentiment	positive mentions	eye size	isize(exp)	Larger eye size = higher positive sentiment
2.	Negative Sentiment	negative mentions	pupil size	psize(exp)	Larger pupil size = higher negative sentiment
3.	Strength	likelihood brand is discussed on social media	face line	fline(exp)	Broader face = higher brand strength
4.	Passion	individuals talking about the brand repeatedly	mouth curve	mcurv(exp)	Higher passion = wider the curve of smile
5.	Reach	range of influence	brow density	bdens(exp)	Wider reach = denser brow
6.	Relative Frequency	frequency of brand mentions	hair density	hdark(exp)	Greater relative frequency = darker hair
7.	Unique Authors	number of authors messaging about the brand	nose line	nose(exp)	Higher unique authors = bigger nose

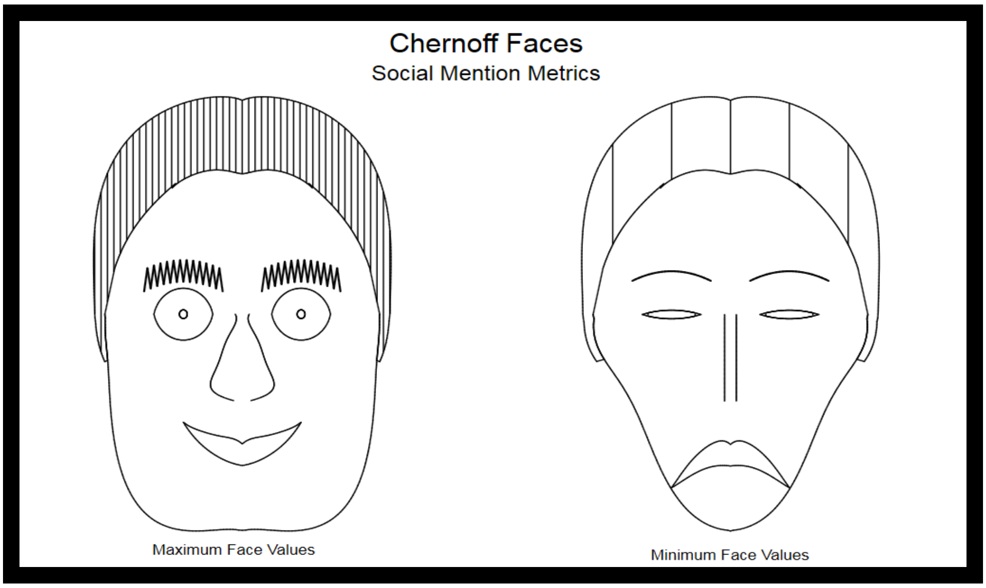

Figure 4.1 Chernoff Faces Maximum and Minimum Value Per Firm

4.3 Sources (Social Mention – Social Media Platforms)

The data collected, in table 3.3, has been manipulated in Microsoft Excel, to generate the descriptive statistics. Bar graphs and pie charts were created to interpret the data in these datasets.

4.4 Data Visualisation

The minimum and maximum values of the specific firms' datasets (table 3.2) were applied in Stata 15, to generate the specific firm's daily faces, as illustrated in figure 4.2. The daily faces were labelled “Day 1” through to “Day 30”. In the next sections, the results of the individual companies, in respect of the faces and social media platforms, are revealed.

4.4.1 AstraZeneca Pharmaceuticals

Figure 4.2 is a graphical representation of the Chernoff Faces, the results of the diary study conducted for 30 days from 5 October 2018 to 4 November 2018. The datasets for the various variables, as explained in table 4.1, were uploaded in Stata 15, to generate the 30 individual faces.

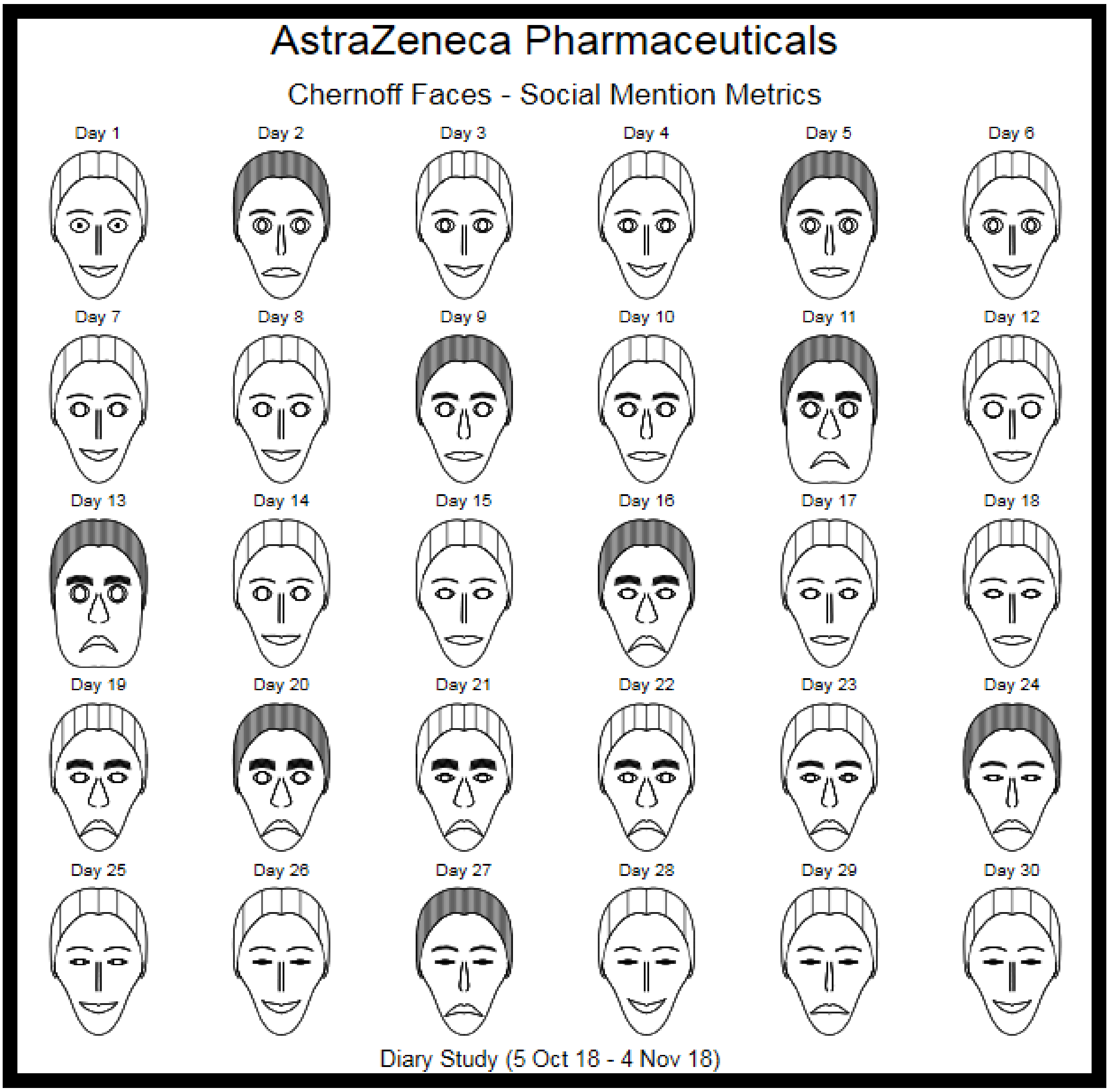

Figure 4.2 AstraZeneca Pharmaceuticals, Chernoff Faces Graph (Diary Study 5 Oct 18 – 4 Nov 18)

4.4.1.1 Chernoff Faces

The faces for AstraZeneca over the 30 days are all noticeably heterogenous, even though the face lines are generally similar. Day 11 and day 13 seem to be the only outliers. On these days the Social Mention attributes all appear to be at maximum values, except for passion (likelihood that individuals talking about the brand will do so repeatedly). Passion is the only attribute at minimum value, hence showing the sad mouth curve. While negative sentiment, depicted by the pupil size, is at maximum value, the implication is negative, hence the pupils are dilated.

The individual faces are analysed as per the following 7 variables:

1) Positive Sentiment (positive mentions) – Eye Size: larger eyes = higher positive sentiment

Days 26,27,28,29 and 30 are the only days the firm's positive mentions are at minimum value, therefore the eyes are narrow. Hence, an indication that for most of the days the mentions are more positive than negative.

2) Negative Sentiment (negative mentions) – Pupil Size: larger pupil size = higher negative sentiment

Day 1 is the only day that there was no negative sentiment hence the pupil is constricted.

3) Strength (Likelihood brand is discussed on social media) - Face line: broader face = higher brand strength

What is visibly apparent is that the variable strength at minimum value is the most prominent. Over the 30 days, for 27 of the days (. i.e. 90% of the period) the brand strength has performed relatively poorly, therefore showing the trend of the long, thin faces. This means that it is very unlikely that the brand is being discussed on social media.

4) Passion (Individuals talking about the brand repeatedly) – Mouth Curve; higher passion = wider smile

For most of the days the passion was high, indicating that the same individuals were generally talking about the brand. 10 of the days reveal that the passion was low, which means that only 30% of the individuals were not the same people.

5) Reach (range of influence) – Brow Density: wider reach = denser brow

The reach for the was only 23% during the 30 days.

6) Relative Frequency (frequency of brand mentions) – Hair Density: greater relative frequency = darker hair

Days 2,5,9,11,13,16,20,24 and 27 are the only days when the relative frequency of the brand mention, are at maximum value. Relative frequency was measured based on the frequency with which mentions of the brand would appear. This was measured in hours. Mentions recorded within 1 to 6 hours would gain maximum value. Mentions within 7 – 12 hours would get slightly less value and mentions 13 – 24 hours would receive even lesser value. Mentions between 25 hours and more received the least value. This is represented by the hair density. Implying that for less than 30% of the period, the mentions were within 1 to 6 hours, meaning mentions of the brand were not very frequent. Hence some many faces with sparse hair.

7) Unique Authors (number of authors messaging about the brand) – Nose Size: higher unique authors = bigger nose

The unique authors messaging about the brand was only 23% of the period.

4.4.1.2 Sources (Social Mention - Social Media Platform)

The results, as represented in figure 4.3, revealed the social media platforms that users were most active on.

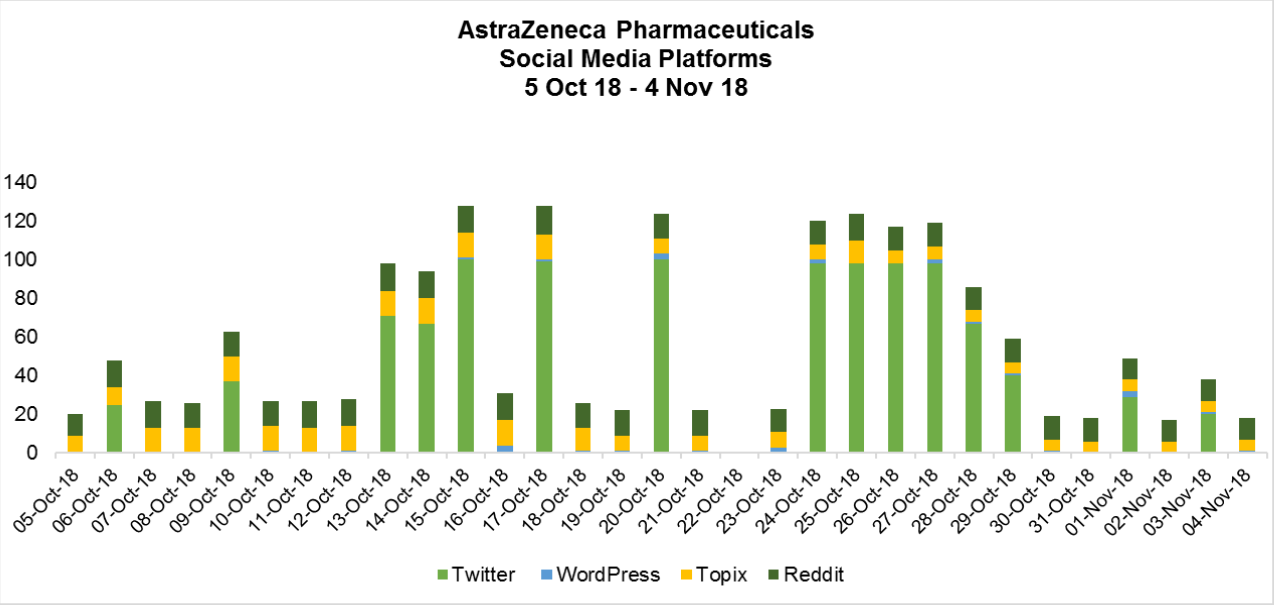

Figure 4.3 Social Mention Sources - Social Media Platforms for 5 Oct 18 – 4 Nov 18

What is apparent is that, although Twitter generates 60% (see figure 4.4) of the mentions, users are not posting on the social media platform daily. For 15 of the 30 days, there is no activity on Twitter. While Topix and Reddit generates only 16% and 22% (see figure 4.4) of the mentions respectively, users are active on these platforms daily.

During this period, as demonstrated in figure 4.4, the results show that users were only active on four social media platforms, namely: Twitter, WordPress, Topix and Reddit. The most mentions on social media platform is Twitter. However, the most visited social media platforms Reddit and Topix, as shown in figure 4.3, did not generate the most engagements.

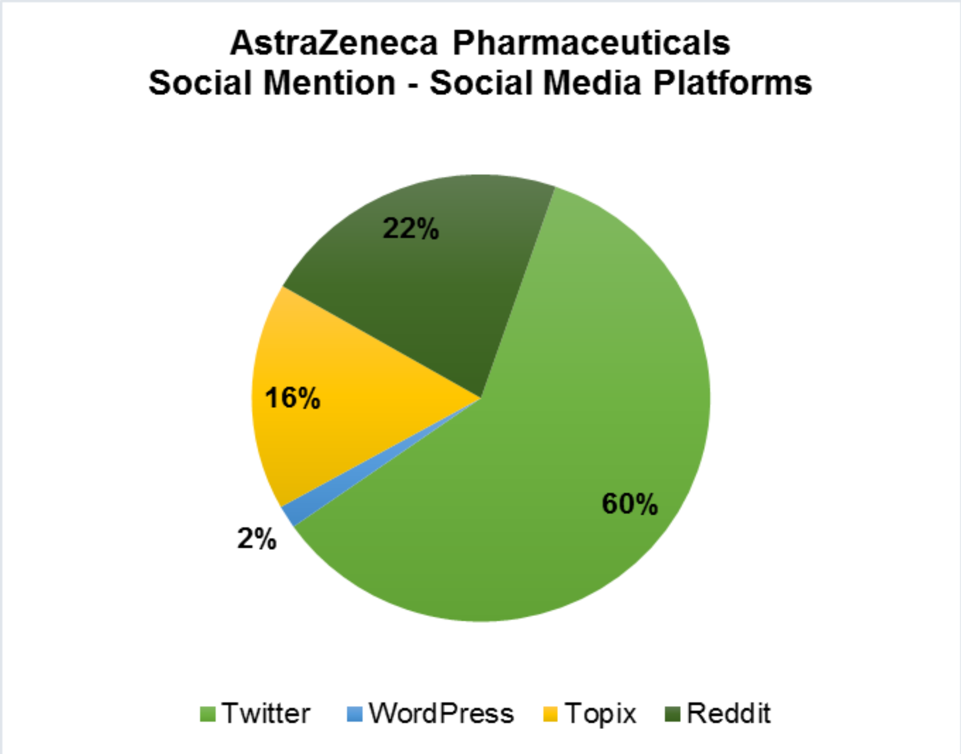

Figure 4.4 Social Mention – Social Media Platforms

4.4.2 Bayer Pharmaceuticals

Figure 4.5 is a graphical representation of the Chernoff Faces, the results of the diary study conducted for 30 days from 5 October 2018 to 4 November 2018. The datasets for the various variables as explained in table 4.1 were uploaded in Stata 15, to generate the 30 individual faces.

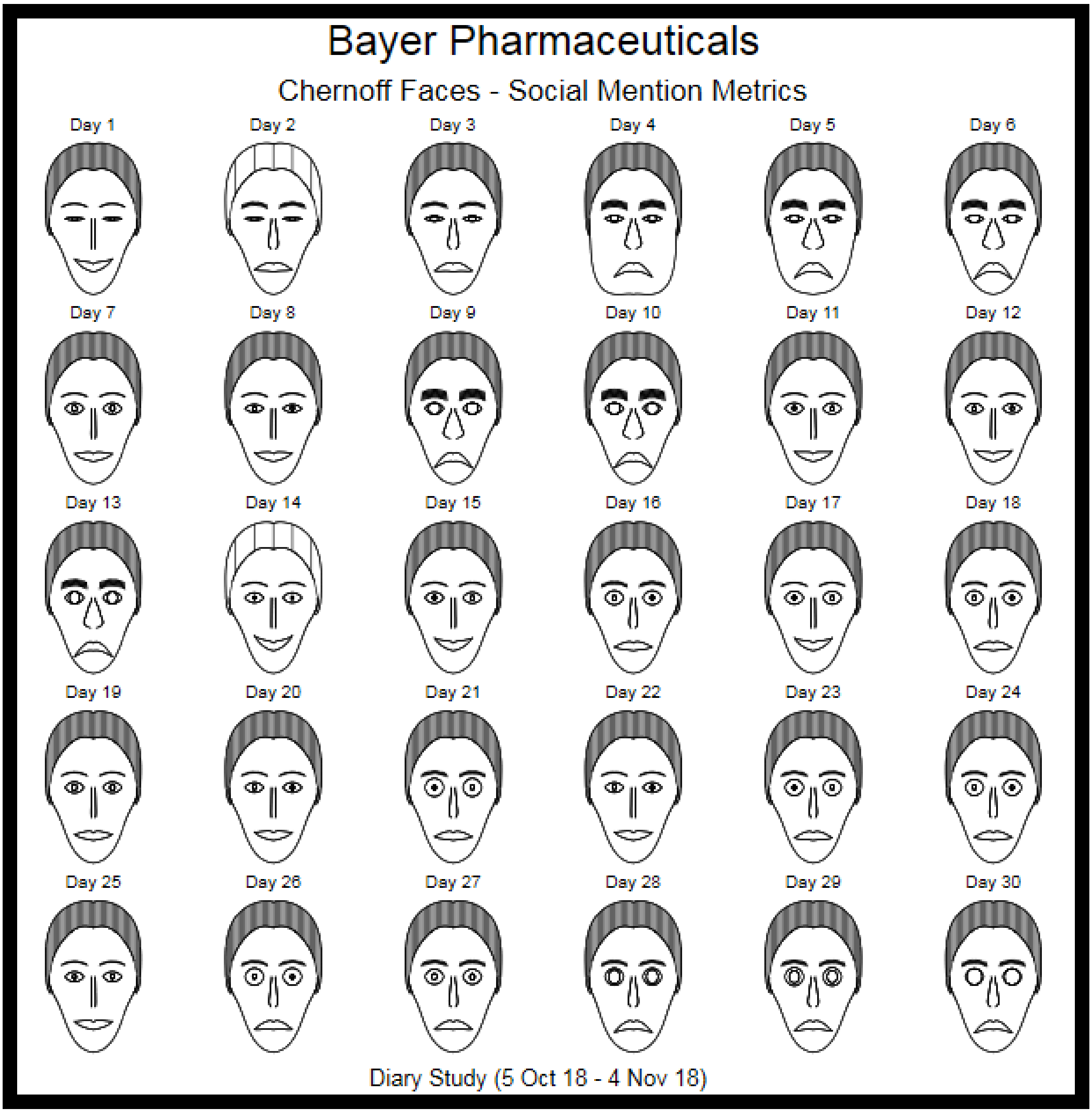

Figure 4.5 Bayer Pharmaceuticals, Chernoff Faces Graph (Diary Study 5 Oct 18 – 4 Nov 18)

4.4.2.1 Chernoff Faces

The faces for Bayer over the 30 days are all strikingly similar. Most of the images sport dark hair and long thin faces. Implying that the brand's strength is at a low, however the dense brows indicates that the brand's reach is wide. Day 4 and 5 and day 2 and 14 look as if it's the only outliers. On days 4 and 5 the Social Mention attributes all give the impression that all are at maximum values, apart from for passion and positive and negative sentiments. Days 2 and 14 are the only days brand reach is at a minimum.

The individual faces are analysed as per the following 7 variables:

1) Positive Sentiment (positive mentions) – Eye Size: larger eyes = higher positive sentiment

The firm's positive mentions are at maximum value for 11 of the days, namely: 9,10,13, 21, 23, 24, 26, 27, 28, 29 and 30, therefore the eyes are big and round. Which is an indication that for these days, the mentions are mostly positive.

2) Negative Sentiment (negative mentions) – Pupil Size: larger pupil size = higher negative sentiment

In 18 of the faces, that is 60% of the pupils are obviously constricted, indicating that the values of the negative sentiment are at its minimum.

3) Strength (Likelihood brand is discussed on social media) - Face line: broader face = higher brand strength

What is very conspicuous, is that the variable strength is the most blatant. Over the 30 days, for 28 of the days the brand strength has performed relatively poorly, therefore showing the trend of the long, thin faces. This suggests that it is highly unlikely that the brand is being talked about on social media.

4) Passion (Individuals talking about the brand repeatedly) – Mouth Curve; higher passion = wider smile

For most of the days the passion was high, indicating that the same individuals were generally talking about the brand. 5 of the days reveal that the passion was low, which means that about 17% of the individuals were not the same people.

