Who talks about diabetes on Twitter, what do they say, and why are there so many jokes?

Amy K McLennan[1], Stanley J Ulijaszek[1] & Mariano Beguerisse-Díaz[2,3]

[1] School of Anthropology and Museum Ethnography, University of Oxford, UK.
[2] Department of Mathematics, Imperial College London, UK.
[3] Mathematical Institute, University of Oxford, UK.

Introduction

Governments across the world are increasingly using social media platforms like Twitter to disseminate health information and advice. Growing numbers of health departments and organisations have social media policies, and social media are now used by many as a low-cost tool for addressing so-called 'lifestyle diseases' such as obesity and Type 2 diabetes. The effectiveness of these initiatives is generally measured in terms of the number of subscribers following social media accounts (for example, see Public Health England 2014). Government social media policies tend to focus on legal concerns such as regulating staff use and ensuring privacy protection, rather than citizen health outcomes or experiences (Fast et al. 2015). There is little acknowledgement that, for citizens, Twitter contains a multitude of messages, with public health messages alongside marketing for unhealthy products (Kelly et al. 2015), or that it may be used for a wide range of reasons, including information seeking or dissemination (Scanfeld et al. 2010), a source of emotional support and community acceptance (Pew Research Center 2013), and even as a place of activism (Beguerisse-Díaz et al. 2014; González-Bailón et al. 2011). Instead, government policies emphasise the importance of self-responsibility (Fast et al. 2015), stating that it is citizens' responsibility to navigate the broader social media landscape for themselves.

Content from social media that appears to be 'unrelated' to health advice does not tend to be considered by health researchers either. Analysis carried out for studies relating to Twitter and health tends to filter out content such as irreverent chatter or jokes that the researchers view as unrelated 'noise' and irrelevant to health research (for example, Harris et al. 2013; Hawn 2009; Paul & Dredze 2011). However, there is ample evidence that marketing (Kelly et al. 2015), social values (McLennan & Ulijaszek 2015), emotional connection, community (Ferzacca 2004) and humour (McCreaddie & Wiggins 2008) can all contribute to lifestyle patterns and health outcomes. Laughter has even been shown to inhibit blood glucose level rises in patients with diabetes (Hayashi et al. 2003).

Unlike typical policy approaches or public health studies, our approach to understanding the significance of Twitter in public health is not limited to formal health messaging or content that aligns with broader public health aims. Instead, we take a large collection of tweets containing the term 'diabetes', and employ techniques from network science and information retrieval to determine who are the most influential Twitter users, what are the most common messages, and what is the content that attracts the most attention. We identify patterns that emerge, and interrogate their significance in the public health context.

We begin this chapter by introducing Twitter and explaining why it, and content on it relating to diabetes, is our focus for our research. We then ask three questions. First, *who generates the content (messages) that has the most impact?* To answer this question we examine the structure of the time-

changing network of retweets, and introduce the notion of hub and authority scores to describe different types of influence and impact. Second, *what is the most common content?* Here, we examine the content generated by the top authority accounts, and explore the type of messages they disseminate. Finally, we ask *why does so much of the content involve humour?* We look more closely at the content that appears more persistently, and explore what it tells us about health information seeking, community building and activism.

Twitter is more than a news platform for people with diabetes

Twitter is a social media platform with approximately 320 million monthly active users around the world (Desilver 2016). Users can post short public messages or *tweets* up to 140 characters in length. In addition, users can subscribe (*follow*) to receive the tweets of other users. When users log in, they can see in their *timeline* the tweets that have been posted by all the people they follow, ordered chronologically with the most recent tweets on top. According to Twitter, the 'microblogging' platform is 'like being delivered a newspaper whose headlines you'll always find interesting' (Twitter Inc. 2016). Many users do use Twitter exclusively to read messages: over 25% of users have never posted a tweet.

Twitter emphasises its importance as a global, real-time public information source; providing information about health management and diet is not an exception. As diabetes is a growing global phenomenon (Figure 1) that is linked strongly to so-called 'individual choice', particularly in diet, it is no surprise that Twitter and other social media have been taken up as platforms through which to seek, share and disseminate advice about diabetes.

Diabetes is a clinical condition associated with blood sugar regulation by the hormone insulin and other endocrine factors. Type two diabetes (T2D), in which the pancreas produces insufficient insulin, accounts for around 90% of cases worldwide. Type one diabetes (T1D) accounts for the remaining 10% of cases, in which the pancreas produces no insulin at all. Insulin regulates blood sugar levels in a person's blood; without insulin, these levels can fluctuate significantly. If a person's blood sugar becomes too low, their brain will cease to function. Over a person's lifetime, fluctuating and/or consistently high blood sugar levels can cause problems with the heart, blood vessels, eyes, kidneys and nerves. People who have diabetes must constantly monitor and regulate their blood sugar levels; administering insulin or being active can reduce blood sugars, while consuming most types of food and drink can raise blood sugars.