5) Reach (range of influence) – Brow Density: wider reach = denser brow

The reach for the brand was 20% during the 30 days.

6) Relative Frequency (frequency of brand mentions) – Hair Density: greater relative frequency = darker hair

93% of the mentions were at maximum value. Relative frequency was measured based on the frequency with which mentions of the brand would appear. This was measured in hours and mentions within 1 to 6 hours would gain maximum value. This is represented by the hair density. Implying that the mentions of the brand were frequent.

7) Unique Authors (number of authors messaging about the brand) – Nose Size: higher unique authors = bigger nose

The unique authors messaging about the brand was about 17% for the period.

4.4.2.2 Sources (Social Mention - Social Media Platform)

The results revealed, as per the datasets in chapter 3, table 3.3, shown here in figure 4.6, that the following social media platforms that users were most active on, on the specific days. Topix and Reddit are visited daily, while Twitter is the most active, with more users visiting on certain days.

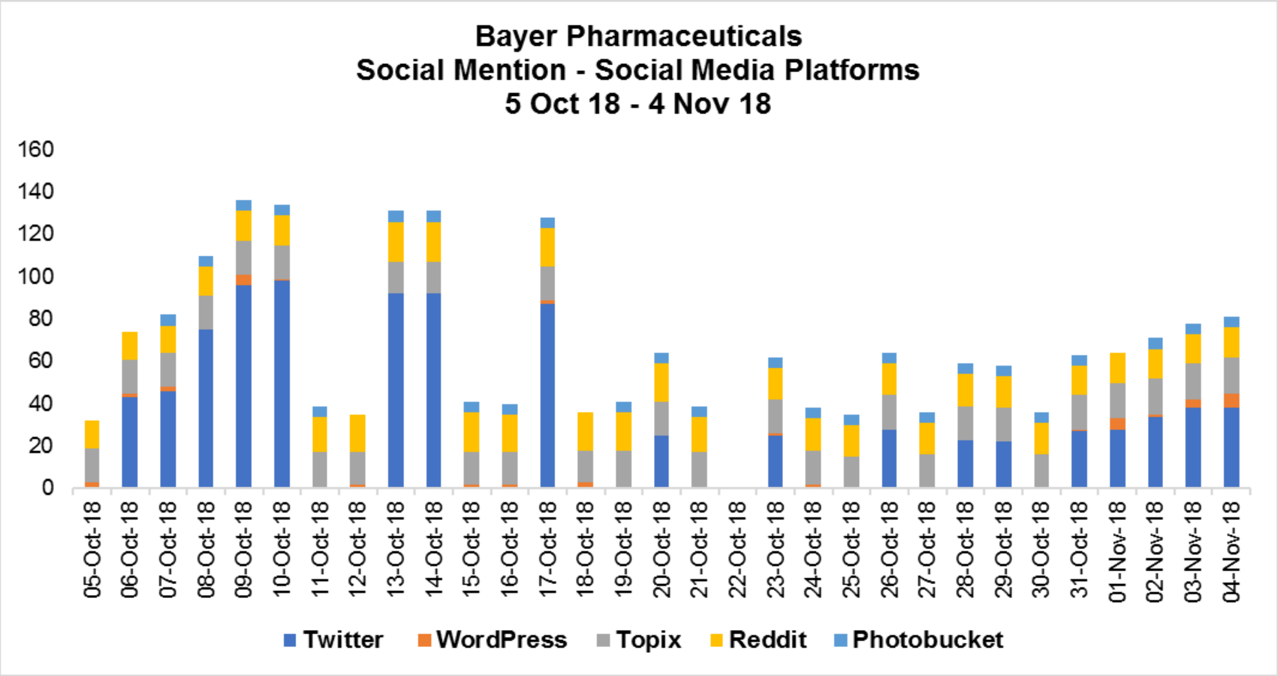

Figure 4.6 Social Mention Sources - Social Media Platforms for 5 Oct 18 – 4 Nov 18

During this period, as demonstrated in figure 4.7, the results show that users were only active on five social media platforms, namely: Twitter, WordPress, Topix, Reddit and Photobucket. The most active social media platform being Twitter. The most frequented platforms are Topix and Reddit as seen in figure 4.6.

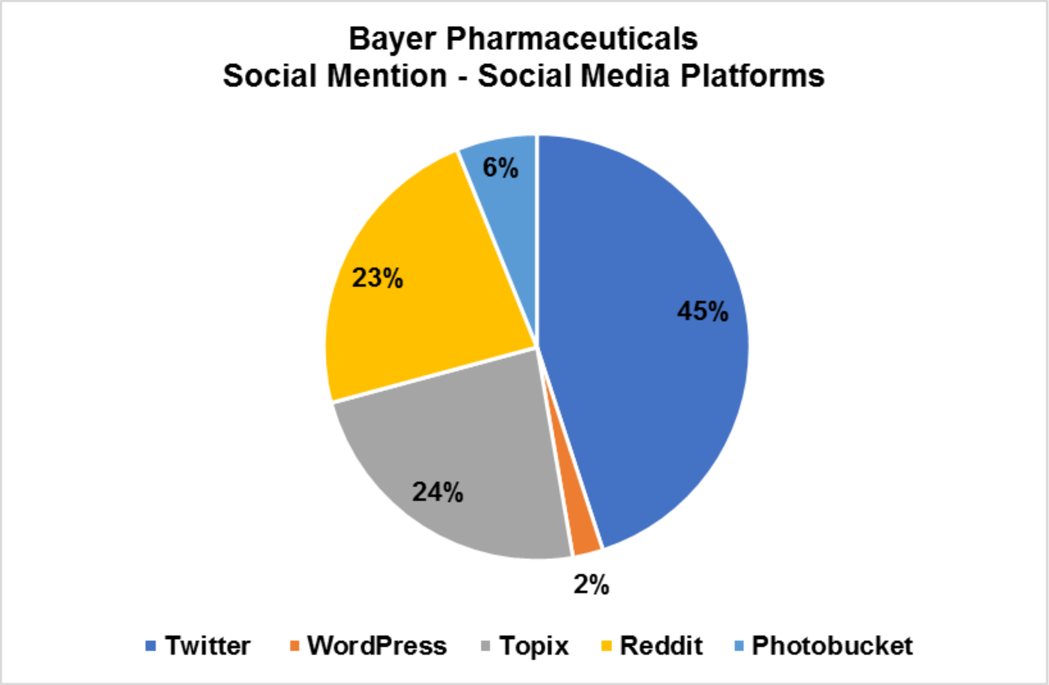

Figure 4.7 Social Mention – Social Media Platforms

4.4.3 Boehringer Ingelheim Pharmaceuticals

Figure 4.8 is a graphical representation of the Chernoff Faces, the results of the diary study conducted for 30 days from 5 October 2018 to 4 November 2018. The datasets for the various variables as explained in table 4.1 were uploaded in Stata 15, to generate the 30 individual faces.

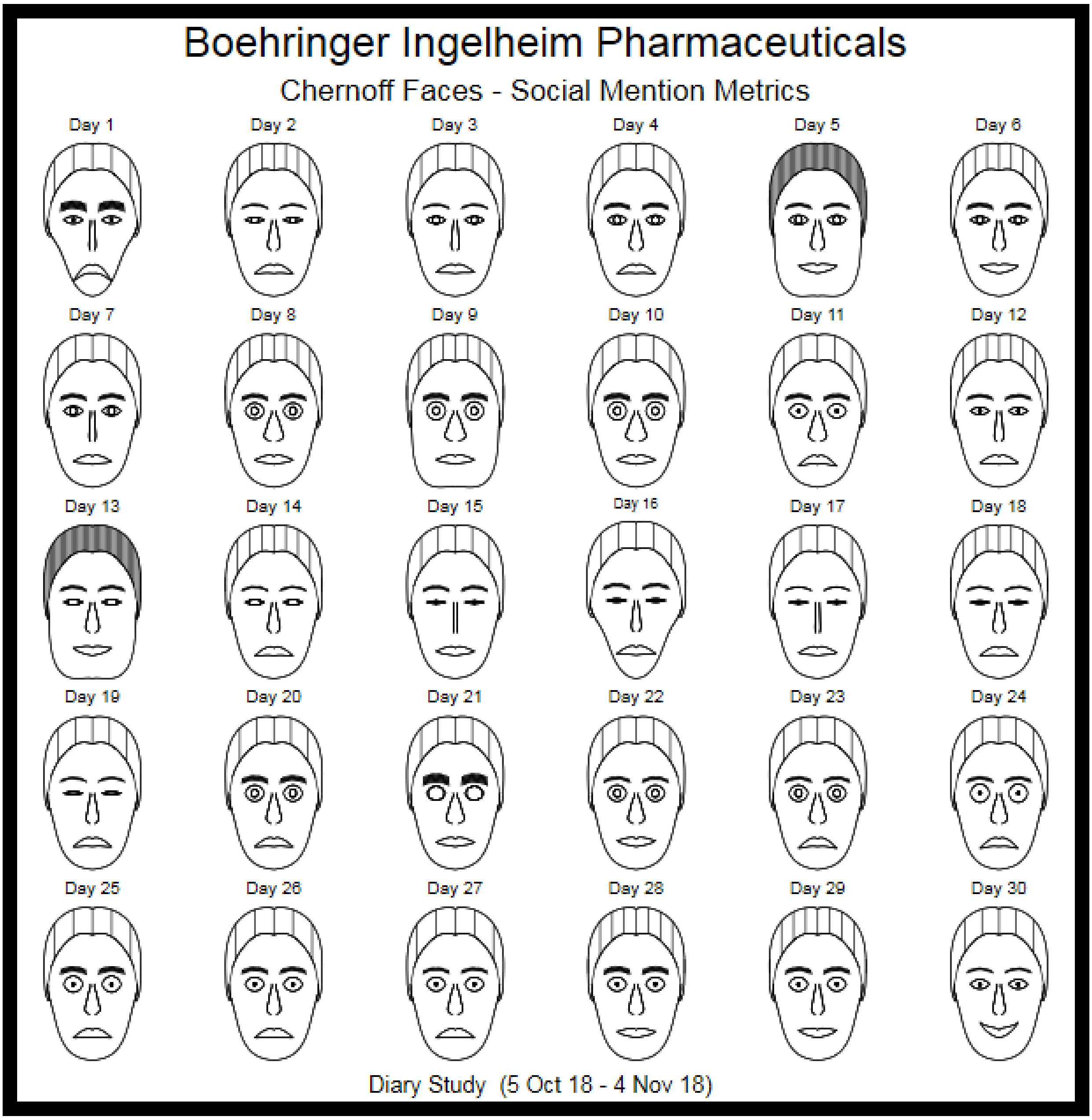

Figure 4.8 Boehringer Ingelheim Pharmaceuticals, Chernoff Faces Graph (Diary Study 5 Oct 18 – 4 Nov 18)

4.4.3.1 Chernoff Faces

The faces for Boehringer Ingelheim over the 30 days are all vividly analogous. Days 1 and 16 and days 5 and 13 appear to be the only outliers. While days 1 and 16 are at minimum values, the narrow long thin faces display that, on these two days, the brand strength was at its weakest. On days 5 and 13, the Social Mention attributes all seem to be at maximum values, with relative frequency being very pronounced. However, the overall general brand strength for the period, reflects that the values were at mid-point, as expressed by more than 90% of the faces. This means that the brand strength was average over the period.

The individual faces are analysed as per the following 7 variables:

1) Positive Sentiment (positive mentions) – Eye Size: larger eyes = higher positive sentiment

At a quick glance about half of the image's eyes are round and big.

2) Negative Sentiment (negative mentions) – Pupil Size: larger pupil size = higher negative sentiment

There are 8 days where the pupils are constricted conveying that there were no negative sentiments on these days.

3) Strength (Likelihood brand is discussed on social media) - Face line: broader face = higher brand strength

The multitude of normal looking facing expresses that the strength variable values were all average. Over the 30 days, this is so for 26 of the days implying that the brand strength for this period was mediocre.

4) Passion (Individuals talking about the brand repeatedly) – Mouth Curve; higher passion = wider smile

On day 1 and day 30 the mouth curve is at its two extremes. The rest of the time its neutral implying that the values for passion were run-of-the-mill.

5) Reach (range of influence) – Brow Density: wider reach = denser brow

The reach for the was only at maximum for 7% of the time i.e. day 1 and 21. Numerous of the days the social media influence was sparse. Meaning that the unique authors referencing the brand were at minimum value for 93% of the time.

6) Relative Frequency (frequency of brand mentions) – Hair Density: greater relative frequency = darker hair

Days 5 and 13 are the only days when the relative frequency of the brand mention, are at maximum value.

7) Unique Authors (number of authors messaging about the brand) – Nose Size: higher unique authors = bigger nose

The unique authors messaging about the brand was only at about 13% for the period.

4.4.3.2 Sources (Social Mention - Social Media Platform)

The results revealed, as per the datasets in chapter 3, table 3.3, shown here in figure 4.9, that the following social media platforms that users were most active on, on specific days. It is interesting to note that, on the last four days, there was no activity on any of the social media platforms.

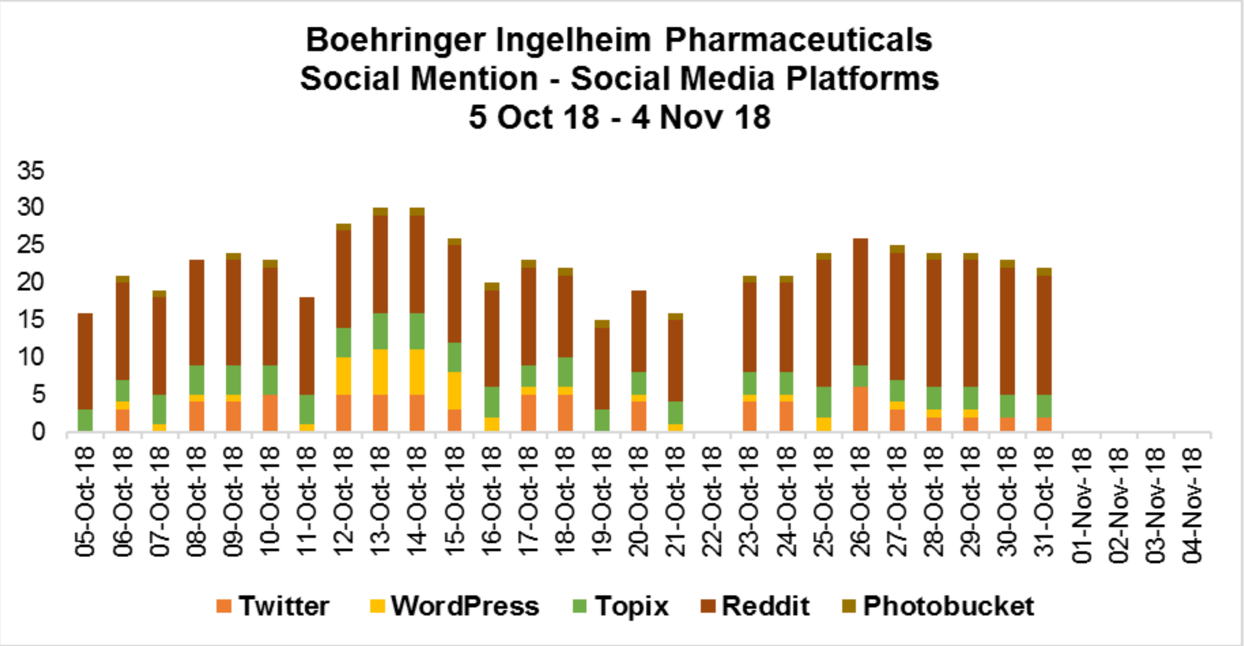

Figure 4.9 Social Mention Sources - Social Media Platforms for 5 Oct 18 – 4 Nov 18

During this period, as demonstrated in figure 4.10, the results show that users were only active on five social media platforms, namely: Twitter, WordPress, Topix, Reddit and Photobucket. The most frequented social media platform being Twitter.

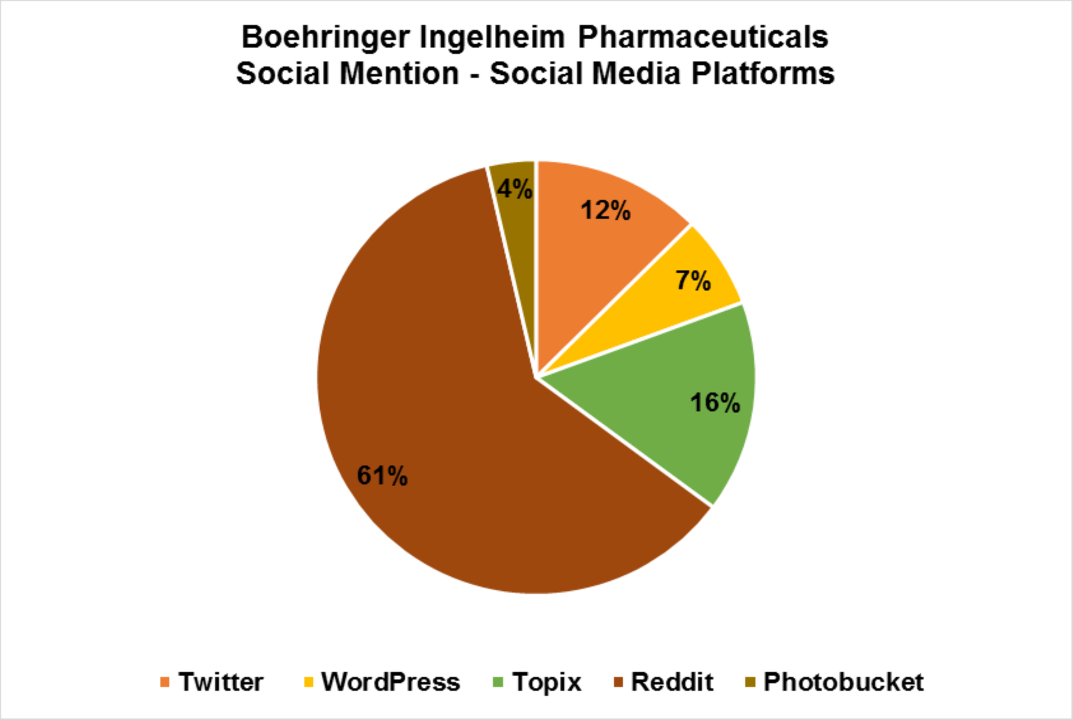

Figure 4.10 Social Mention – Social Media Platforms

4.4.4 Eli Lilly Pharmaceuticals

Figure 4.11 is a graphical representation of the Chernoff Faces, the results of the diary study conducted for 30 days from 5 October 2018 to 4 November 2018. The datasets for the various variables, as shown in table 4.1, were uploaded in Stata 15, to generate the 30 individual faces.

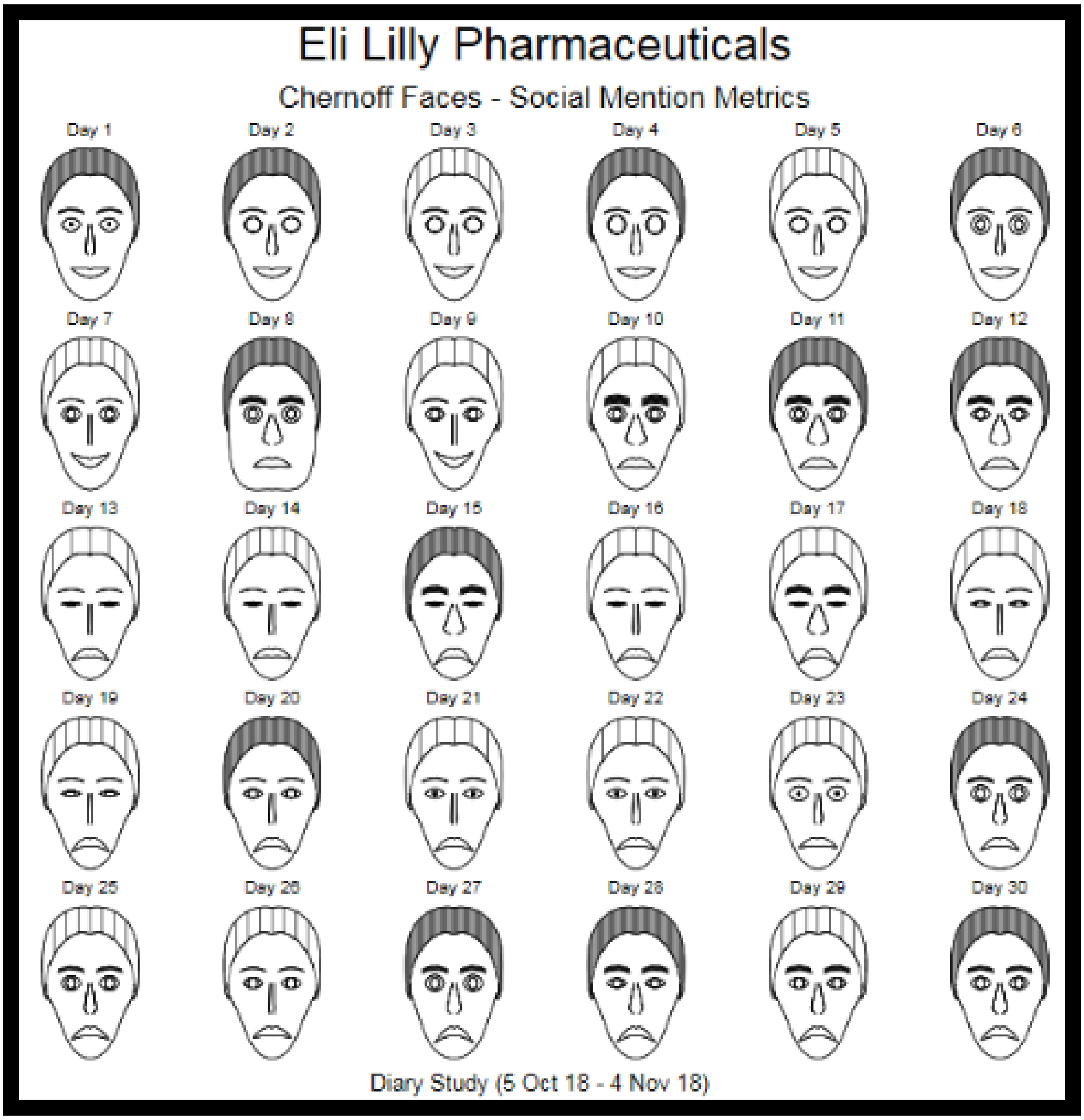

Figure 4.11 Eli Lilly Pharmaceuticals, Chernoff Faces Graph (Diary Study 5 Oct 18 – 4 Nov 18)

4.4.4.1 Chernoff Faces

The faces for Eli Lilly over the 30 days are all slightly varied. The general trend of the long, thin, sad-looking faces is obvious. During the diary study, only day 8 reflects an image at almost maximum values. The overall trend points out that the firms brand strength during this time is relatively weak.

The individual faces are analysed as per the following 7 variables:

1) Positive Sentiment (positive mentions) – Eye Size: larger eyes = higher positive sentiment

At a glance the eyes seem evenly round and big, compared to the narrow and long ones. Thus, indicating the positive versus negative mentions are almost the same.

2) Negative Sentiment (negative mentions) – Pupil Size: larger pupil size = higher negative sentiment

At a glance, the eyes seem evenly round and big, compared to the narrow and long ones. Thus, indicating the positive versus negative mentions are almost the same.

3) Strength (Likelihood brand is discussed on social media) - Face line: broader face = higher brand strength

The general trend of long thin sad looking faces is obvious. During the diary study, only day 8 reflects an image at almost maximum values. The overall trend points out that the firm’s brand strength, during this time, is relatively weak.

4) Passion (Individuals talking about the brand repeatedly) – Mouth Curve; higher passion = wider smile

From day 11 to day 30, the down curve of the mouth is evident. On these days the passion is at minimum values, indicating that the same individuals are not talking about the brand.

5) Reach (range of influence) – Brow Density: wider reach = denser brow

The overall reach is not wide as for only 3 of the days it was at maximum values.

6) Relative Frequency (frequency of brand mentions) – Hair Density: greater relative frequency = darker hair

The relative frequency of brand mentions is average, as indicated by the faces, with about 50% having dense hair.

7) Unique Authors (number of authors messaging about the brand) – Nose Size: higher unique authors = bigger nose

The unique authors messaging about the brand was around 13% of the period.

4.4.4.2 Sources (Social Mention - Social Media Platform)

The results revealed, as per the datasets in chapter 3, table 3.3, shown here in figure 4.12 that the following social media platforms that users were most active on, on specific days.

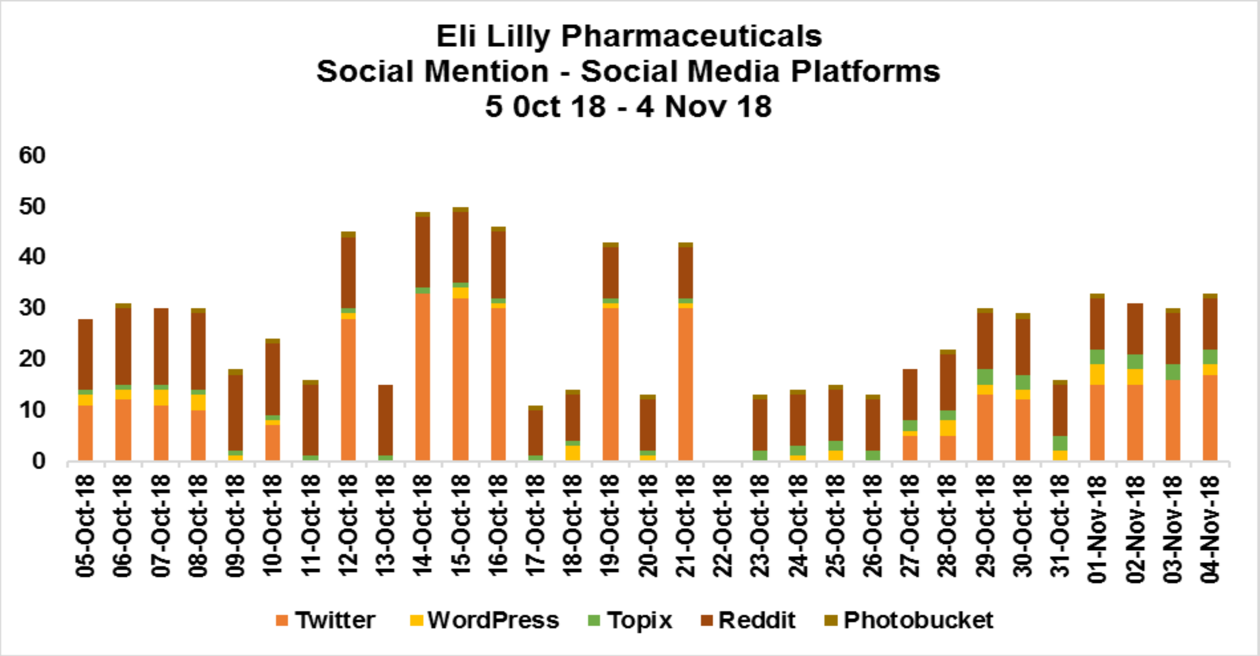

Figure 4.12 Social Mention Sources - Social Media Platforms for 5 Oct 18 – 4 Nov 18

During this period, as demonstrated in figure 4.13, the results show that users were only active on five social media platforms, namely: Twitter, WordPress, Topix and Reddit. The most active and visited social media platform being Reddit.

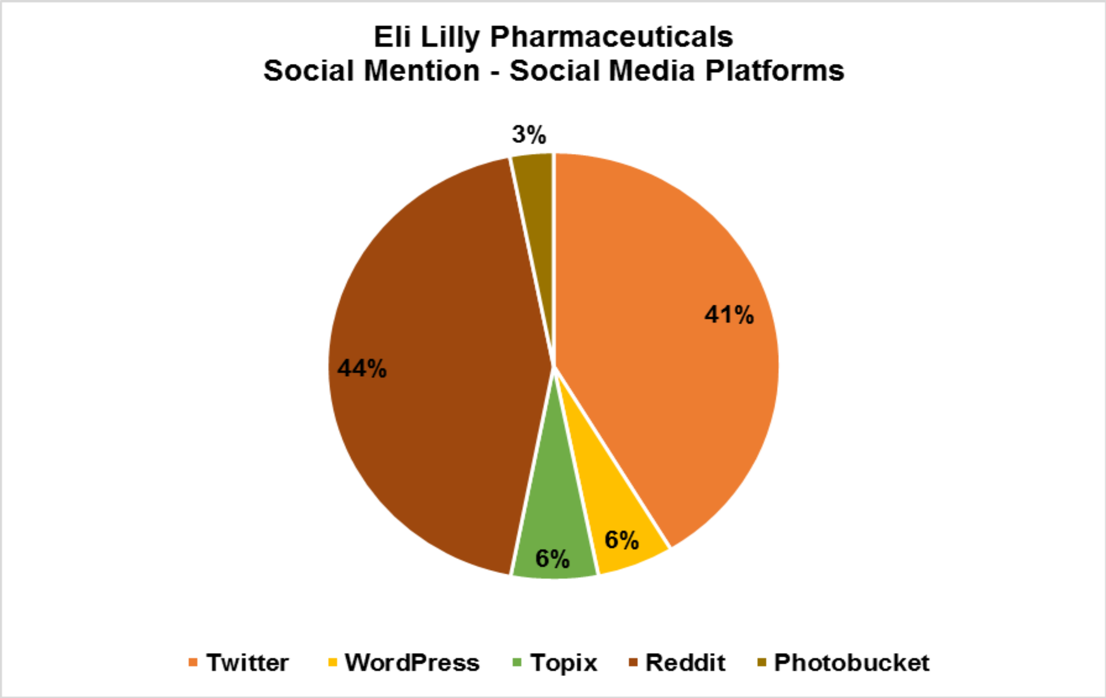

Figure 4.13 Social Mention – Social Media Platforms

4.4.5 Johnson and Johnson Pharmaceuticals

Figure 4.14 is a graphical representation of the Chernoff Faces, the results of the diary study conducted for 30 days from 5 October 2018 to 4 November 2018. The datasets for the various variables, as explained in table 4.1, were uploaded in Stata 15, to generate the 30 individual faces.

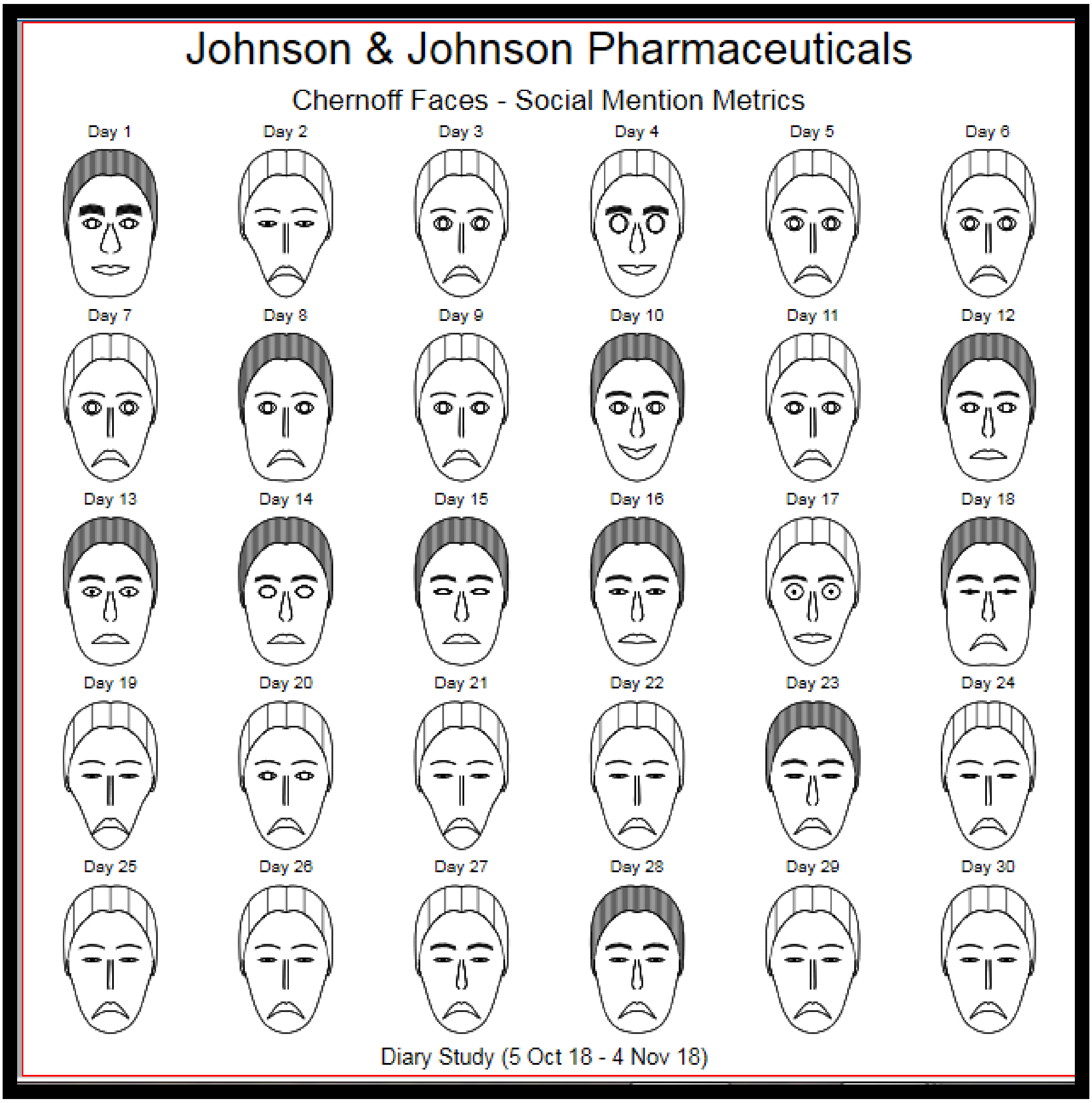

Figure 4.14 Johnson & Johnson Pharmaceuticals, Chernoff Faces Graph (Diary Study 5 Oct 18 – 4 Nov 18)

4.4.5.1 Chernoff Faces

The faces for Johnson and Johnson, over the 30 days, are all markedly similar. Day 1, 8 and 18 seem to be the only outliers. On these days, the Social Mention attribute strength, appears to be at maximum values, compared to the rest of the faces.

The individual faces are analysed as per the following 7 variables:

1) Positive Sentiment (positive mentions) – Eye Size: larger eyes = higher positive sentiment

The round big eyes are far and few between thus an indication that for most of the days the mentions are more negative than positive.

2) Negative Sentiment (negative mentions) – Pupil Size: larger pupil size = higher negative sentiment

Day 2, 13, 16 and 17 are the only days that there was no negative sentiment hence the pupils of these specific eyes are constricted.

3) Strength (Likelihood brand is discussed on social media) - Face line: broader face = higher brand strength

What is visibly apparent is that the variable strength is the most prominent. With only three faces being at maximum value. This means that brand strength overall all is mediocre for this period.

4) Passion (Individuals talking about the brand repeatedly) – Mouth Curve; higher passion = wider smile

For most of the days (21 days) the passion was low, indicating that the same individuals were not talking about the brand. Hence the many sad looking faces.

5) Reach (range of influence) – Brow Density: wider reach = denser brow

Day 1 was the only day with reach at its maximum point. The overall reach was low indicating that very few unique authors are referencing the brand.

6) Relative Frequency (frequency of brand mentions) – Hair Density: greater relative frequency = darker hair

The frequency of brand mentions was spread evenly across the various intervals for the period.

7) Unique Authors (number of authors messaging about the brand) – Nose Size: higher unique authors = bigger nose

The unique authors messaging about the brand was at maximum value only on day 1. For the rest of the period the number of users messaging within a specified time period was very few.

4.4.5.2 Sources (Social Mention - Social Media Platform)

The results revealed, as per the datasets in chapter 3, table 3.3, shown here in figure 4.15 that the following social media platforms that users were most active on, on the specific days.

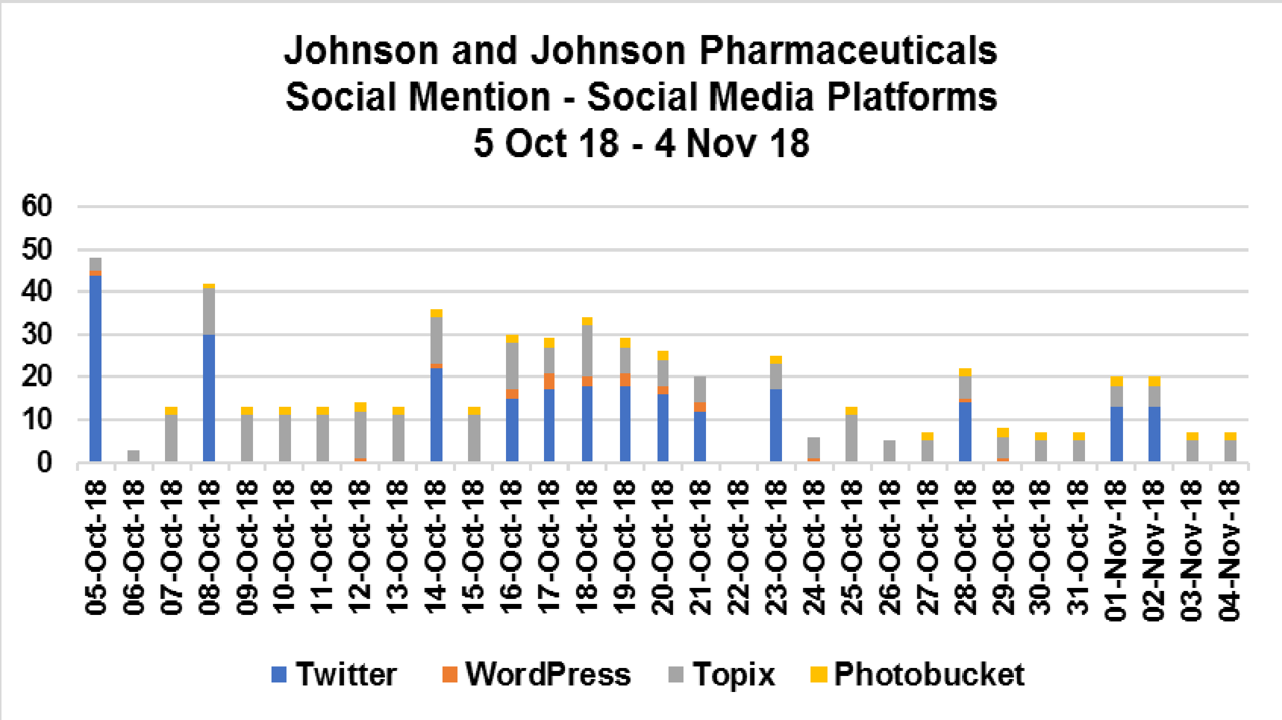

Figure 4.15 Social Mention Sources - Social Media Platforms for 5 Oct 18 – 4 Nov 18

During this period, as demonstrated in figure 4.16, the results show that users were only active on four social media platforms, namely: Twitter, WordPress, Topix and Reddit. The most active social media platform being Twitter see figure 4.16. The most visited platforms being Topix followed by Photobucket, as can be seen in figure 4.15.

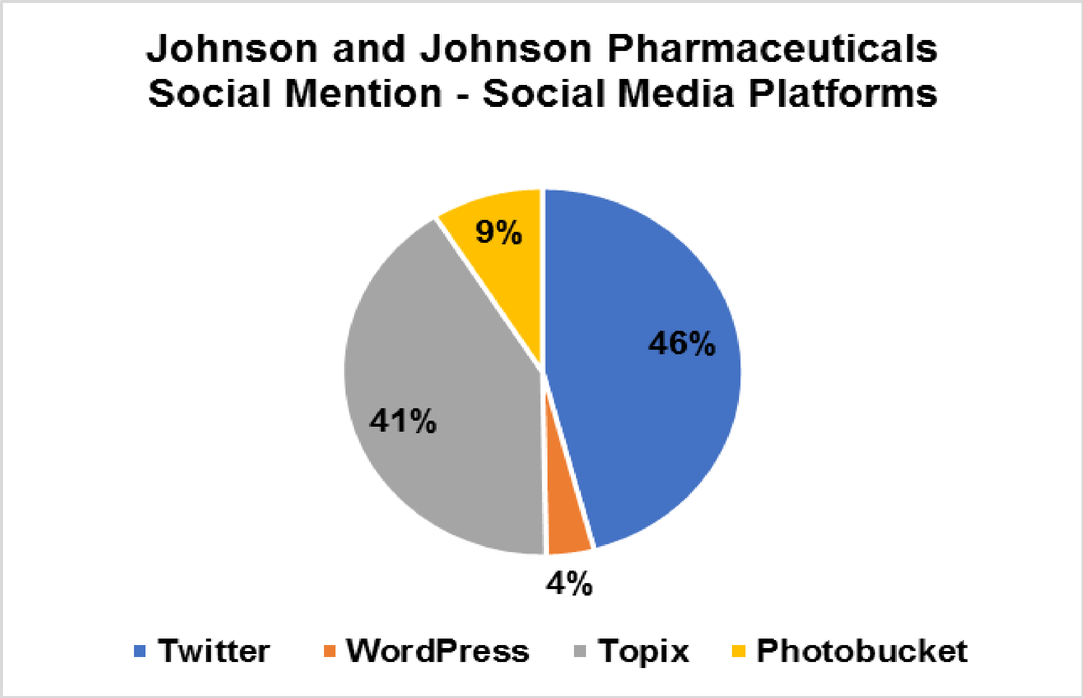

Figure 4.16 Social Mention – Social Media Platforms

4.4.6 Merck Pharmaceuticals

Figure 4.17 is a graphical representation of the Chernoff Faces, the results of the diary study conducted for 30 days from 5 October 2018 to 4 November 2018. The datasets for the various variables as explained in table 4.1 were uploaded in Stata 15, to generate the 30 individual faces.