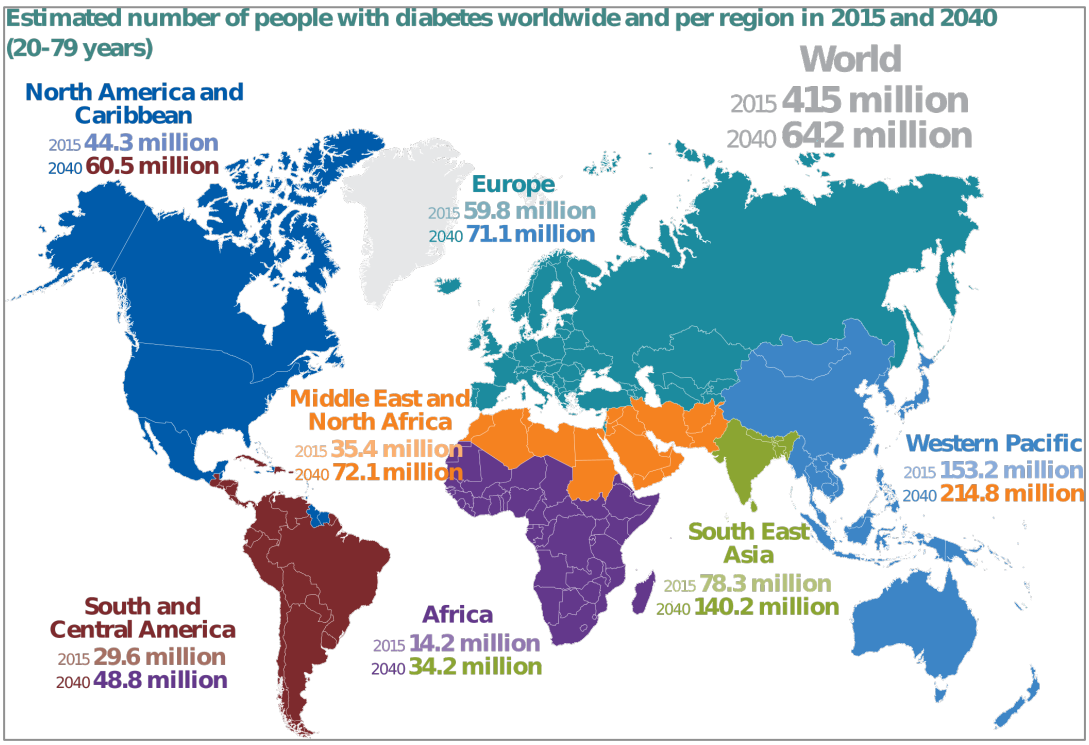

Figure 1: Estimated number of people with diabetes worldwide and per region in 2015 and 2040 (20-79 years). Reproduced with permission from (International Diabetes Federation 2015).

Yet in practice, Twitter, like other social media platforms, is more than a news source that facilitates a unidirectional flow of news information. It is also a venue for engaging in social interaction, seeking or giving emotional support, stigmatising or ostracising others, and organising collective action (Kwak et al. 2010).

Americans living with chronic diseases use the internet more for communication than for seeking health information (Fox & Purcell 2010). While they are less likely than those without chronic disease to have access to the internet or to seek health information online, once they are online, people living with chronic disease are more likely than others to access user-generated health content such as blog posts, hospital reviews, doctor reviews, and podcasts (Fox & Purcell 2010). If people seek emotional support when dealing with a health issue, or a quick remedy for an everyday health issue, they more commonly seek information from fellow patients, friends and family than from professional sources (Pew Research Center 2013). Caregivers are also more likely than non-caregivers to seek health information and support online (72% versus 50%) (Pew Research Center 2013).

Social media such as Twitter allow users to build community through their interactions. Twitter users can interact with each other in a variety of ways, including *liking* (expressing approval of a tweet), *replying* to a tweet, *mentioning* a specific user in a tweet, and *retweeting* (forwarding a tweet posted by someone else to one's own followers). Through retweeting, the flow of messages can create local or global concentrations of a particular type of information at a particular moment in time. For example, disclosures of diabetes by celebrities such as film actor Tom Hanks spread rapidly and widely on Twitter in a short amount of time (Beguerisse-Díaz et al. 2015). Twitter users generated a wealth of messages about this news; some messages were supportive while others were critical or sarcastic, some were gossip or jokes, while others were factual (such as news headlines). The

structure of these interactions naturally lends itself to being investigated using quantitative and computational methods, such as network science.

Twitter has also been used to organise collective action, from diabetes fundraising and awareness events to massive protests and rallies. The immediacy and ease of messaging, the large number of users around the world, and the need for community consent and endorsement for the platform to succeed as an advocacy tool, have led to Twitter and other social media platforms enable what is known as 'organisation without organisations' (Shirky 2008). This new form of collective organisation permits individualised collective action, and social connectivity and consensus-formation around personal lifestyle values, without group loyalties that characterised pre-digital social movements (Bennett 2012).

Methods

This chapter is the result of an iterative process of collaborative research between mathematicians and anthropologists. The analysis is based on over 2.5 million English-language tweets that contain the term 'diabetes' posted between March 26 2013 and January 19 2014. The tweets were collected by Sinnia, a data analytics company.[1] Along with the text of the tweets, we collected the following information about the users who produced them:

- Followers: Twitter users that are subscribed to the user's tweets.
- Friends: Users whom the user has followed.
- Retweets: Tweets composed by other users that the user of interest has passed along to his/her followers.
- Biography: A user's self-description, where Twitter allows users to describe themselves in up to 140 characters.