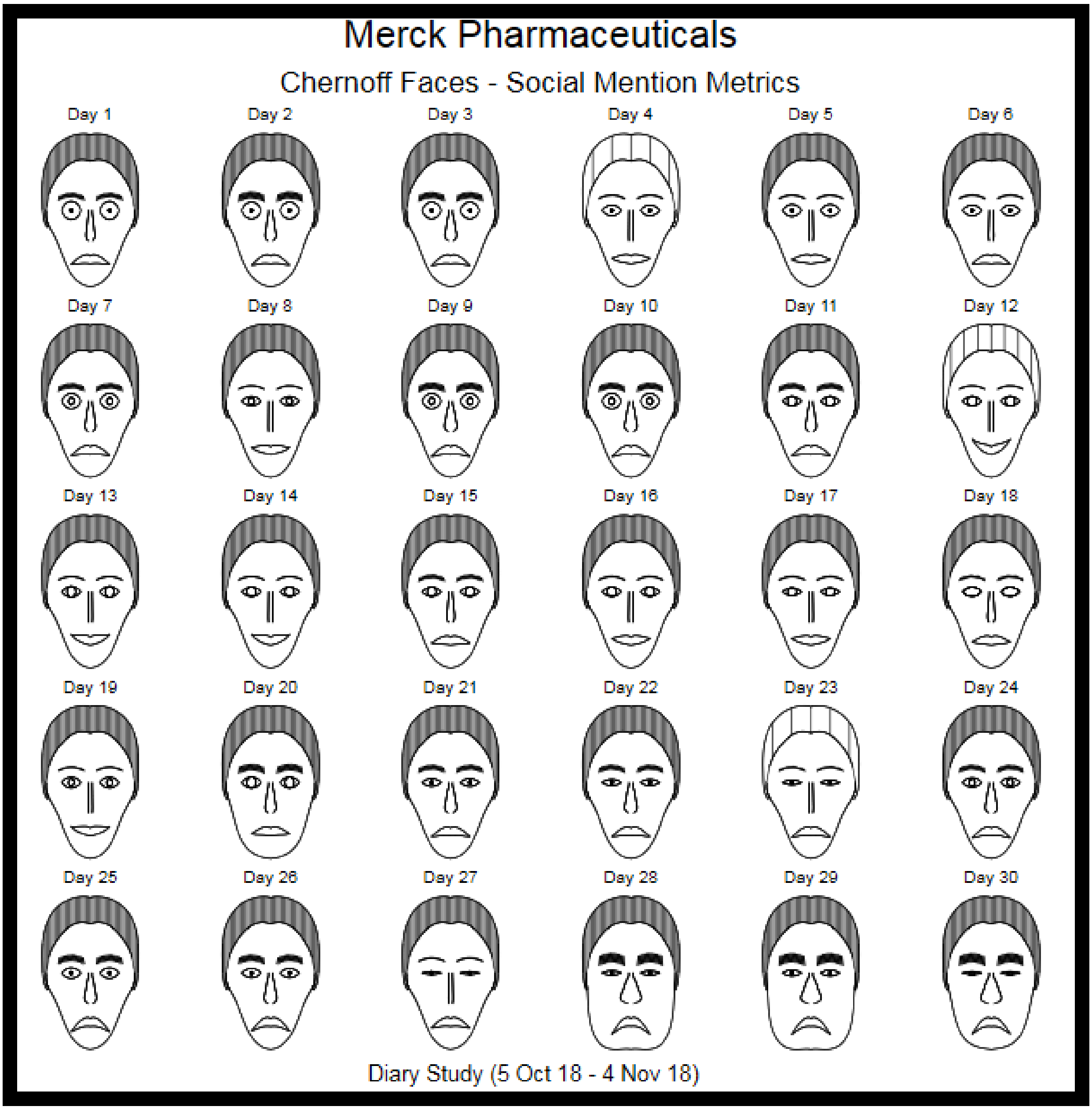

Figure 4.17 Merck Pharmaceuticals, Chernoff Faces Graph (Diary Study 5 Oct 18 – 4 Nov 18)

4.4.6.1 Chernoff Faces

The faces for Merck over the 30 days are all strikingly matching. Day 4,12 and 23 appear to be the only outliers, at a quick glance. On these days, the hair density is at minimum value, meaning that the brand experienced the frequent mention of its brand, for most of this period.

The individual faces are analysed as per the following 7 variables:

1) Positive Sentiment (positive mentions) – Eye Size: larger eyes = higher positive sentiment

For most of the days the mentions are evenly spread between maximum and minimum values. Hence a third of the faces sport big round eyes, a third normal looking eyes and a third narrow eyes.

2) Negative Sentiment (negative mentions) – Pupil Size: larger pupil size = higher negative sentiment

For thirteen of the days there are no negative sentiment hence these eyes pupils are visibly constricted.

3) Strength (Likelihood brand is discussed on social media) - Face line: broader face = higher brand strength

What is visibly apparent, is that the variable strength is the most noticeable. Over the 30 days, for 26 of the days the brand strength has performed relatively poorly, therefore, the trend of the long, thin faces stands out. This means that it is very unlikely that the brand is being discussed on social media.

4) Passion (Individuals talking about the brand repeatedly) – Mouth Curve; higher passion = wider smile

For most of the days the passion was low, indicating that the same individuals are not generally talking about the brand. Only on day 12 do we encounter a face with a smile.

5) Reach (range of influence) – Brow Density: wider reach = denser brow

The reach was even during this period, some days at maximum, some at minimum and some at midpoint.

6) Relative Frequency (frequency of brand mentions) – Hair Density: greater relative frequency = darker hair

90% of the mentions were at maximum value. Relative frequency was measured based on the frequency with which mentions of the brand would appear. This was measured in hours and mentions within 1 to 6 hours would gain maximum value. This is represented by the hair density. Implying that the mentions of the brand were frequent.

7) Unique Authors (number of authors messaging about the brand) – Nose Size: higher unique authors = bigger nose

The unique authors messaging about the brand was only 10% of the period.

4.4.6.2 Sources (Social Mention - Social Media Platform)

The results revealed, as per the datasets in chapter 3, table 3.3, shown here in figure 4.18 that the following social media platforms that users were most active on, on the specific days.

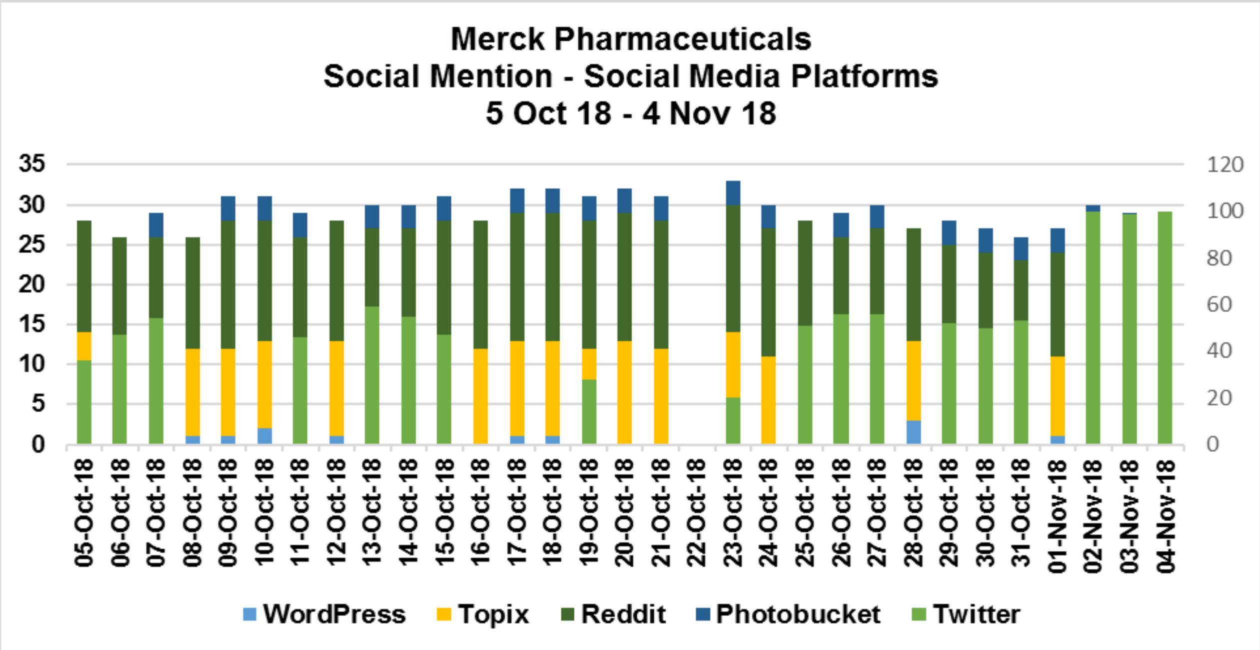

Figure 4.18 Social Mention Sources - Social Media Platforms for 5 Oct 18 – 4 Nov 18

During this period, as demonstrated in figure 4.19, the results show that users were only active on five social media platforms, namely: Twitter, WordPress, Topix, Reddit and Photobucket. The most active social media platform being Twitter as seen in figure 4.19. even though the was not activity on Twitter for twelve of the days during this period. Reddit was the most visited social media platform during this period for 29 out of the 30 days users frequented Reddit.

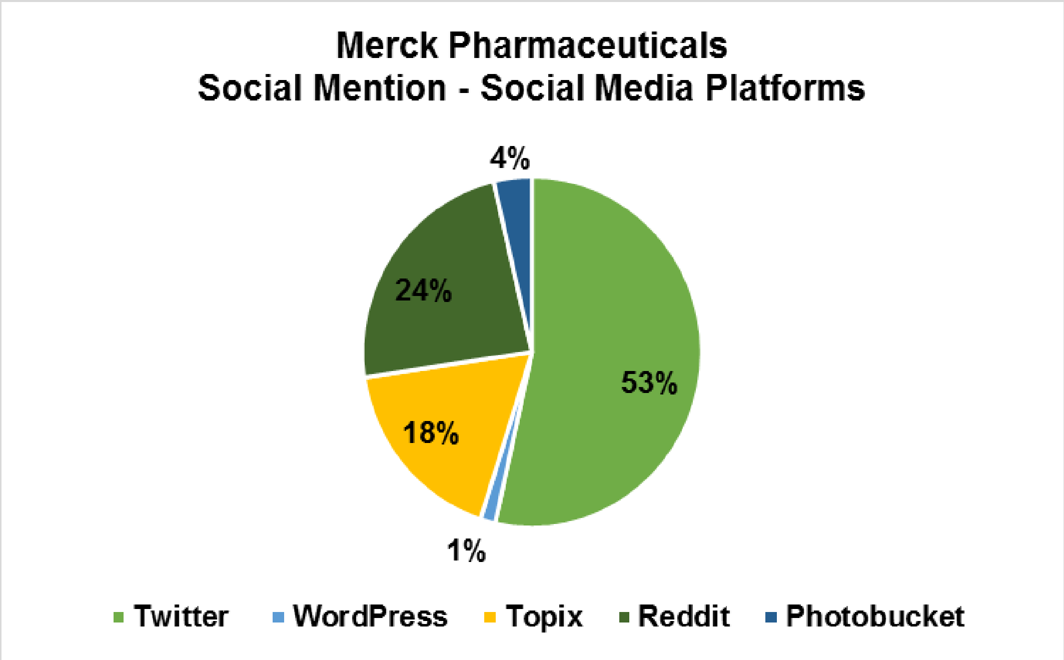

Figure 4.19 Social Mention – Social Media Platforms

4.4.7 Novartis Pharmaceuticals

Figure 4.20 is a graphical representation of the Chernoff Faces, the results of the diary study conducted for 30 days from 5 October 2018 to 4 November 2018. The datasets for the various variables as explained in table 4.1, were uploaded in Stata 15, to generate the 30 individual faces.

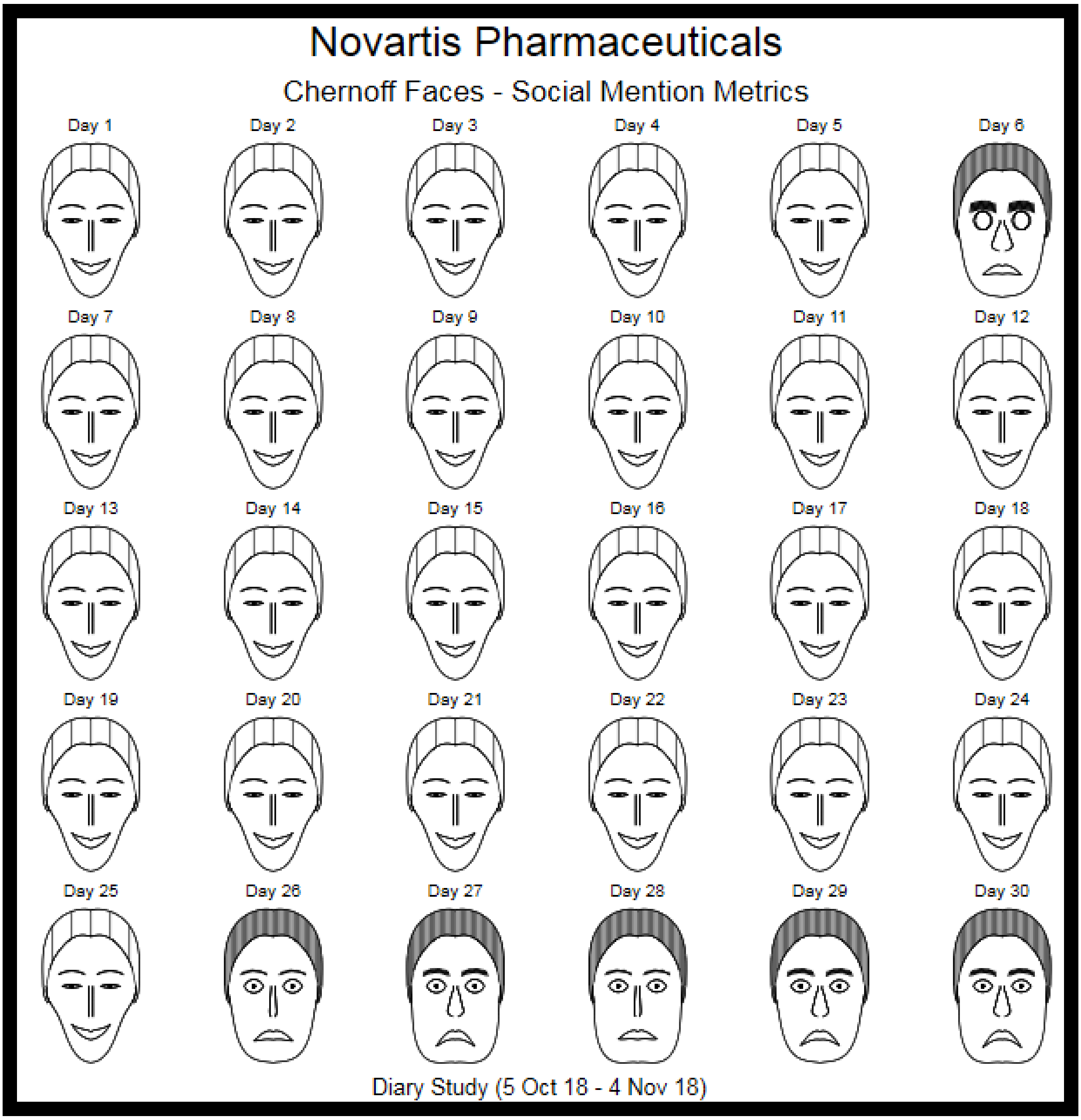

Figure 4.20 Novartis Pharmaceuticals, Chernoff Faces Graph (Diary Study 5 Oct 18 – 4 Nov 18)

4.4.7.1 Chernoff Faces

The faces for Novartis over the 30 days are all conspicuously homogeneous. For 24 of the days, the faces are identical. With 6 days being the outliers, with the faces being virtually identical.

The individual faces are analysed, as per the following 7 variables:

1) Positive Sentiment (positive mentions) – Eye Size: larger eyes = higher positive sentiment

For 24 of the days, positive sentiment is at minimum value, hence the sea of narrow thin eyes.

2) Negative Sentiment (negative mentions) – Pupil Size: larger pupil size = higher negative sentiment

For 29 of the days, the pupils of the images are constricted, due to negative sentiment being at minimum value. The negative mentions are virtually zero or very low. Only on day 6 are the pupils dilated, meaning, negative sentiment was at maximum value.

3) Strength (Likelihood brand is discussed on social media) - Face line: broader face = higher brand strength

What is visibly apparent is that the variable strength is the most pronounced. Over the 30 days, for 24 of the days, the brand strength has performed relatively poorly, therefore showing the trend of the long, thin faces. This means that it is very unlikely that the brand is being discussed on social media.

4) Passion (Individuals talking about the brand repeatedly) – Mouth Curve; higher passion = wider smile

For most of the days the passion was high, indicating that the same individuals were generally talking about the brand. 6 of the days reveal that the passion was low, which means that only 30% of the individuals were not the same people.

5) Reach (range of influence) – Brow Density: wider reach = denser brow

The reach for the period was only at maximum value on day 6. For the rest of the time the reach was generally narrow.

6) Relative Frequency (frequency of brand mentions) – Hair Density: greater relative frequency = darker hair

The brand mention frequency was generally sporadic only 6 of the days the mentions were within the 1- 6- hour band.

7) Unique Authors (number of authors messaging about the brand) – Nose Size: higher unique authors = bigger nose

The unique authors messaging about the brand was only 3% of the period.

4.4.7.2 Sources (Social Mention - Social Media Platform)

The results revealed, as per the datasets in chapter 3, table 3.3, shown here in figure 4.21 that the following social media platforms that users were most active on, on specific days. The activity on social media in general for this firm is very infrequent for this period. The activity happening on the six days in figure 4.21 is a clear tell tail sign of the faces presented in figure 4.20. The authors that were active were the same people hence the smiling mouths. It is evident that these few users were active on the social media platform, WordPress. While WordPress was the most frequented social media platform, users on Twitter were the most active.

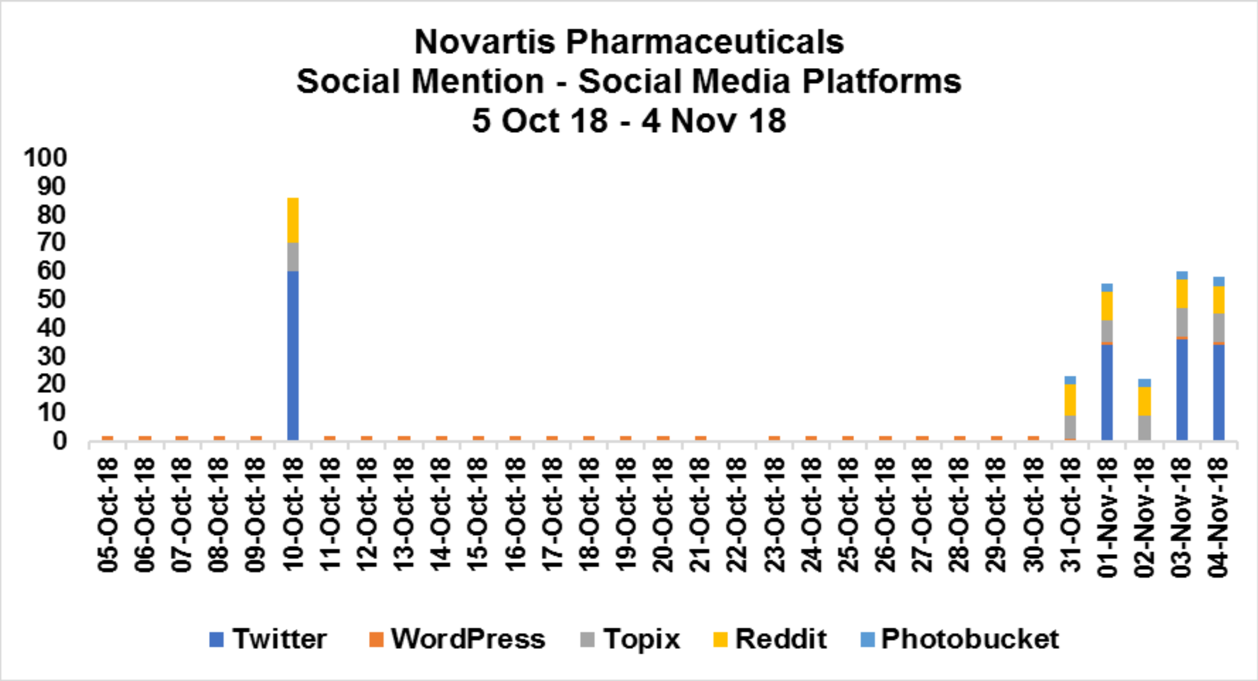

Figure 4.21 Social Mention Sources - Social Media Platforms for 5 Oct 18 – 4 Nov 18

During this period, as demonstrated in figure 4.22, the results show that users were only active on four social media platforms, namely: Twitter, WordPress, Topix, Reddit and Photobucket. The most active social media platform being Twitter.

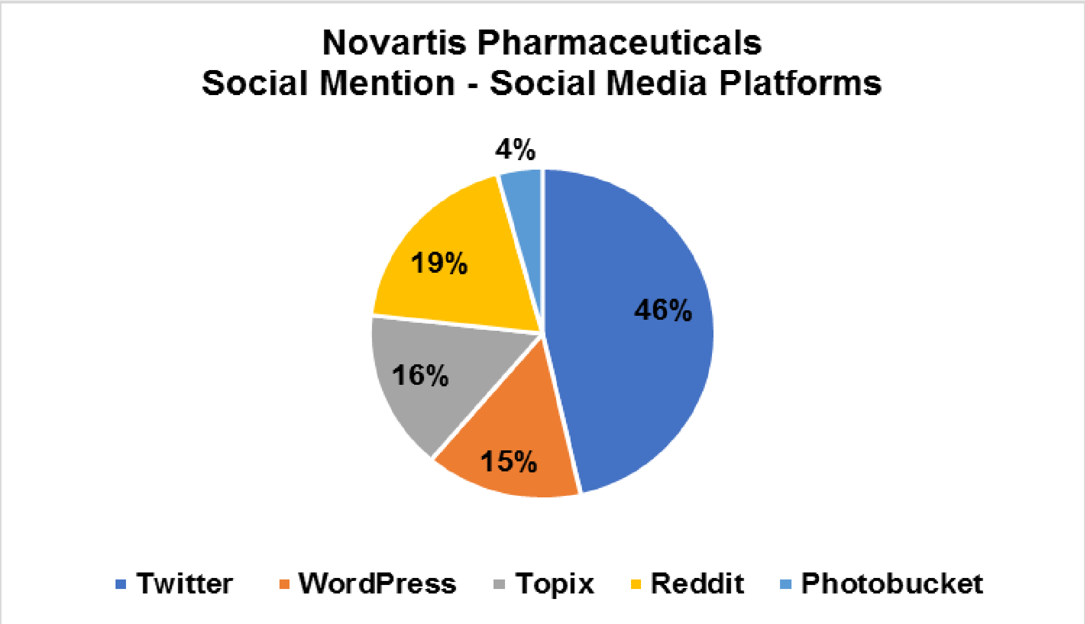

Figure 4.22 Social Mention – Social Media Platforms

4.4.8 Novo Nordisk Pharmaceuticals

Figure 4.23 is a graphical representation of the Chernoff Faces, the results of the diary study conducted for 30 days from 5 October 2018 to 4 November 2018. The datasets for the various variables as explained in table 4.1 were uploaded in Stata 15, to generate the 30 individual faces.

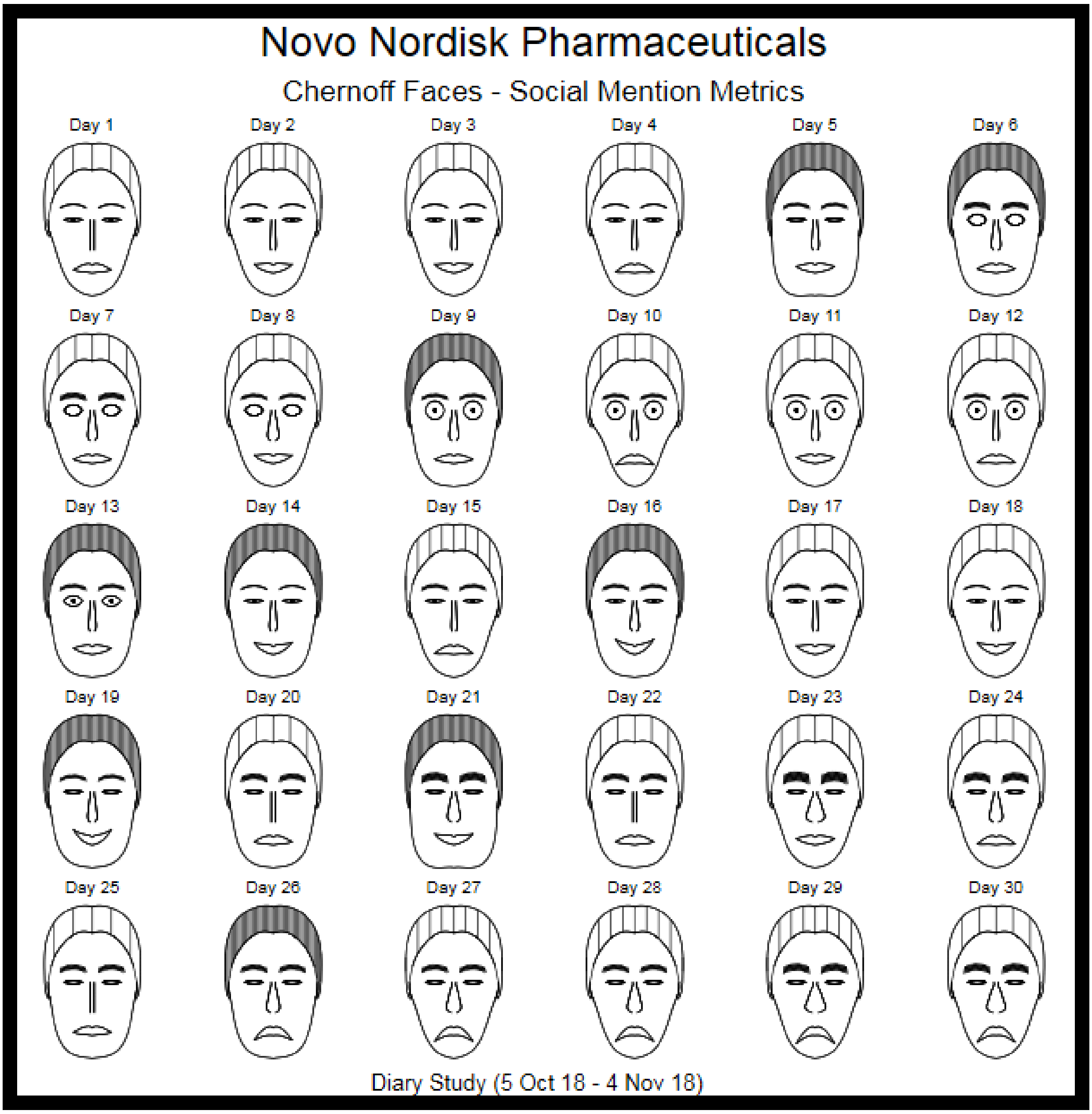

Figure 4.23 Novo Nordisk Pharmaceuticals, Chernoff Faces Graph (Diary Study 5 Oct 18 – 4 Nov 18)

4.4.8.1 Chernoff Faces

The faces for Novo Nordisk over the 30 days are generally similar. On day 10 and 21 the faces are at its extremes. Day 10 being at minimum value and day 21 being at maximum value for most of the variables.

The individual faces are analysed as per the following 7 variables:

1) Positive Sentiment (positive mentions) – Eye Size: larger eyes = higher positive sentiment

On days 9,10, 11, and 12 the eyes are big and round suggesting the variable is at maximum value for those days. For most of the time the eyes are at low values and therefore the many narrow appearing eyes.

2) Negative Sentiment (negative mentions) – Pupil Size: larger pupil size = higher negative sentiment

For almost half of the period there was no negative sentiment for this reason the faces on day 1,2,3,4,9,10,11,12,13,14,15,16,17 and 18 the pupils are constricted.

3) Strength (Likelihood brand is discussed on social media) - Face line: broader face = higher brand strength

The faces all mostly appear normal looking. This meaning that the values were all generally at midpoint. Suggesting that the brand strength was mediocre.

4) Passion (Individuals talking about the brand repeatedly) – Mouth Curve; higher passion = wider smile

For most of the days, the passion was average. The last 5 days shows that the passion was at its minimum values and 2 of the days it was at maximum values.

5) Reach (range of influence) – Brow Density: wider reach = denser brow

The reach was generally narrow as only 5 of the 30 days reflects dense brows.

6) Relative Frequency (frequency of brand mentions) – Hair Density: greater relative frequency = darker hair

The images show that for 60% of the time the mentions exceeded the 1-6-hour range hence so many faces with sparse hair.

7) Unique Authors (number of authors messaging about the brand) – Nose Size: higher unique authors = bigger nose

The unique authors messaging about the brand was only 10%.

4.4.8.2 Sources (Social Mention - Social Media Platform)

The results revealed, as per the datasets in chapter 3, table 3.3, shown here in figure 4.24 that the following social media platforms that users were most active on, on specific days.

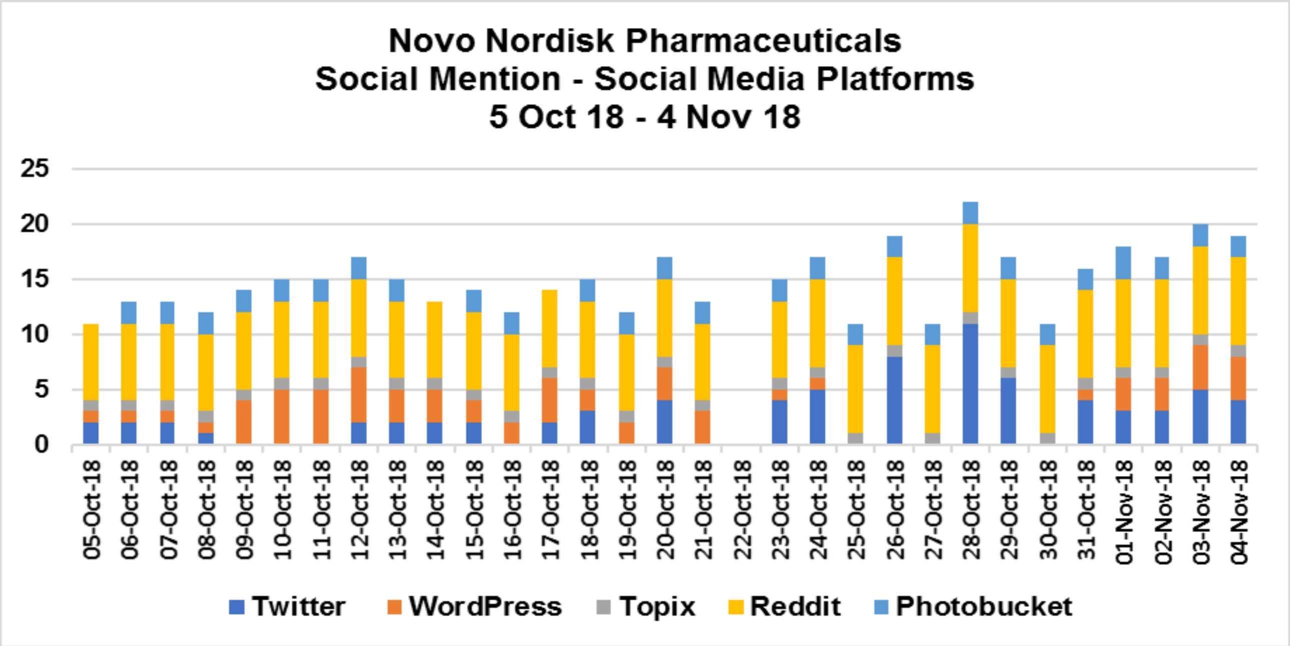

Figure 4.24 Social Mention Sources - Social Media Platforms for 5 Oct 18 – 4 Nov 18

During this period, as demonstrated in figure 4.25, the results show that users were only active on five social media platforms, namely: Twitter, WordPress, Topix and Reddit. The most active and frequented social media platform being Reddit refer figure 4.24.

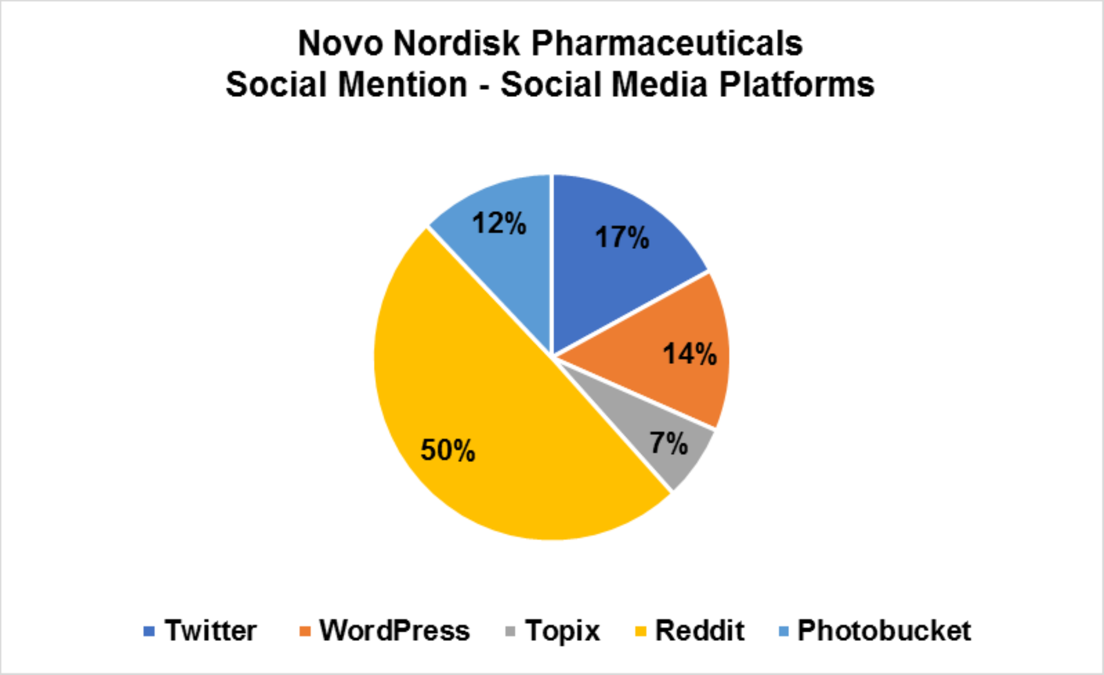

Figure 4.25 Social Mention – Social Media Platforms

4.4.9 Sanofi Pharmaceuticals

Figure 4.26 is a graphical representation of the Chernoff Faces, the results of the diary study conducted for 30 days from 5 October 2018 to 4 November 2018. The datasets for the various variables as explained in table 4.1 were uploaded in Stata 15, to generate the 30 individual faces.

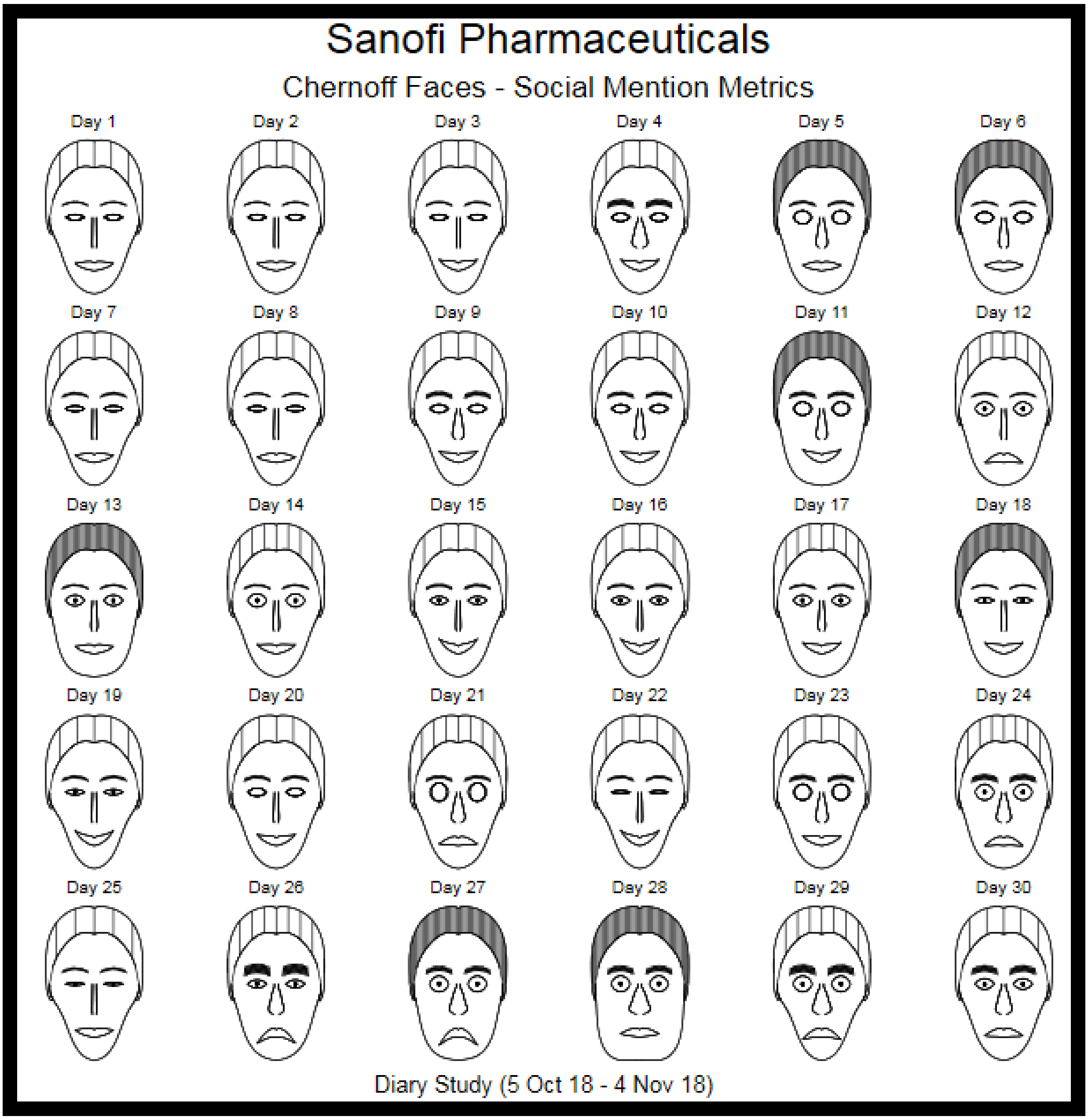

Figure 4.26 Sanofi Pharmaceuticals, Chernoff Faces Graph (Diary Study 5 Oct 18 – 4 Nov 18

4.4.9.1 Chernoff Faces

The faces for Sanofi over the 30 days are all generally different with most of the faces leaning towards narrow thin face lines. This meaning that the brand strength ranges from mediocre too low for the period.

The individual faces are analysed as per the following 7 variables:

1) Positive Sentiment (positive mentions) – Eye Size: larger eyes = higher positive sentiment

For 8 of the days the positive sentiment is at its highest values as on these days the eyes are generally round and big. The rest of the time the eyes are in between narrow and oval.

2) Negative Sentiment (negative mentions) – Pupil Size: larger pupil size = higher negative sentiment

50% of the pupils are constricted indicating that half of the negative sentiments for the time frame was at minimum values. Intimating that there were no negative mentions for 15 of the days.

3) Strength (Likelihood brand is discussed on social media) - Face line: broader face = higher brand strength

Most of the faces are leaning towards narrow thin face lines. This meaning that the overall brand strength ranges from mediocre too low for the period.

4) Passion (Individuals talking about the brand repeatedly) – Mouth Curve; higher passion = wider smile

Two of the days reveal that the passion was at minimum values, which means that about 7% of the individuals were not the same people. However, on the other days the variable values were generally at midpoint suggesting that individuals talking about the brand were generally the same people.

5) Reach (range of influence) – Brow Density: wider reach = denser brow

The reach was relatively narrow as only one face has brows at maximum thickness.

6) Relative Frequency (frequency of brand mentions) – Hair Density: greater relative frequency = darker hair

The brand mentions were infrequent as only 7 days show hair at maximum density.

7) Unique Authors (number of authors messaging about the brand) – Nose Size: higher unique authors = bigger nose

The unique authors messaging about the brand was only 13% for the period.

4.4.9.2 Sources (Social Mention - Social Media Platform)

The results revealed, as per the datasets in chapter 3, table 3.3, shown here in figure 4.27 that the following social media platforms that users were most active on, on specific days.

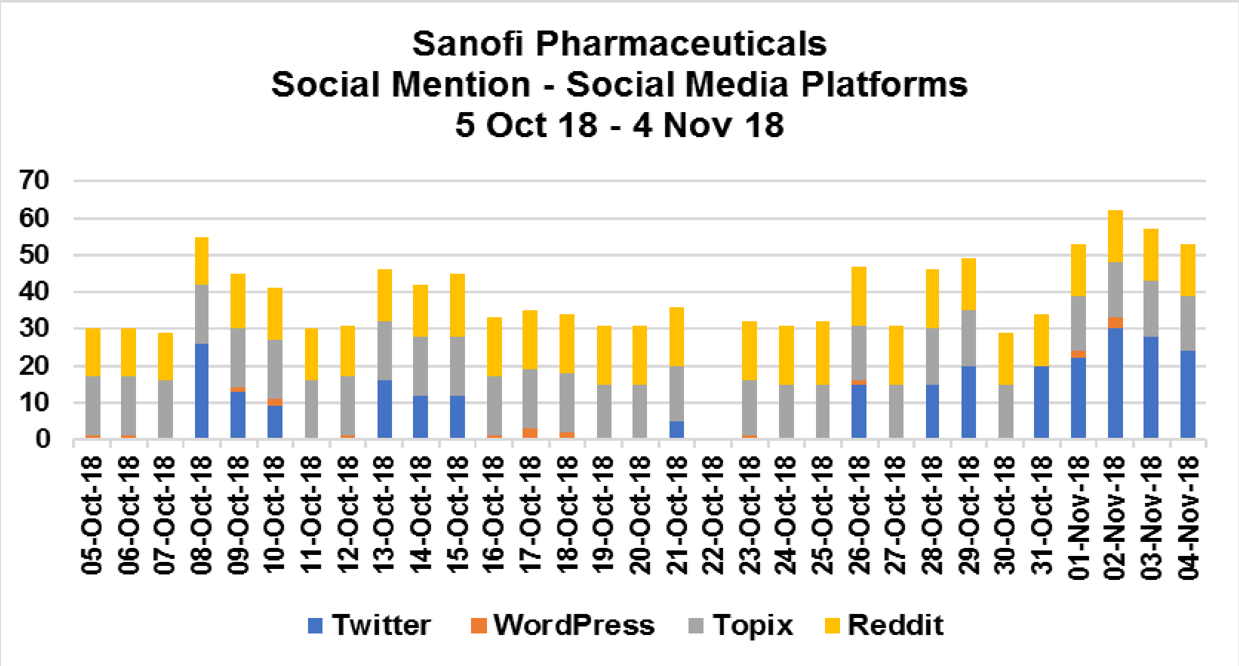

Figure 4.27 Social Mention Sources - Social Media Platforms for 5 Oct 18 – 4 Nov 18

During this period, as demonstrated in figure 4.28, the results show that users were only active on four social media platforms, namely: Twitter, WordPress, Topix and Reddit as reflected in figure 4.27. The most frequented and active social media platforms being Topix and Reddit figure 4.27 and figure 4.28.