A *network* (or graph) is an information structure (more precisely, a mathematical object) in which pairs of *nodes* (or vertices) can be connected to each other by *edges* (or arcs) (Newman 2010). From Twitter, one can obtain several networks of in which the users are nodes by creating, for example, connections that represent friend/follower relationships, or the event in which one user has retweeted someone else's tweet. In these cases, the edges are *directional,* that is, they distinguish the user who follows (retweets) from the user who is being followed (retweeted). The information (i.e., the content) flows in the opposite direction to the declared direction of interest. In other words, if a user 'follows' another user, the interest goes from the source of the connection to the target, and the target's tweets are received by the source. Likewise, in a retweet, the person who retweets is expressing interest in a particular message, which then is transmitted to the retweeter's followers.

We used methods from network science (analysis of centralities in temporal networks, community detection) and information retrieval (topic detection) to identify the main patterns in the content of tweets, and the interactions among the users. Importantly, only 10% of user accounts in our data produced tweets that elicit any form of response (a retweet or reply). Furthermore, the intensity of the response was extremely unevenly distributed: relatively few users attained a disproportionally high amount of attention. The technical details of our work and ethical considerations are elaborated in detail in Beguerisse-Díaz et al. 2015.

[1] Sinnia is a data analytics company operating from Mexico. More information can be found at their website: www.sinnia.com

Whose content has the largest impact?

Anthropologists, among others, have begun to explore how the rise of digital and interactive media reflects, or can change, the landscape of power and authority. Some have argued that existing power structures are destabilised and democratised as information becomes more widely available (Lien & Nerlich 2004; Blue 2010). In this case, attribution of authority is broadened, with knowledge production becoming more participatory and widely distributed. However, others have found that a specific curatorial process is required to achieve such distributed power, and that the use of participatory platforms in the absence of careful guidelines and principles can limit their capacity to achieve the imagined ideals of open access and information democratisation (Geismar 2012). In this case, digital media can just be another channel through which existing power structures are reinforced.

Quantitative approaches to investigating Twitter offer a different perspective on influence, power and authority. Using techniques from network science, researchers can find which users are more 'central' in a network of interactions. There are many different notions of centrality that range from the simple (number of connections) to the sophisticated (using the properties of random walks on the network) (Newman 2010). The 'hub' and 'authority' score of a user (or node) are two examples of node centrality that are defined recursively (Kleinberg 1999):

- A good authority is a node that receives many connections from good hubs; i.e. authorities are users who produce tweets that attract the most attention (in the form of retweets).
- A good hub is one that connects to good authorities; i.e. passes information from authorities to other users.

For example, in the worldwide web (in which web pages are nodes and hyperlinks are directed connections between them) and example of a hub would be Google, which contains links to other sites but has no content itself. An authority would be a site such as Wikipedia, which contains content that many web users seek.

We have identified the users with the highest hub an authority scores in a retweet network from our diabetes Twitter data (Beguerisse-Díaz et al. 2015).

Authorities send the messages with the biggest impact

The top ten authorities in our data set are a mixture of bloggers, advocacy groups, companies and a health information firm (Figure 2). Four of the top ten authorities are directly linked to T1D. The onset of T1D is typically much earlier than T2D's and tends to affect people who have it much more severely, as it is related to inability of the pancreas to produce insulin rather than a reduction in the pancreas' capacity to produce insulin. The top authorities tend to have a relatively sustained presence over the data observation period.