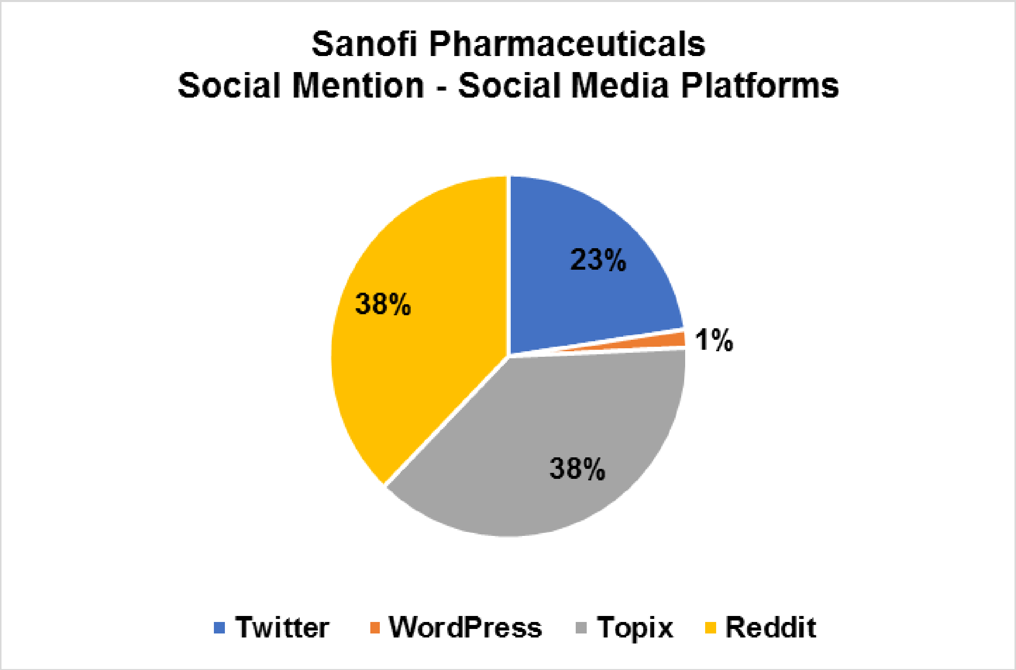

Figure 4.28 Social Mention – Social Media Platforms

4.4.10 Takeda Pharmaceuticals

Figure 4.29 is a graphical representation of the Chernoff Faces, the results of the diary study conducted for 30 days from 5 October 2018 to 4 November 2018. The datasets for the various variables as explained in table 4.1 were uploaded in Stata 15, to generate the 30 individual faces.

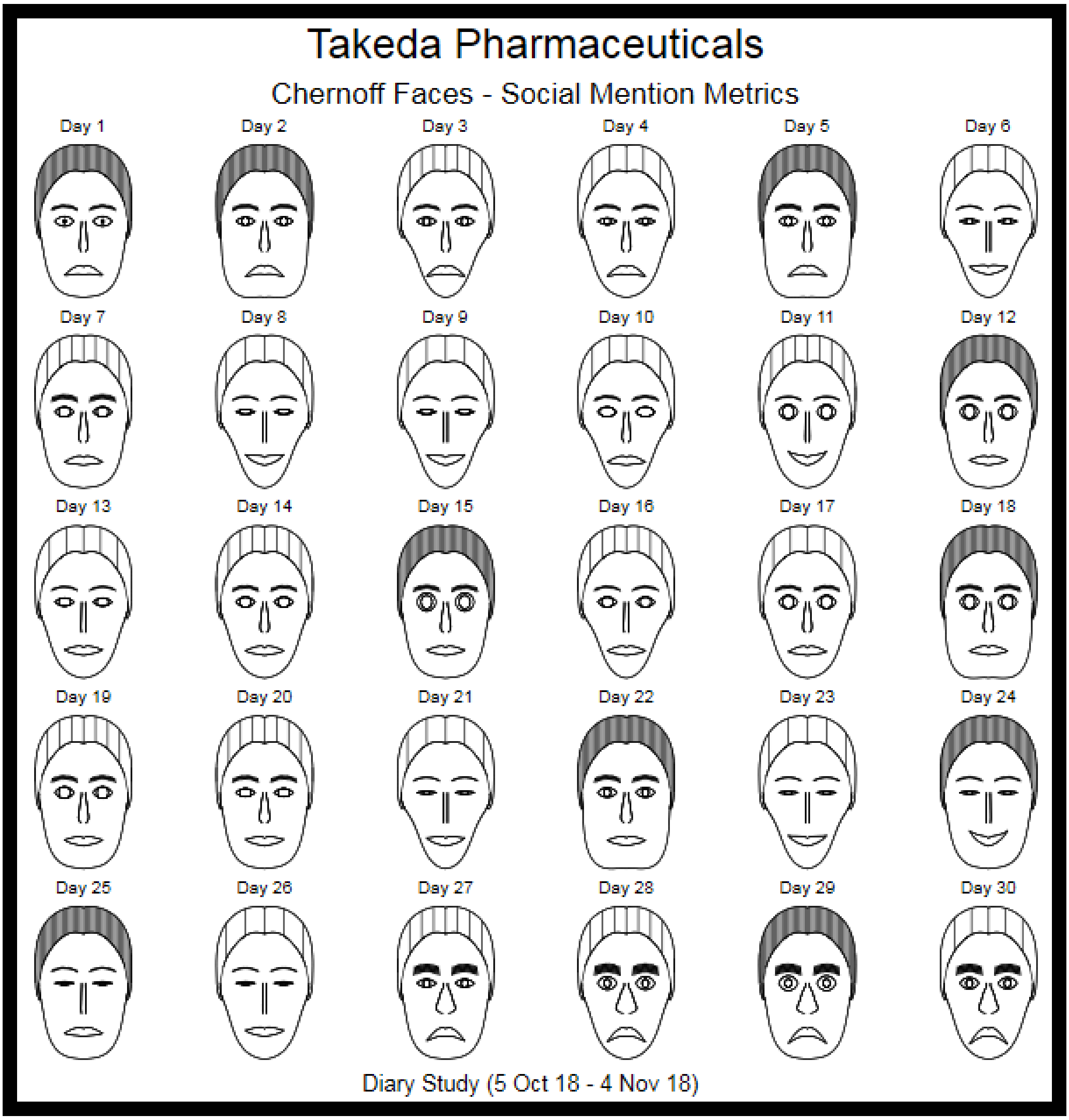

Figure 4.29 Takeda Pharmaceuticals, Chernoff Faces Graph (Diary Study 5 Oct 18 – 4 Nov 18)

4.4.10.1 Chernoff Faces

The faces for Takeda over the 30 days are all noticeably different. While most of the faces don sparse hair only one face has constricted pupils.

The individual faces are analysed as per the following 7 variables:

1) Positive Sentiment (positive mentions) – Eye Size: larger eyes = higher positive sentiment

Only 5 days of the firm's positive mentions are at minimum value, therefore the eyes are narrow. Hence, an indication that for most of the days the mentions leaning towards positive mentions than negative.

2) Negative Sentiment (negative mentions) – Pupil Size: larger pupil size = higher negative sentiment

Day 1 is the only day that there was no negative sentiment hence the pupil is constricted.

3) Strength (Likelihood brand is discussed on social media) - Face line: broader face = higher brand strength

The brand strength shows no images at extreme values meaning that the variable was relatively average over the period.

4) Passion (Individuals talking about the brand repeatedly) – Mouth Curve; higher passion = wider smile

For most of the days the passion was average. The last four days show the variable at the minimum values.

5) Reach (range of influence) – Brow Density: wider reach = denser brow

The reach was only 13% during the 30 days. This was evident during the last four days.

6) Relative Frequency (frequency of brand mentions) – Hair Density: greater relative frequency = darker hair

For one third of the time the relative frequency was at maximum values.

7) Unique Authors (number of authors messaging about the brand) – Nose Size: higher unique authors = bigger nose

The unique authors messaging about the brand was only 13% of the period.

4.4.10.2 Sources (Social Mention - Social Media Platform)

The results revealed, as per the datasets in chapter 3, table 3.3, shown here in figure 4.30 that the following social media platforms that users were most active on, on specific days.

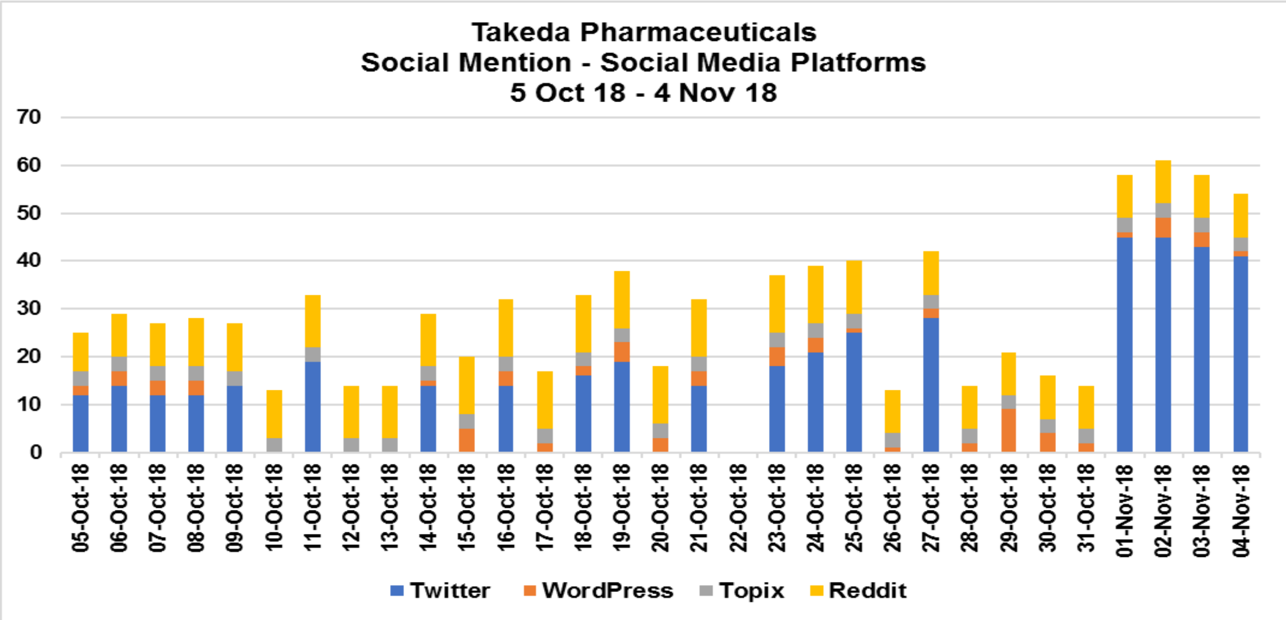

Figure 4.30 Social Mention Sources - Social Media Platforms for 5 Oct 18 – 4 Nov 18

During this period, as demonstrated in figure 4.31, the results show that users were only active on four social media platforms, namely: Twitter, WordPress, Topix and Reddit. The most active social media platform being Twitter and the most visited platform being Reddit refer figure 4.30 and 4.31.

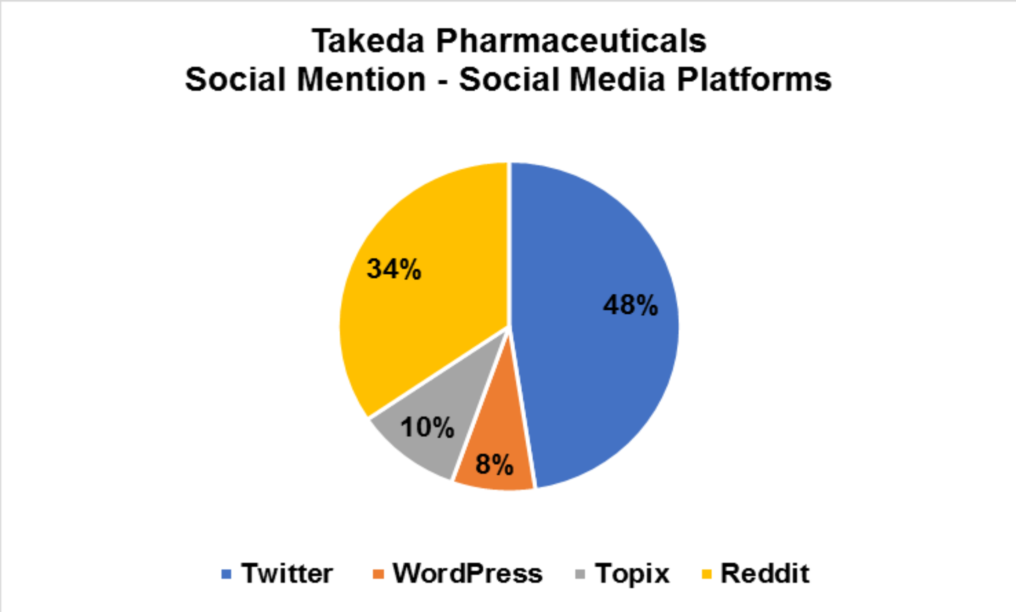

Figure 4.31 Social Mention – Social Media Platforms

4.5 Comparisons of Firms

4.5.1 Chernoff Faces

The researcher created a table listing all the firms, showing the sum of their Social Mention metric scores per variable. The lowest values for each variable was used as the minimum value and the highest score per variable was used as the maximum value as shown in table 4.2.

Table 4.2 Total metric scores per variable per firm

No.	Firm	Positive Sentiment	Negative Sentiment	Strength	Passion	Reach	Relative Frequency	Unique Authors
1	AstraZeneca	494	73	154	1061	674	1522.5	1159
2	Bayer	640	87	155	783	592	2900	1254
3	Boehringer Ingelheim	404	85	31	202	260	1083	530
4	Eli Lilly	306	90	57	510	296	1939.5	631
5	Johnson & Johnson	97	19	65	58	190	1474	303
6	Merck	676	27	217	686	630	2800	1292
7	Novartis	62	3	15	1294	155	948	237
8	Novo Nordisk	222	16	43	626	162	1529	329
9	Sanofi	410	45	38	626	412	1439.5	581
10	Takeda	227	134	40	931	328	1672.5	587
	Minimum Value	62	3	15	58	155	948	237
	Maximum Value	676	134	217	1294	674	2900	1292

The maximum and minimum values were uploaded in Stata 15 to generate the two extreme Chernoff Faces to use as reference points refer figure 4.32.

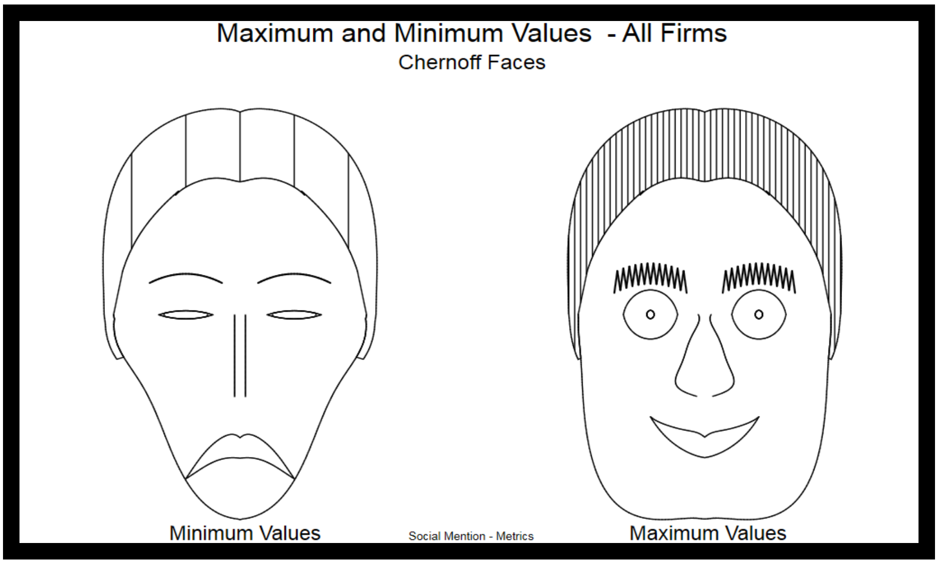

Figure 4.32 Chernoff Face: Maximum and Minimum Value Total Social Mention Metric Scores

Table 4.2 is a representation of the individual firms total scores, per variable over the period 5 Oct 18 – 4 Nov 18. These scores were taken from the individual firm's datasets totals, to create table 4.2. The data was then uploaded in Stata 15 to generate the 10 Chernoff Faces depicting the various firms as seen in figure 4.33.

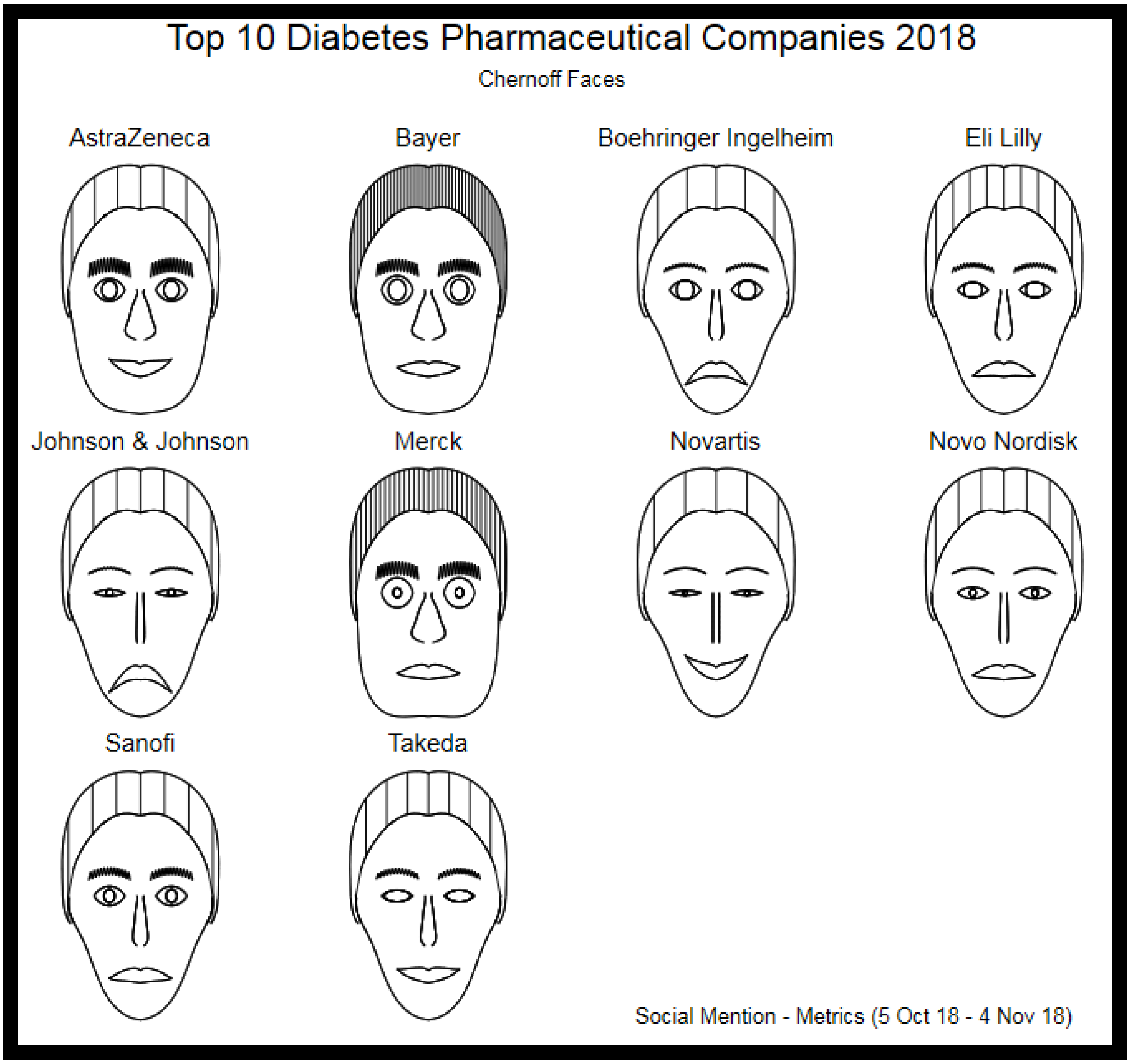

Figure 4.33 Top 10 Pharmaceutical Firms (Total Social Mention Metric Scores 5 Oct 18 – 4 Nov 18)

At face value 7 of the 10 firm's Social Mention Metric scores, as depicted by their faces in figure 4.33, indicates that the firms scores are at minimum value. Three of the firms' scores are at maximum values, noting however, that this does not apply to all the variables. Glaring missing from Merck Pharmaceuticals is the smiling mouth representing passion (individuals talking about the brand repeatedly). When studying the 30 faces of each firm the combined scored faces is not significantly different from the trends identified.

4.5.2 Sources (Social Mention – Social Media Platforms)

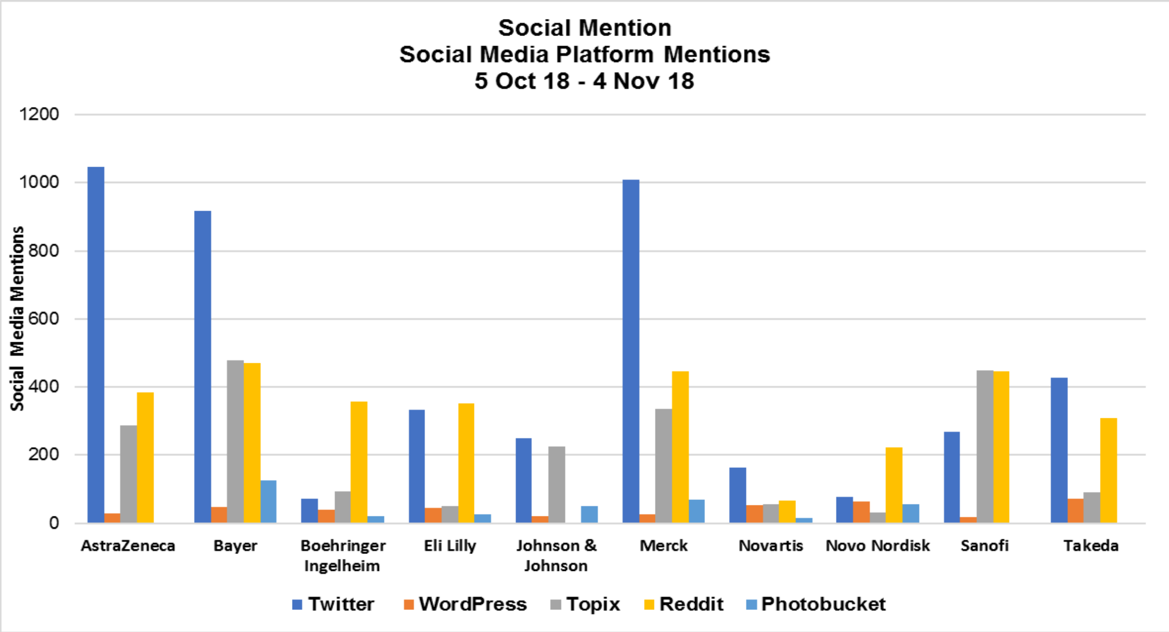

Figure 4.34 Social Mention – Social Media Platforms Mentions Per Firm

Even though Twitter was not the most visited platform, the users on this platform was the most active. With AstraZeneca, Bayer and Merck leading the pact by far with mentions, as visibly observed in figure 4.34.

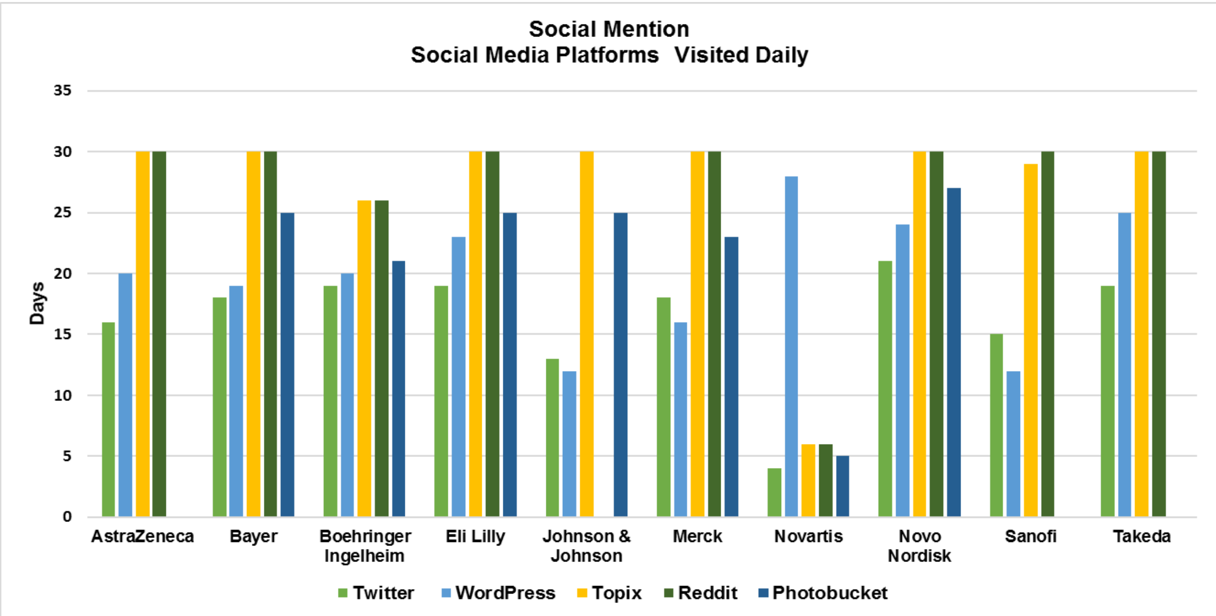

Figure 4.35 Social Mention – Social Media Platforms Visited Daily

The social media platforms visited daily by users during the period 5 Oct 18 – 4 Nov 18 were Topix and Reddit as seen in figure 4.35. While these platforms were frequently visited, they were not the most active as indicated in figure 4.34. AstraZeneca, Bayer and Merck are the firms leading with posts on Twitter, figure 4.35 shows that during the study period Twitter was not visited daily. The data shows that AstraZeneca, Bayer and Merck were visited for 16,18 and 18 days respectively during this time.

4.6 Social Media Platforms and Pharma

4.6.1 Active Platforms

The study results indicates that Internet users were active on an array of platforms, other than the social media platforms Pharma is generally posting on namely: Twitter, Facebook, YouTube, LinkedIn, Instagram and Pinterest as identified in the literature (Canvin, Toms, Evans, McClure, Polimeno and Shah, 2016; Narayan, 2017; Spitz and Einarsen, 2017; Partikas, Toms, Evans, Demuren and Shah, 2018; Cornejo, 2018; Kemp, 2018) with the exception of Twitter. Social Mention (2018) claims to amass data from more than 80 social media platforms. The results of this research shows that, during the study period, the user-generated content, was aggregated from 5 social media platforms namely: Twitter, WordPress, Topix, Reddit and Photobucket across the 10 firms, as revealed in figure 4.35.

A recent report by Partikas, et al. (2018), which included the 10 firms of this study sample, showed that the average weekly posts on Facebook, Twitter, YouTube and Instagram was less, compared to the 2016 report (Canvin, et al., 2016). Citing the reasons for the decline as Pharma's strategy of focusing on posting valuable quality content rather than the quantity of posts, hence the concept of "be visible but not noisy". In other words, post only when there is a powerful narrative to communicate. The report showed that companies with the most posts on the various social media platforms, did not necessarily achieve the highest engagement scores. The reported results, Partikas, et al. (2018), indicated that Novo Nordisk achieved the highest engagement scores, even though their number of posts was below average. While AstraZeneca had a high number of posts, their engagement score was below average.

The social media platform results in this study confirms that firms with the most posts (figure 4.36) does not necessarily have the highest level of engagements (figure4.35) as in the case of AstraZeneca, Bayer and Merck's Twitter results. It is therefore essential that firms invest more time creating and custom-building social media content that is relevant for the various platforms and that will appeal to their respective target audiences, rather than spending time posting frequently (Partikas, et al., 2018).

4.6.2 Platform Content

As a result of the heavily regulated environment Pharma is not only choosy on which social media platforms it is operating on. Firms are also very judicious about the content they share. According to Narayan (2017, p.2) there are 4 areas that Pharma generally engage in i.e. corporate social profiles, careers in Pharma, OTC (over the counter) brand profiles and branded community properties. Table 5.1 is a reconstructed highlighting only the 10 firms in this study showing the areas they are mostly conversing about on social media.

Table 4.3 Firms social media landscape

No.	Firms	Corporate Social Profiles	Careers in Pharma	OTC Brand Profiles	Branded Community Properties
1	AstraZeneca	*	*	*	
2	Bayer	*		*	
3	Boehringer Ingelheim	*		*	
4	Eli Lilly	*			
5	Johnson & Johnson	*	*	*	*
6	Merck	*	*		*
7	Novartis	*	*		
8	Novo Nordisk	*			*
9	Sanofi	*		*	*
10	Takeda	No data available			

Cornejo (2018) reckons that robust social media engagement can be accomplished in a regulated milieu. Recognising that risks associated with unmoderated social media is not just about the company's reputation but also the possible exorbitant regulatory infringements. Cornejo (2018) says and Spitz and Einarsen (2017) concur, to address these risks Pharma must do the following proactively and not when a catastrophe erupts:

- Firms regulatory units must be involved in social media strategies from the onset
- Make sure key topics are addressed timeously e.g. adverse events, product complaints, negative and positive comments, off-label use etc.
- Plan how to respond to possible events in advance. This allows for quick action and damage control

Pharma social media content that span the entire social media space achieves greater engagement (Partikas, et al., 2018). With social media content trends for 2017 showed :

- Highlighting the human side of the business was common among the high performing posts
- Awareness day e.g. World Diabetes Day, content drove greater engagement across the social space
- Celebrity involvement in causes increased engagement
- Innovative technologies keep content stimulating

The various social media platforms have made it possible to access the opinions, perspectives and experiences of a cosmic pool of people. Conversely people are also able to express their opinions, share their experiences (good and bad) and articulate their perspectives (Pang and Lee, 2008; Gupta, Tyagi and Sharma, 2013). While this study does not display a definitive result of the user generated content. Firms can use the results as a quick gauge of users social media sentiment (Raciborski, 2009; Farshid, Chan and Nel, 2012). Understanding social media sentiment and constantly analysing it, is vital for firms. This user generated content can assist firms with making important strategic business decisions (Mukhopadhyay, 2018; Mgudlwa and Iyamu, 2018b).

4.7 Big Data and Pharma

The exponential upsurge of social media has been at the epicentre of the megatrend of big data. The literature indicates that big data can be a source of social value when it provides people with better healthcare and economic value when organisations can measure an increase in profits, business growth and competitive advantage (Günther, et al., 2017). The McKinsey Global Institute (Manyika, et al., 2011) has projected that the adoption of big data strategies, for sound decision-making, could generate 300 billion dollars per annum for the US healthcare economy. The latest report (Henke, et al., 2016) shows that only between 10-20% of the potential identify in the 2011 report was realised by the US healthcare industry. Citing a lack of incentives, the difficulty of process and organizational changes, a shortage of technical talent, data-sharing challenges and regulations as impediments to adoption. Cornejo (2018) has identified how firms can overcome regulation impediments.

The RBT and DCT has illuminated what tangible and intangible resources and capabilities Pharmaceutical firms will require to leverage social media big data as a source of sustained competitive advantage (Wernerfelt, 1984; 1989; Barney, 1991; Grant, 1991; Barney, 1995; Teece, Pisano and Shuen, 1997; Teece, 2018). What is apparent in the literature is that social media big data on its own is not a source of sustained competitive advantage for firms (Gupta and George, 2016; Wang and Hajli, 2017; Grover et al., 2018). The literature indicates that firms wanting to adopt big data initiatives need to firstly develop a business strategy that includes all the business units (Mazzei and Noble, 2017; Pisano, 2015; Teece, 2018). LaValle, et al. (2011) cautions that big data initiatives should not be embarked on without strategic business direction.

Grant (1991) provides a practical framework with five ground rules for resource strategy analysis as depicted in figure 2.2 . Barney (1991) VRIN framework provides a dais for resources to be analysed as sources of sustained competitive advantage figure 2.1. These frameworks allows for internal corporate strategy discussions and the development of social media big data initiatives. Clarifying how firms can leverage resources and capabilities to improve the business value (McAfee, Brynjolfsson and Dearstyne, 2012).

Mazzei and Noble (2017) posits that once firms understand what value they want to derive from there social media big data initiatives only then the technical aspects of the process should be considered. Data sources, capture, storage, intergration, transformation, analysis and visualisation (Chen, Chiang and Storey, 2012; Wang, Wang and Alexander, 2015; Wamba et al., 2015; Oussous,Benjelloun, Lahcen and Belfkih, 2018). Sivarajah, et al. (2017) says the biggests challenge of big data is in the analysing of it in a way that firms will derive big value. Yi, et al. (2014) argues that with the availaibility of innovative big data technologies valuable insights in sectors like healthcare can be achieved. Noting that this value is derived from a big data technological aspect. This advantage lies in the firms ability to turn the data into useful business intelligence and then take the relevant action to realise this value . It is the firm's ability namely: big data analytics that will lead to superior performance (Gupta and George, 2016; Wang and Hajli, 2017; Braganza, et al., 2017; Teece, 2018). Braganza, et al. (2017) drawing on empirical evidence gives a step by step business process that can be followed to ensure successful adoption of big data initiatives.

4.8 Big Data Analytics and Pharma

Big data and big data technology are not sources of sustained competitive advantage (Mazzei and Noble, 2017). Teece (2018) dynamic capabilities framework and Gupta and George (2016) recommended seven resources that will allow firms to develop a big data analytic capability enables firms to identify and exploit unique resources that will lead to sustained competitive advantage. Big data and big data analytics meticulous synergy are resources that will lead to valuable business insights (Günther, et al., 2017).

4.9 Chernoff Faces and Big Data

The need for data to be easily understood and act upon timeously is key for firms to be more data driven. Leaders want improved ways of communicating complex information (LaValle, et al., 2011). Data visualisation allows for the valuable insights to be interpreted effortlessly resulting in decisions to be made timeously and opportunities to be seized swiftly. Manyika, et al. (2011) defines data visualisation as technologies used for producing images, diagrams, or animations to convey a point used to synthesize the outcomes of big data assays. Data is displayed in numerous ways e.g. scatter diagrams, pie charts, histograms and bar charts. A typical graph generally depicts absolute numerical data whereas images are designed to spot clusters, categorize and arrange variables (Farshid, Chan and Nel, 2012). One of the images studied is a facial technique by Chernoff (1973).

The Chernoff Faces (Raciborski, 2009) were used to represent the data in this study. The faces proved to be a novel and easy way to interpret the data. Two extreme images were created to use as a reference point for the individual companies and for the companies' combined results. According to

4.10 Conclusion

While Chernoff Faces is a very novel way to analyse big data, recognising the variables and what they represent comprehensively, is imperative. This allows for the reader to analyse the faces swiftly.

Overall, the faces for the 30 days across the ten firms, all lean towards a narrow, longish, thin image. The reason for this, is that the face line of an image is the most pronounced facial feature, therefore spotting a broad or thin face, is effortless. This implies that more than 70 percent of the firms' social media strength over the period, was generally weak as depicted in figure 4.33, page 95.

The literature shows that social media big data and big data analytics, through the lens of RBT and DCT, can be a source of sustained, competitive advantage. The research results and literature has shown that, while the firms in the study all have a presence on social media, it is on selective platforms and the content that is posted is on very specific topics (Narayan, 2017; Cornejo, 2018).

The overall results of the Chernoff Faces depict that the firms' Social Mention metrics, over the diary study period, was generally at low values. Hence, the thin, long, generally sad looking faces, refer figure 4.33 page 95.

The Pharmaceutical industry is described as being slow in taking up the social media big data initiatives, due to several barriers. The literature indicates ways in which the industry can circumvent these impediments and realise substantial business value from this unprecedented, global phenomenon (Cornejo, 2018).

5 Chapter 5: Conclusion

5.1 Introduction

This paper suggests that Pharma companies can use social media big data and big data analytics, to gain a competitive advantage in the healthcare sector. Ray (2017), purports that 1 in 20 google searches are enquiries about healthcare matters. Therefore, for firms to derive benefit from the proliferation of information on the Internet, they need to develop their internal resources and capabilities to analyse the data, for it to lead to a sustained competitive advantage.

5.2 Discussion

The goal of the research is to demonstrate how Pharma can use social media big data, to make strategic business decisions, through the lens of RBT and DCT, that could lead to a sustained competitive advantage. In and of its own, big data, does not constitute a competitive advantage. It may hold value for the firm, but lacks rarity, inimitability, and is not substitutable (Braganza, et al. 2017; Mata, Fuerst and Barney, 1995; Delmonte, 2003). It is in the analysis of this data, through RBT and DCT, which turns the information into useful business intelligence (Amit and Schoemaker, 1993; Barney, 1991; 1995; Marr, 2015; Gupta and George, 2016; Kurtmollaiev, et al., 2018). Most importantly, firms must constantly reconfigure their resources in line with the dynamic business environment to ensure superior performance (Teece, Pisano and Shuen, 1997; Helfat, et al., 2007; Teece, 2014; 2018).

The researcher has illustrated how Pharmaceutical firms can use social media platforms, to gather and analyse existing data. ‘Social Mention’ is the platform that was used to gather various social media sentiments, extracted from more than 80 social media sites (Social Mention, 2018). These media sentiments, namely: positive mention, negative mention, strength, passion, reach, relative frequency and unique authors, have been depicted by animated faces, known as Chernoff Faces. The Chernoff Faces graphically represented the multivariate data, with each Social Mention metric depicted by different facial features, e.g. Positive Mention = eye size, Negative Mention = pupil size, Strength = facial line, Passion = mouth curve, Reach = eyebrow density, Relative Frequency = hair density, Unique Authors = nose size.

The key actors in the healthcare industry are Pharmaceutical Companies, Regulatory Authorities, Health Care Practitioners and Patients. They are an integral part of the Pharmaceutical firm’s external environment. The Regulatory Authorities are the main reason why Pharma is so cautious in engaging social media and patients are the most avid users and creators of health-related content. They play a pivotal role in the generation and creation of big data on social media platforms (Aitken, 2014; Gupta, Tyagi and Sharma, 2013; Hawkins, DeLaO and Hung, 2016; Unmetric, 2017). For firms to make sound business decisions, data sourced externally, must be used in conjunction with the firm’s internal data and resources. Thus, the researcher shows how the amalgamation of external data and internal resources and capabilities, can lead to a company’s sustained competitive advantage.

Qualitative and quantitative research are often presented as two fundamentally different paradigms. However, in this study, a qualitative approach was used to examine the RBT (VRIN Framework) and DCT, to describe and understand the relevant theories and to build upon the quantitative results (de Vos, Strydom, Fouche and Delport, 2011), while a quantitative approach was used to analyse the social media sentiment, as depicted by Social Mention metrics and interpret the Chernoff Faces.

The research results show that, while the 10 firms in the study all have a presence on social media, it is on selective platforms and the content that is posted, is on very specific topics (Narayan, 2017; Cornejo, 2018). The overall results of the Chernoff Faces indicate that the firms’ Social Mention metrics, over the 30 day period, was generally at low values. Since strength of social mention is depicted by the face line, the thin, long, generally sad looking faces implies that more than 70 percent of the firms’ social media strength over the study period, was generally weak.

5.3 Limitations of the Study

This is not a definitive study of social media big data. It is an illustration of how Pharma can use secondary data, available on social media platforms, like Social Mention, to integrate, develop and reconstruct their dynamic capabilities and resources, for efficient and effective strategic, digital, multichannel marketing decisions. There is a plethora of social media big data repository firms, which can be accessed for free or a fee. Notably missing from the data collected from Social Mention is data from social media platforms such as Facebook, Instagram, LinkedIn, Pinterest and YouTube. On its website, Social Mention indicates that the information is available for free, for personal and non-commercial use (Social Mention, 2018). This may explain the omission of the missing platforms. The various big data sources enable firms to correlate and find ties between the copious amounts of data. This enables them to draw valuable insights, to make sound business decisions. The availability and inclusion of data from these social media platforms, may have resulted in a completely different outcome for the 10 firms.

5.4 Recommendations for further research

According to Gupta and George (2016), there has been an over-emphasis on technical aspects of big data in the literature and not sufficient on other resources, such as human skills and organizational culture. LaValle, et al. (2011), emphasizes that the big data adoption challenges practitioners face, is not related to data and technology, but managerial and cultural. Mazzei and Noble (2017) agrees by saying that the literature concentrates on how big data will influence management research, instead of investigating how big data is revolutionising the critical thinking processes of business leaders. The literature shows that sustained competitive advantage can mainly be derived from how firms align and fit their resources. This process is not static, it is robust and fluid (Wernerfeldt,1984; Barney, 1991; Grant, 1991; Teece, 1997; 2018; Helfat and Martin, 2015; Kurtmollaiev, Pedersen, Fjuk and Kvale, 2018). One could argue that this is so, because practioners may be reluctant to provide detail, as this may reveal pertainent resources, business processes and dynamic cababilities to rivals.

According to a report by the McKinsey Global Institute, by 2018, the US will confront a dearth of 140 000 to 190 000 people with natural analytical skills and 1,5 million supervisors with the aptitude to analyse big data, to make sound business decisions (Manyika, Chui, Brown, Bughin, Dobbs, Roxburgh and Hung Byers, 2011; Cattell, Chilikuri and Levy, 2013 and Henke, Bughin, Chui, Manyika, Saleh, Wiseman and Sethupathy, 2016). Research could be done to understand what and how the various education systems and learning institutions are augmenting their curricula to address this phenomenon.

For the purposes and augmenting of this study, further research could be pursued to evaluate what the 10 firms' social media standing is on the platforms that they are most active on, namely: Facebook, Instagram, YouTube, LinkedIn and Pinterest.

5.5 Conclusion

The Pharmaceutical industry has been described as being slow in taking up the social media big data initiatives, due to several barriers. The most obvious barrier being industry regulation. For the HCP, barriers such as patient confidentiality, malpractise lawsuits, etc. represents risk. The literature and the results of this study, indicates ways in which the industry can circumvent these impediments and realise substantial business value from this unprecedented, global phenomenon of big data and big data analytics. The data visualisation, as depicted by the Chernoff Faces, shows that the overall results and findings of the firms' social media sentiment values on Social Mention for the 30 days, were generally low. The literature indicates that the true value of the results can only be realised if firms make sound decisions and act swiftly.

References

Aitken, M., 2014. *Engaging patients through social media*. [pdf] IMS Institure for Healthcare Economics. Available at: <http://magazine.imshealth.it/wp-content/uploads/2015/09/IIHI_Social_Media_Report.pdf> [Accessed 5 May 2018].

Amit, R. and Schoemaker, P.J.H., 1993. Strategic assets and organizational rent. *Strategic Management Journal*,14(1), pp.33-46.