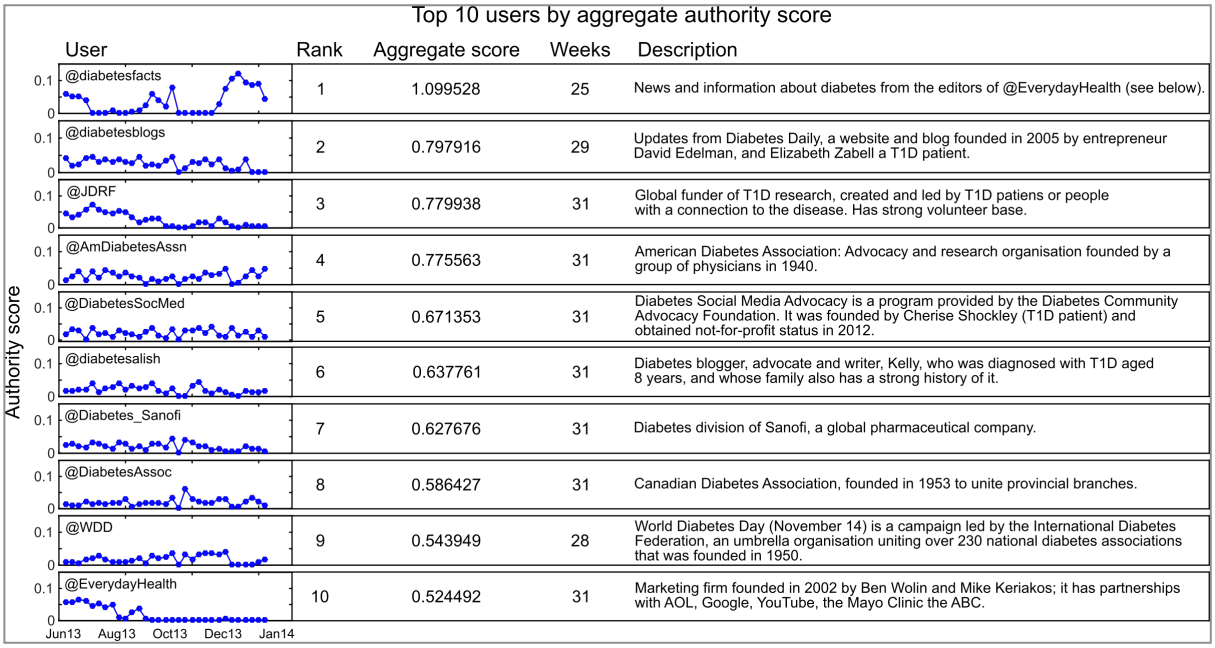

Figure 2: Top ten users by aggregate authority score, number of weeks with non-zero authority score, and brief description (Beguerisse-Díaz et al. 2015)**.**

We then create the follower network of the top 1,000 authorities in our data (who follows whom from within this group). These users can be divided into six distinct communities (Figure 3):

C0. Health and medicine related accounts.
C1. A diabetes-related group of advocates, patients and families.
C2. Accounts related to lifestyle and wellbeing.
C3. Accounts related to news and media.
C4. Celebrities.
C5. A group of accounts specifically related to retailer Tesco.
C6. Humour and parody accounts.

A community in this context is a group of nodes that contain many more connections within the group than with the rest of the network (Beguerisse-Díaz et al. 2014; Porter et al. 2009). The word clouds in Figure 3 are made from the biographies of the members of each community show a remarkable consistency in the vocabulary used; the exception is the community of humoristic and parody accounts whose members do not the same vocabulary to describe themselves. Despite Tesco's location in the UK, it has large presence on the global Twitter platform (in English). Aside from Tesco, food industry representatives and lobbies (which are heavily active in debates about food policy and legislation) are notably absent from Twitter debates about diabetes.

Together, authority nodes represent a variety of advocacy positions – health advocates tweet about different lifestyle choices, the wellbeing group promotes new diets and fads, and the pharmaceutical industry advocates for pharmaceutical intervention rather than dietary change. However, not all users give a clear signal of what underpins their position; Tesco's interest in diabetes, for instance, promotes its brand and generates opportunities for it to also promote its products, such as health insurance and food products. In practice, there is an unclear distinction between marketing, sponsorship and advocacy on Twitter. This lack of clarity is a characteristic of many electronic media: anyone can establish a presence, and there is little scrutiny of their objectives and their effects on the

broader population. Users may be expected to be discerning and responsible, but it is unclear on what they should base their judgement, given that many user profiles appear equally credible and the information equally authoritative.

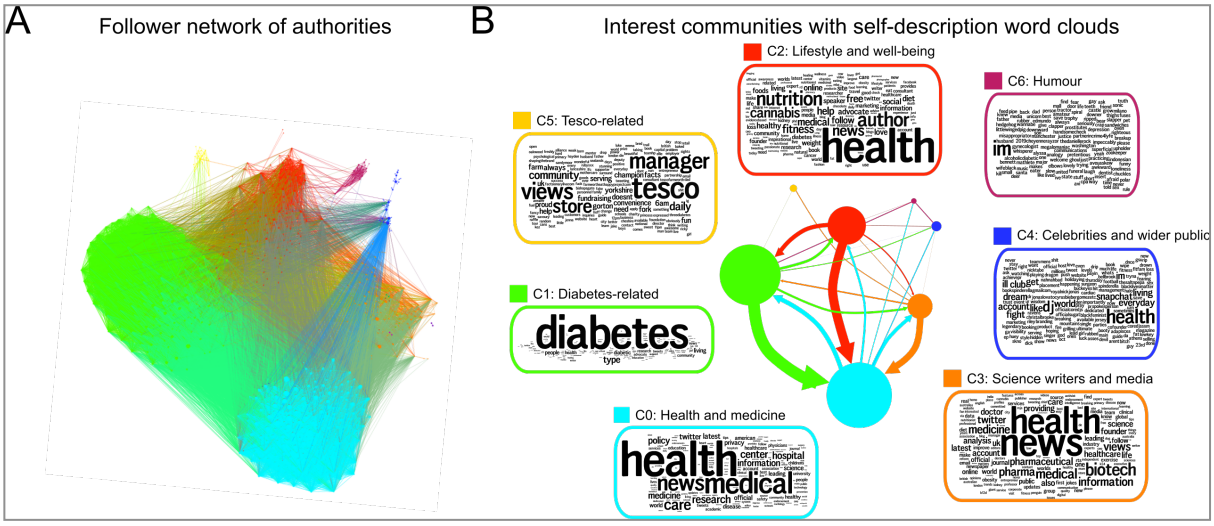

Figure 3: A: The follower network of the top authority account. Nodes are coloured according to the community to which they belong. B: The follower network coarse-grained by communities. The word clouds contain the words that most frequently appear in the members' self-descriptions. The greatest number of users overwhelmingly pays attention to the health and medical advice community (C0), some pay attention to the diabetes community (C1, which includes funding agencies and patients) (Beguerisse-Díaz et al. 2015)**.**

Hubs connect users to the most important messages

The top hub accounts in our data set change from week to week, and tend to be a mixture of bloggers, automated accounts, users with no specific or declared interest in health, and accounts which have since been closed. Hubs, unlike authorities, do not have a sustained high presence over time (Figure 4). This means that there is no account that is routinely and consistently linking users with sources of information. Instead, hubs tend to have a flash of brilliance and then dissipate. Our data do not allow any inference about why this is the case.