Barney, J.B., 1991. Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), pp.99–120.

Barney, J.B., 1986a. Organizational Culture: Can It Be a Source of Sustained Competitive Advantage? *Academy of Management Review*, 11(3), pp.656–665.

Barney, J.B., 1986b. Strategic Factor Markets: Expectations, Luck, and Business Strategy. *Management Science*, 32(10), pp.1231–1241.

Barney, J.B., 1986c. Types of Competition and the Theory of Strategy: Toward an Integrative Framework. *Academy of Management Review*, 11(4), pp.791–800.

Barney, J.B., 1995. Looking inside for competitive advantage. *Academy of Management Perspectives*, 9(4), pp.49-61.

Barney, J.B. and Arikan, A.M., 2001. The Resource Based View: Origins and Implications.In: M.A. Hitt. *The Blackwell HandBook of Strategic Management.* Oxford: Blackwell Publishing. Ch.5.

Barney, J.B. and Clark, D.N., 2007. *Resource-Based Theory: Creating and Sustaining Competitive Advantage*. New York: Oxford University Press.

Barney, J.B., Ketchen, D.J. and Wright, M., 2011. The future of resource-based theory: Revitalization or decline? *Journal of Management*, 37(5), pp.1299-1315.

Benetoli, A., Chen, T.F. and Aslani, P., 2018. How patients' use of social media impacts their interactions with healthcare professionals. *Patient Education and Counseling*, 101(3), pp.439–444.

Bharadwaj, A.S., 2000. A Resource Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation. *MIS Quartely*, 24(1), pp.169-196.

Braganza, A., Brooks, L., Nepelski, D., Ali, M. and Moro, R., 2017. Resource management in big data initiatives: Processes and dynamic capabilities. *Journal of Business Research*, 70, pp.328–337.

Canvin, R., Toms, C., Evans, R., McClure, P., Polimeno, G. and Shah, S., 2016. *The Social Check-up September 2016*. [online] Available at: <https://www.pulsarplatform.com/resources/social-check-up-the-top-20-Pharma-companies-on-social-media/pdf> [Accessed 13 May 2018].

Casadesus-Masanell, R. and Ricart, J.E., 2011. How to design a winning business model. *Harvard Business Review*. [online] Abstract only. Available through: https://www.hbs.edu/faculty/Pages/item.aspx?num=38951 [Accessed 18 May 2018].

Cattell, J., Chilikuri, S. and Levy, M., 2013. *How Big Data Can Revolutionize PharmaceuticalR&D*.[online]Availableat:<https://www.mckinsey.com/industries/Pharma

ceuticals-and-medical-products/our-insights/how-big-data-can-revolutionize-Pharmaceutical-r-and-d> [Accessed 18 May 2018].

Chen, Chiang and Storey, 2012. Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, 36(4), pp.1165–1168.

Chen, J., Chen, Y., Du, X., Li, C., Lu, J., Zhao, S. and Zhou, X., 2013. Big data challenge: A data management perspective. *Frontiers of Computer Science*, 7(2), pp.157-164.

Chernoff, H., 1973. The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association*, 68(342), pp.361-368.

Cornejo, C., 2018. *Pharma Social Media Best Practices: Reaching Out Within Regulatory Restrictions*.[online] Available at:<https://www.pulsarplatform.com/resources/social-check-up-the-top-20-Pharma-companies-on-social-media/ [Accessed 11 Nov. 2018].

Côrte-Real, N., Oliveira, T. and Ruivo, P., 2017. Assessing business value of Big Data Analytics in European firms. *Journal of Business Research*, 70, pp.379–390.

Davenport, T., 2014. Big Data at work. Dispelling the myths, uncovering the opportunities. *Harvard Business Review*.[webinar]. Available at: https://hbr.org/2014/03/big-data-at-work-dispelling-the-myths-uncovering-the-opportunities/webinar [Accessed 18 November 2018].

Davenport, T.H., 2013. *Big Data and the Role of Intuition*. [online] *Harvard Business Review Digital Articles*. Available at: <https://hbr.org/2013/12/big-data-and-the-role-of-intuition> [Accessed 18 November 2018].

Delmonte, A.J., 2003. Information Technology and the Competitive Strategy of Firms. *Journal of Applied Management and Entrepreneurship*, 8(1), pp.115-129.

Dezzani, L., 2018. *Top 10 Pharmaceutical Companies 2018 - Diabetes*. [online] Available at: <https://www.igeahub.com/2018/05/19/top-10-Pharmaceutical-companies-2018-diabetes> [Accessed 17 June 2018].

Dierick, I. and Cool, K., 1989. Asset Stock Accumulation and Sustainability of Competitive Advantage. *Management Science*, 35(12), pp.1504–1511.

Doherty, N.F. and Terry, M., 2009. The role of IS capabilities in delivering sustainable improvements to competitive positioning. *Journal of Strategic Information Systems*, 18(2), pp.100-116.

Farshid, M., Chan, A. and Nel, D., 2012. A sweet face man: using Chernoff Faces to portray social media wine brand images. *International Journal of Wine Business Research*, 24(3), pp.183–195.

Fischbach, S. and Zarzosa, J., 2018. Big data on a smaller scale: A social media analytics assignment. *Journal of Education for Business*, 93(3), pp.142–148.

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G. and Gnanzou, D., 2015. How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165, pp.234-246.

Fosso Wamba, S.F., Gunasekaran, A., Akter, S., Ren, S.J. fan, Dubey, R. and Childe, S.J., 2017. Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70(2017), pp.356-365.

Fox, S., 2011. *The Social Life of Health Information*. [online] Pew Internet & American Life Project. Available at: <http://pewInternet.org/Reports/2011/Social-Life-of-Health-Info.aspx> [Accessed 18 Sep. 2018].

Gandomi, A. and Haider, M., 2015. Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2015), pp.137-144.

Gaudiano, P., 2017. *The Best Approach to Decision Making Combines Data and Managers Expertise*. [online] Harvard Business Review Digital Articles. Available at: <https://hbr.org/2017/06/the-best-approach-to-decision-making-combines-data-and-managers expertise ? referra l = 03759 & cm _ vc = rr _ item _ page . bottom > [Accessed 15 Nov. 2018].

Grant, R.M., 1991. The Resource-Based Theory of Competitive Advantage: Implications for Strategy Formulation. *California Management Review*, 33(3), pp.114–135.

Grover, V., Chiang, R.H.L., Liang, T.P. and Zhang, D., 2018. Creating Strategic Business Value from Big Data Analytics: A Research Framework. *Journal of Management Information Systems*, 35(2).

Günther, W.A., Rezazade Mehrizi, M.H., Huysman, M. and Feldberg, F., 2017. Debating big data: A literature review on realizing value from big data. *Journal of Strategic Information Systems*, 26(2017), pp.191-209.

Gupta, A., Tyagi, M. and Sharma, D., 2013. *Use of Social Media Marketing in Healthcare*. [online] Journal of Health Management. Available at: <https://www.ru.ac.za/library/> [Accessed 6 May 2018].

Gupta, M. and George, J.F., 2016. Toward the development of a big data analytics capability. *Information and Management*, 53(8), pp.1049–1064.

Hawkins, C.M., DeLaO, A.J. and Hung, C., 2016. Social Media and the Patient Experience. *Journal of the American College of Radiology*, [online] 13(12), pp.1615–1621. Available at: http://dx.doi.org/10.1016/j.jacr.2016.09.006 [Accessed 18 May 2018].

Helfat, C., Finkelstein, S., Mitchell, W., Peteraf, M., Singh, H., Teece, D., Winter, S. and Maritan, C., 2007. Dynamic capabilities and organizational processes. In: *Dynamic Capabilities: Understanding Strategic Change in Organizations*. Oxford: Blackwell Publishing. Ch.3.

Helfat, C.E., 2011. Dynamic Capabilities and Strategic Management: Organizing for Innovation and Growth. By David J. Teece. *R&D Management*, 41(2), pp.217–218.

Helfat, C.E. and Martin, J.A., 2015. *Dynamic Managerial Capabilities: Review and Assessment of Managerial Impact on Strategic Change*. *Journal of Management*, 41(5), pp.1281-1314.

Helfat, C.E. and Peteraf, M.A., 2003. The dynamic resource-based view: Capability lifecycles. *Strategic Management Journal*, 24, pp.997-1010.

Henke, N., Bughin, J., Chui, M., Manyika, J., Saleh, T., Wiseman, B. and Sethupathy, G., 2016. *The age of analytics: Competing in a data-driven world*. Washington DC: McKinsey& Company.

Hunter, R.F., Gough, A., O’Kane, N., McKeown, G., Fitzpatrick, A., Walker, T., McKinley, M., Lee, M. and Kee, F., 2018. Ethical issues in social media research for public health. *American Journal of Public Health*, 108(3), pp.343–348.

Janssen, M., van der Voort, H. and Wahyudi, A., 2017. Factors influencing big data decision-making quality. *Journal of Business Research*, 70(2017), pp.338-345.

Kaplan, A.M. and Haenlein, M., 2010. Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53(1), pp.59–68.

Kapoor, K.K., Tamilmani, K., Rana, N.P., Patil, P., Dwivedi, Y.K. and Nerur, S., 2018. Advances in Social Media Research: Past, Present and Future. *Information Systems Frontiers*, 20(3), pp.531-558.

Kemp, S., 2018. *Digital in 2018: World's Internet users pass the 4 Billion Mark*. [online] We Are Social. Available at: <https://wearesocial.com/blog/2018/01/global-digital-report-2018> [Accessed 6 May 2018].

Kemp, S., 2019. *Digital 2019: Internet Use Accelerates*. [online] We Are Social. Available at: <https://wearesocial.com/blog/2019/01/digital-2019-global-Internet-use-accelerates> [Accessed 9 February 2019].

Kim, G., Shin, B., Kim, K. and Lee, H.G., 2011. IT Capabilities, Process-Oriented Dynamic Capabilities, and Firm Financial Performance. *Journal of the Association for Information Systems*, 12(7), pp.487-517.

Cornejo, C., 2018. *Pharma Social Media Best Practices: Reaching Out Within Regulatory Restrictions*. [online] Available at: <https://www.pulsarplatform.com/resources/social-check-up-the-top-20-pharma-companies-on-social-media/> [Accessed 11 Nov. 2018].

Koscielniak, H. and Puto, A., 2015. BIG DATA in decision making processes of enterprises. *Procedia Computer Science*, 65(2015), pp.1052–1058.

Koumpouros, Y., Toulias, T.L. and Koumpouros, N., 2015. The importance of patient engagement and the use of Social Media marketing in healthcare. *Technology and Health Care*, 23(4), pp.495–507.

Kurtmollaiev, S., Pedersen, P.E., Fjuk, A. and Kvale, K., 2018. Developing managerial dynamic capabilities: A quasi-experimental field study of the effects of design thinking training. *Academy of Management Learning and Education*, 17(2), pp.184-202.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S. and Kruschwitz, N., 2011. Big data, analytics and the path from insights to value.(The New Intelligent Enterprise)(Report). *MIT Sloan Management Review*, 52(2), pp.21-31.

Li, Y., Wang, X., Lin, X. and Hajli, M., 2018. Seeking and sharing health information on social media: A net valence model and cross-cultural comparison. *Technological Forecasting and Social Change*,126, pp.28–40.

Lin, Y. and Wu, L.Y., 2014. Exploring the role of dynamic capabilities in firm performance under the resource-based view framework. *Journal of Business Research*, 67(2014), pp.407-413.

Livinec, M., 2018. *Global Sector Report: Pharmaceuticals*. [online] Euler Hermes Economic Research. Available at: <http://www.eulerhermes.com/economic-research> [Accessed 5 May 2018].

Lockett, A., O'Shea, R.P. and Wright, M., 2008. The development of the resource-based view: Reflections from Birger Wernerfelt 1. *Organization Studies*, 29(8–9), pp.1125–1141.

Longest, K.C., 2015. *Using Stata for Quantitative Analysis*. Second ed. Los Angeles: Sage Publications Inc.

Makadok, R., 2001. Toward a synthesis of the resource-based and dynamic-capability views of rent creation. *Strategic Management Journal*, 22(5), pp387-401.

Mallinger, M., 2015. Big Data Desicion Making: Is There Room for Intuition in the Era of Big Data. *Graziadio Business Review*, 18(2), pp1-6.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Hung Byers, A., 2011. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. [online] Available at:

<https://bigdatawg.nist.gov/pdf/MGI_big_data_full_report.pdf> [Accessed 5 May 2018].

Marr, B., 2015. *Big data : using SMART big data, analytics and metrics to make better decisions and improve performance*. 1st ed. West Sussex: Wiley & Sons.

Mata, B.F.J., Fuerst, W.L. and Barney, J.B., 1995. Information Technology and Sustained Competitive Advantage : A Resource-Based Analysis *MIS Quarterly*, December(1995), pp.487-505.

Mazzei, M.J. and Noble, D., 2017. Big data dreams: A framework for corporate strategy. *Business Horizons*, 60(3), pp.405–414.

McAfee, A., Brynjolfsson, E. and Dearstyne, B.W., 2012. Big Data: The Management Revolution. *Harvard Business Review*.[e-journal] October. Available at: https://hbr.org/2012/10/big-data-the-management-revolution [Accessed 5 May 2018].

Mgudlwa, S. and Iyamu, T., 2018. Integration of social media with healthcare big data for improved service delivery. *South African Journal of Information Management*, 20(1), p.894.

Miller, K.D. and Waller, H.G., 2003. Scenarios, real options and intergrated risk management. *Long Range Planning*, 36, pp.93–107

Mukhopadhyay, S., 2018. Opinion mining in management research: the state of the art and the way forward. *OPSEARCH*, 55, pp.221-250.

Narayan, L., 2017. *7 Social Media Trends in the Pharmaceutical Industry*. [online] Available at: <https://www.klick.com/health/news/blog/social/the-top-7-Pharma-social-media-best-practices-of-2017/> [Accessed 17 May 2018].

Ohly, S., Sonnentag, S., Niessen, C. and Zapf, D., 2010. Diary Studies in Organizational Research: An Introduction and Some Practical Recommendations. *Journal of Personnel Psychology*, 9(2), pp.79–93.

Oussous, A., Benjelloun, F.Z., Ait Lahcen, A. and Belfkih, S., 2017. Big Data technologies: A survey. *Journal of King Saud University - Computer and Information Sciences*, 2017, pp1-17.

Panahi, S., Watson, J. and Partridge, H., 2014. Social media and physicians: Exploring the benefits and challenges. *Health Informatics Journal*, 22(2), pp.99–112.

Pang, B. and Lee, L., 2008. Opinion Mining and Sentiment Analysis. *Foundation and Trends Information Retrieval.* 2(1-2), pp1-135.

Partikas, C., Toms, C., Evans, R., Demuren, E. and Shah, S., 2018. *The Social Check-up March 2018*. [online] Available at: <https://www.pulsarplatform.com/resources/social-check-up-the-top-20-Pharma-companies-on-social-media/> [Accessed 13 May 2018].

Peteraf, M.A., 1993. The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal*, 14(3), pp.179-191.

Pisano, G.P., 2015. A Normative Theory of Dynamic Capabilities: Connecting Strategy, Know-How, and Competition. *HBS Working Knowledge*, [online] Available at: <https://hbswk.hbs.edu/item/a-normative-theory-of-dynamic-capabilities-connecting-strategy-know-how-and-competition > [Accessed 17 June 2018].

Porter, M.E., 1980. *Competitive.Strategy.-Techniques for Analyzing Industries and Competitors: with a New Introduction.* First ed. *Techniques for Analyzing Industries and Competitors: with a New Introduction*. New York: The Free Press.

Raciborski, R., 2009. Graphical representation of multivariate data using Chernoff Faces. *Stata Journal*, 9(3), pp.374–387.

Raghupathi, W. and Raghupathi, V., 2014. Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*, 2(3), pp.1-10.

Ram, J., Zhang, C. and Koronios, A., 2016. The Implications of Big Data Analytics on Business Intelligence: A Qualitative Study in China. In: *Procedia Computer Science*, 87(2016), pp.221-226.

Ray, G., Barney, J.B. and Muhanna, W.A., 2004. *Capabilities, business processes, and competitive advantage: Choosing the dependent variable in empirical tests of the resource-based view*. *Strategic Management Journal*, 25(2004), pp.23-37.

Ray, S., 2017. In sickness and in health : Clinical research and social media. *Perspectives in Clinical Research*, 8(2), pp.73–78.

Ross, J.W., Beath, C.M. and Quaadgras, A., 2013. You may not need big data after all. *Harvard Business Review,* [online] Available at: https://hbr.org/2013/12/you-may-not-need-big-data-after-all [Accessed 5 May 2018].

Schroeder, R., 2014. Big Data and the brave new world of social media research. *Big Data & Society*, [online] 1(2), p.20. Available at: http://bds.sagepub.com/content/1/2/2053951714563194.abstract%5Cnhttp://bds.sagepub.com/lookup/doi/10.1177/2053951714563194 [Accessed 5 May 2018].

Shah, T., Rabhi, F. and Ray, P., 2015. Investigating an ontology-based approach for Big Data analysis of inter-dependent medical and oral health conditions. *Cluster Computing*, 18(1), pp.351-367.

Sivarajah, U., Kamal, M.M., Irani, Z. and Weerakkody, V., 2017. Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70(2017), p.263-286.

Social Mention, 2018. *Social Mention*. [online] Available at: <www.socialmention.com> [Accessed 5 May 2018].

Spitz, M. and Einarsen, B., 2017. *The Top 7 Pharma Social Media Best Practices of 2017*. [online] Available at: <https://www.klick.com/health/news/blog/social/the-top-7-Pharma-social-media-best-practices-of-2017/> [Accessed 18 May 2018].

Statista, 2018. • *Global social media ranking 2018.* [online] Available at: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users [Accessed 11 November 2018]

Statista, 2018. *Global Pharmaceutical Industry - Statistics & Facts*. [online] Available at: <https://www.statista.com/topics/1764/global-Pharmaceutical-industry/> [Accessed 11 Nov. 2018].

Sullivan, T., 2018. *FDA Sends Warning Letter to Duchesnay Inc. Over Social Media Posts From Kim Kardashian*. [online] Policy and Medicine. Available at: <https://www.policymed.com/2015/08/fda-sends-warning-letter-to-duchesnay-inc-over-social-media-posts-from-kim-kardashian.html> [Accessed 11 Nov. 2018].

Syrdal, H.A. and Briggs, E., 2018. Engagement With Social Media Content: a Qualitative Exploration. *Journal of Marketing Theory and Practice*, 26(1–2), pp.4–22.

Teece, D.J., 2010. Business models, business strategy and innovation. *Long Range Planning*, 43(2–3), pp.172–194.

Teece, D.J., 2014. The Foundations of Enterprise Performance: Dynamic and Ordinary Capabilities in an (Economic) Theory of Firms. *Academy of Management Perspectives*, 45(1), pp.8-37.

Teece, D.J., 2018. Business models and dynamic capabilities. *Long Range Planning*, 51(1), pp.40–49.

Teece, D.J., Pisano, G. and Shuen, A., 1997. Dynamic capabilities & strategic management. *Strategic Management Journal*, 18(7), pp.509–533.

Thirathon, Wieder, B., Ossimitz, M. and Matolcsy, 2017. Impact of Big Data Analytics on Decision Making and Performance. *TBA, International Conference on Enterprise Systems, Accounting and Logistics, (*2017-09–01), pp.1–16. Available at:< http://hdl.handle.net/10453/98428> [5 May 2018].

Tutelman, P.R., Dol, J., Tougas, M.E. and Chambers, C.T., 2018. Navigating Your Social Media Presence: Opportunities and Challenges. *Clinical Practice in Pediatric Psychology*. [online] Available at: <http://dx.doi.org/10.1037/cpp0000228> [Accessed 15 May 2018].

Unmetric, 2017. *7 Social Media Trends in the Pharmaceutical Industry*. [online] Web Page. Available at: <https://unmetric.com/Pharma-social-media-trends-report>

[Accessed 11 Nov. 2018].

de Vos, A., Strydom, H., Fouché, C. and Delport, C., 2011. *Research At Grass Roots*. 3rd ed. Pretoria:Van Schaik Publishers.

Wang, L., Wang, G. and Alexander, C.A., 2015. Big Data and Visualization: Methods, Challenges and Technology Progress. *Digital Technologies*.

Wang, Y. and Hajli, N., 2017. Exploring the path to big data analytics success in healthcare. *Journal of Business Research*, 70, pp.287–299.

Wernerfelt, B., 1984. A resource-based view of the firm. *Strategic Management Journal*, 5(2), pp.171–180.

Wernerfelt, B., 1989. From critical resources to corporate strategy. *Journal of General Management*, 14, pp.4–12.

Wernerfelt, B., 2013. *Small forces and large firms: Foundations of the RBV*. *Strategic Management Journal*, 34(6), pp.635-643.

Yang, D., Huang, C. and Wang, M., 2017. A social recommender system by combining social network and sentiment similarity: A case study of healthcare. *Journal of Information Science*, 43(5), pp.635–648.

Yang, Y.-F., 2008. The roles of human resources, information technology, and marketing knowledge capabilities in performance: an extension of the resource-based theory perspective. *Social Behavior and Personality: an international journal*, 36(9), pp.1269–1282.

Yi, X., Liu, F., Liu, J. and Jin, H., 2014. Building a network highway for Big Data: Architechure and Challenges. *IEEE Network*, 28(4), pp.5–13.

Zhao, J.L., Fan, S. and Hu, D., 2014. Business challenges and research directions of management analytics in the big data era. *Journal of Management Analytics*, 1(3), pp.169-174.