The top hubs in our data set are predominantly bloggers and users who have experienced diabetes. It makes sense that they point to many authoritative sources of information, as this is their declared interest both on Twitter and on the blogs they administrate. However, the intention of the top hub – *@1Medical2News* – is less clear. It appears to be a medical doctor named Dr Richard Billiard, and the account appears credible at first glance; however, this doctor has no other online presence, and retweets a large amount of messages per day at regular intervals (an average of 50 tweets per day since 2013), and has never produced an original tweet. It is unclear who might be behind the profile or what they are attempting to achieve.

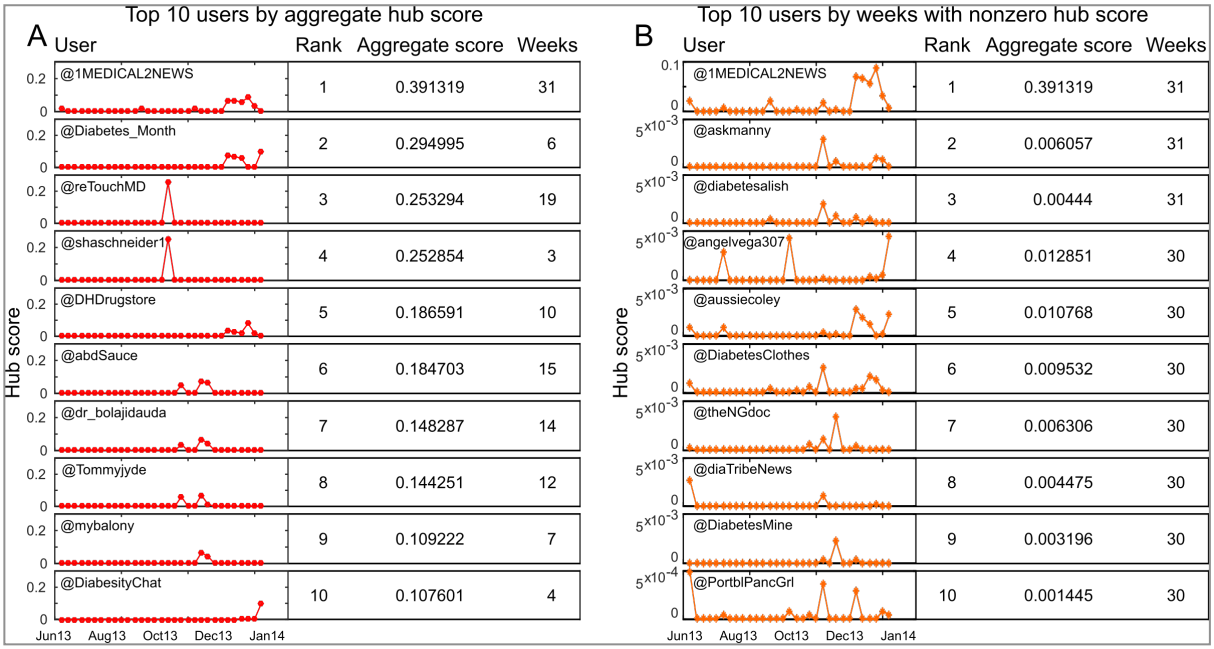

Figure 4: A: Top ten hub accounts by aggregate hub score score. B: Top hub accounts by number of weeks with non-zero authority score (Beguerisse-Díaz et al. 2015)**.**

It is difficult to be a top authority and hub simultaneously

Of the two percent of the most central users in our data set (by hub and authority score) – users who are frequently retweeted by other users – none are at the very top of the authority and hub ranking simultaneously. Authorities tend to push out information of their own, but they retweet messages by other authoritative sources of information far less frequently. Hubs tend to retweet authorities' messages but seldom produce tweets with wide impact themselves.

There are at least three reasons why this might occur. First, authorities may all be advocating slightly different things: with no clear overlap of agendas, one organisation may not opt to pass to its followers with messages from a different source. Second, it is relatively more time-consuming and risky to read, select and retweet information from other sources. This may make it an unsustainable activity, especially for established authorities that must check and verify content before retweeting it. Third, it may suggest that organisations in the public health landscape are seen as atomised units rather than part of a broader network of advocates.

A hub's role of bridging content between users or distributing information can be an important one, but it does not appear to be a sustained activity by any one user on Twitter in relation to diabetes. It may be worth considering whether public health information could be improved if large organisations focused on not only pushing out their own information, but also pointing at relevant information from other sources or engaging more widely with other Twitter users.

Overall, the users who are most influential on Twitter when it comes to diabetes are a diverse group. This challenges the notion that the domain of health is a discrete category: social media content relating to diabetes on Twitter is connected with a diverse range of sectors, organisations, interests and perspectives. Biomedicine and public health must look beyond their boundaries to identify important influencers of people.

What is the most common content?

In addition to knowing who are the most influential users on Twitter, it is also useful to understand the content of the messages being posted. Content in tweets can be aggregated into topics using methods from information retrieval and natural language processing, and then qualitatively analysed to interrogate and explain themes and patterns (Beguerisse-Díaz et al. 2015). The tweets in our data set fall into four broad thematic groups: health information, news, social interaction, and commercial. Further, there is then one anomalous cluster of recurrent tweets that is discussed further in the following section. The word clouds of the top 200 word roots (for example 'obe' stands for 'obese', 'obesity', and so on) in each theme are shown in Figure 5.

Content can be split into four main thematic groups

One of the largest thematic groups consists of health information, research findings, recommendations, advice and warnings, which are all abundantly tweeted and retweeted. The top 10 terms in this thematic group are: *risk, type2, disease, heart, research, month, obesity, fruit, news* and *aware*. Tweets in this group include:

- Public health messages.
- Links to articles, blogs and studies about risks, treatment and cure of diabetes.
- Population health fears.
- Publicity about outreach and awareness events and activities.
- Advice about diabetes management and diagnosis.
- Lifestyle, diet and cookery tips, news and links.
- Life stories and experiences (some for marketing purposes).
- Dangers of sugar, sugar replacements and/or soda.

The advice in these tweets generally appears authoritative in tone and language, making it difficult to distinguish less-credible advice from more credible advice. In general, there is a high turnover in the content that each user is exposed to, even though many messages (for example, those from newspapers and online media) are posted multiple times. Put another way, a 'hot topic' in one week will not necessarily appear in subsequent weeks.

The second group contains news-related content. News tweets in the dataset list a headline of a news article and sometimes the first line of the story, and often provide a link to the complete story. The top 10 terms in this group are: *type2, risk, fruit, type1, eat, people, blueberry, cut, research* and *juice.* Some news-related tweets communicate research breakthrough studies or technologies, which may be reported with messages of hope for those who have diabetes, in particular T2D. Tweets in this group include:

- Headline links to particular 'breakthrough' studies or technologies.
- Celebrity news.
- General news articles about diabetic people or pets.
- News relating to the pharmaceutical industry and the economy.

The third theme is social interaction. These tweets use language differently other thematic groups: they are typically informal in tone, their attention to spelling and grammar is limited, and they often

use exclamation marks to express fun, laughter, exasperation, and abuse. The top 10 terms in this group are: *give, health, food, die, think, fat, year, diet, disease* and *cause*. Tweets typically include:

- Users that joke about how what they have eaten is likely to give them diabetes.
- Chatter and everyday social interchanges that include mentions of diabetes.
- Everyday experiences of diabetes.
- Stigmatising comments.
- Banter, sexual innuendo and humour relating to sweetness and diabetes.

These tweets indicate a baseline level of awareness of dietary guidelines and diabetes aetiology. Users have conversations and interact about a diversity of topics in a chatter that is not necessarily directly related to diabetes but may include references to it. People who have diabetes – particularly T1D – also talk about the daily experiences of their bodies, sugar management, and social acceptance or stigma; such tweets may elicit retweets or messages of support from other users. Some users also talk in terms of a division between 'us' (people with diabetes, especially T1D) and 'them' (people without diabetes). For example, a user talks about T1D as being a feature he/she looks for in a romantic partner:

I haven't stopped thinking about this girl for seriously like...a month. AND she has diabetes! #diabetesperks.

Such content often receives retweets and replies, including messages of support, or appreciation of a joke.

On the other hand, stigmatising comments, especially tweets that blame diabetic people for bringing the disease on themselves through, for example, poor diet or lack of physical activity, are abundant in the dataset. Faced with such messages, users with T1D diabetes point out that it is important to differentiate between T1D and T2D, insinuating that while T1D diabetes is not a person's 'fault', T2D may be. Other tweets include calling other people 'diabetic' as an insult and wishing diabetes upon a person a user does not like.

A distinct theme in this category consists of tweets with sexual innuendo. At their mildest, such tweets refer to boy-band members or other (e.g., celebrity) infatuations, where the person is said to be 'so sweet' they are diabetes-inducing. At their most extreme, such tweets joke that others' bodily fluids and genitals are so sweet they are diabetes-inducing, and these tweets contain links to pornography websites or other explicit material. Like the jokes discussed earlier, these tweets reflect a baseline awareness of the links between sugar and diabetes amongst people who do not appear to have diabetes themselves.

Finally, there are a number of tweets that are commercial in nature. The top 10 terms in this group are: *type2, drug, job, manage, care, health, marijuana, sale, test,* and *for sale*. Common tweets include:

- Advertisements for jobs in the pharmaceutical and care industries.
- Marketing for a specific product, app, treatment, event or service.
- Pharmaceutical, health industry and stockmarket updates and FDA approvals.
- Sales of diabetes drugs, diets or treatment products online.

Diabetes treatment is a lucrative industry because diabetes is a chronic condition that requires regular and life-long treatment (rather than cure), and so the demand for pharmaceutical products and lifestyle aids is inelastic (Simonsen et al. 2015). People with diabetes depend on different technologies, consumables, health services, and pharmaceutical products. Furthermore, the number of people with T2D is projected to increase dramatically in the future as a result of population ageing and obesity

(International Diabetes Federation 2015), which will further expand the market. This commercial dimension of diabetes is reflected in many Twitter messages.

Who contributes what content?

Overall, tweets are posted by users with different claims to expertise: individuals who have first-hand experience of diabetes; personal trainers advertising their services; companies selling lifestyle products or services; other users with an apparent interest in diabetes, cookery and 'healthy' eating; marketing agencies trying to sell a particular food, supplement or device; or hospitals and health agencies attempting to communicate a specific health message. Home remedies and 'miracle cures' appear alongside health tips and recommendations. Other health-related messages include publicity about outreach and awareness events, activities and information.

The topics in which the highest numbers of top ten authorities converge are related to advocacy and awareness. For example, a topic about Diabetes Blog Week in May 2013 gathered 6 of the top ten authorities: *@diabetesalish, @diabetesblogs, @DiabetesSocMed, @Diabetes_Sanofi, @diabetesfacts,* and *@EverydayHealth*. In other weeks, the top ten authorities appear together in topics related to promotion of blogs by diabetics (using the hashtag *#dblogs*, which appears in 15,901 tweets in the data set), and diabetes social media awareness (using the hashtag *#dsma*, which is promoted by @DiabetesSocMed and appears in 10,945 tweets).

All of the top ten authorities post messages relating to health information frequently and consistently. Some also feature news-related tweets, although these are less common. Two accounts, *@Diabetes_Sanofi* and *@diabetesblogs*, post a broad range of health information tweets. Two other accounts, *@WDD* and *@AmDiabetesAssn* primarily contribute tweets related to outreach and advocacy activities, events and news. The not-for-profit organisation and research funding body *@JDRF* produces tweets that contain life stories and experiences of diabetes sufferers more than any other top-ten authority. Importantly this is not interactive or interpersonal in any way, but appears to be just a different framing of news and information.

Two accounts, *@diabetesfacts* and *@EverydayHealth* (owned by the same company) focus predominantly on lifestyle and diet-related tips, hints and advice. Unlike the other authorities, these do not produce outreach or advocacy messages at all. The vast majority of the tweets from these two accounts provide a link back to the company's website, which offers articles containing health and lifestyle advice.

The messages posted by the accounts *@diabetesalish* and *@DiabetesSocMed* are dominated by a mix of social interactions, banter and advocacy. They post content relating to health information and news headlines, but to a lesser degree than the other top-ten authorities. Their tone is different to the others: it is informal and conversational rather than authoritative or informational. For example:

@DiabetesSocMed: Happy Mother's (aunts, fur baby moms, god moms, etc.) to all the women in the diabetes community! Have a great day!

or

Banging My Head Over Hubby's Clueless Doc http://t.co/2EfKqA2qZZ #diabetes #dblog.

Two other users, *@diabetesblogs* and *@DiabetesAssoc*, also tweet some social and interpersonal messages.

Two accounts, *@diabetesblogs* and *@diabetesalish* occasionally feature marketing or product promotion messages; this is to be expected, as some bloggers generate income by advertising goods and services. Furthermore, while marketing might not be made explicit to other users, it can still be the case that other accounts are also practising marketing, especially accounts owned by stockmarket-listed firms. This may be related to regulatory practices in many countries: while company advertising and marketing are often regulated by state authorities, company sponsorship of bloggers is not regulated, and nor is the content that bloggers post.

Figure 5: Word clouds of the five thematic groups. The clouds on the left column are formed by the most frequently used terms in the thematic group (larger words appear more frequently). The clouds on the right column are formed by the names of the users whose tweets appear more frequently in the thematic group (users who appear more in the group appear larger) (Beguerisse-Díaz et al. 2015).

Why does so much of the content involve humour?

The intensity of collective activity (e.g., number of tweets, book sales, internet searches) can follow a pattern of spikes of interest followed by attenuation, typically driven by external (exogenous) events, or internal activity (endogenous) (Sornette et al. 2004). For example, tweets about Tom Hanks in our dataset began to appear rapidly after Hanks revealed he had T2D on a late-night talk show on 8 October 2013, but interest subsequently waned (Figure 6, green line). However, several tweets in our data set have an activity profile that is strikingly distinct from what we would expect to see; they have a high, sustained occurrence rate over a long period of time. For example, a joke about mathematics and diabetes appears consistently (Figure 6, red line). The top 10 terms in this group of recurrent tweets are: *sugar, eat, blood, sweet, risk, type2, drink, high, reduce,* and *health*.

Humorous tweets, jokes and memes generate substantial and sustained interest over time, something that other types of tweets seldom achieve. They also maintain similar phrasing across the duration of the data set. Examples of these tweets include:

- Jokes about the relative healthiness or unhealthiness of a particular food or activity, relative to widely published public health standards.
- Lyrics from two specific rap songs (one making a joke about sex, sugar and diabetes; the other an inspirational song by a rap artist with T1D).
- Viral 'fun facts' or trivia such as tasting urine as a test for diabetes, or moderate consumption of alcohol being linked to reduced diabetes.

One of the most prominent instances of recurrent content in our data corresponds to various versions of a mathematics joke:

Math Problems: If Jim has 50 chocolate bars, and eats 45, what does he have? Diabetes. Jim has diabetes...

This joke appears consistently in our dataset, more than any other specific tweet (44,130 times including retweets). Other common jokes are exclamations that a user's latest meal or snack (typically food products, from soda to cookies to ice cream) was tasty but will likely cause them diabetes:

2 bowls of yogurt, a bowl of oreos, hersheys, chips, cheese and a shitload of mints. My diet consists of diabetes.

Bother! Burger King has arrived. Hello obesity, diabetes, poor nutrition. McD's is bad enough. Grumble Grumble????

The coca cola Christmas advert, because nothing says Christmas quite like diabetes and capitalism. LOL

It would be easy disregard these tweets as 'noise' distracting from more important messages; indeed, one anonymous reviewer of our initial manuscript reporting this finding asked why we were talking about such material. However, we think there are several reasons why it is important to examine these tweets further. First, the striking consistency and volume of these messages, especially compared to other tweets in the data set, is too great to ignore. Second, just because these tweets contain content that does not fit into the narrow biomedical definition of health and nutrition, does not mean the content is unrelated to it. Indeed, these tweets represent how a large proportion of Twitter users engage in conversations about diabetes. For this reason, we look further at humour in the following pages.

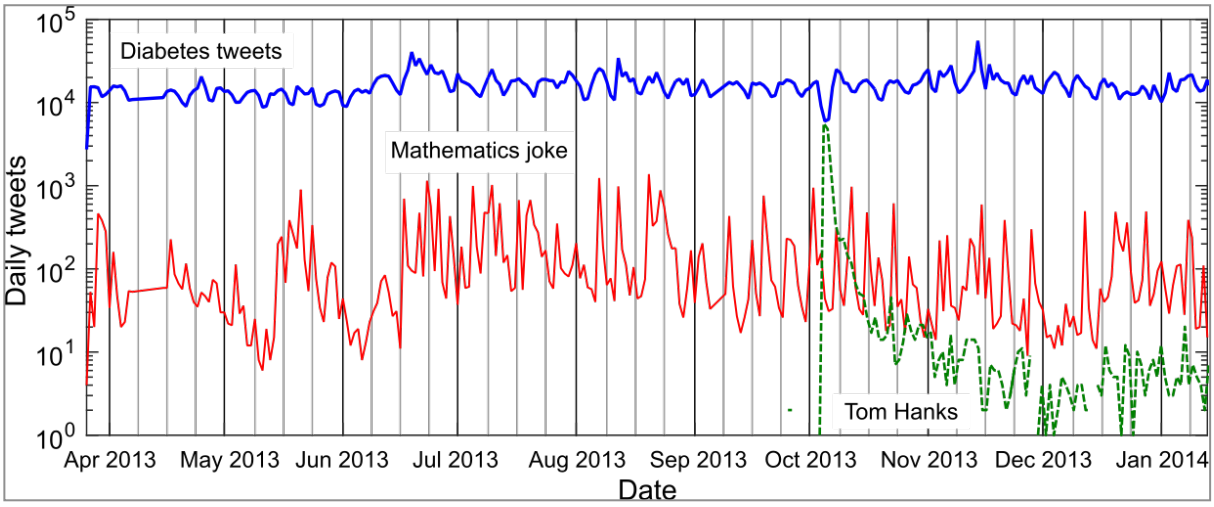

Figure 6: The number of daily tweets in English containing the term 'diabetes' (blue line), the number of tweets containing some version of the 'mathematics joke' (red line, appears in a total of 44,130 tweets in our data set), and the number of tweets mentioning the actor Tom Hanks, who disclosed that he had diabetes in October 2013 (green line, appears in a total of 13,454 tweets in our data set) (Beguerisse-Díaz et al. 2015)**.**

Humour is not a well-studied topic in anthropology and the social sciences (Wasilewska 2013). Some anthropologists have written about it, predominantly because jokes or humour inevitably arise in the course of fieldwork and so ethnographers detour onto talking about them briefly (Carty & Musharbash 2008). Sustained enquiry on the topic has been much less common.

The investigations that have been carried out suggest that humour largely appears to relate to social inclusion and exclusion, where laughing *with* some people inevitably leads to others being excluded from the joke or even being laughed *at*. Digital platforms permit users to create their own categories, and jokes can illuminate what those categories are.

A number of other common themes arise (Carty & Musharbash 2008). In particular:

- Laughter and humour are often used to express unease about discrimination, domination and power imbalances.
- Laughter often appears to have a role in mediating social rupture.
- Laughter and humour often only make sense in a particular context (time, place and/or social context).

These anthropological insights highlight just how significant the relatively high and enduring popularity of diabetes-related jokes on Twitter might be for understanding the success – or failure – of health messaging.

Joking on Twitter is set in a global context, at a time where there is good public awareness about rising diabetes levels, and the links between diet and diabetes. Jokes are also set in a context of rising neoliberal forms of governance that emphasise free markets, consumerism and self-responsibility. Individuals are expected to navigate markets responsibly and to avoid doing things that might harm them, and at the same time to be consumers of products and services. We have discussed this consumer-citizen tension elsewhere (Ulijaszek & Mclennan 2016).

The majority of diabetes jokes on Twitter point to a sense of powerlessness among users in this context. The chocolate bar joke, for example, as with many of the jokes about what people have eaten

that is likely to give them diabetes, makes a subtle protest, a mockery of official dietary advice. Excluded by these jokes are health agencies and their scientific advice. Included are everyday people, who demonstrate that they understand the scientific advice but also recognise the cultural dominance of the foods they invoke in their everyday lives (soft drinks, cookies, ice cream, chocolate, junk food and so on). Scientific advice is quickly shown to be almost ridiculous when placed in the context of a world in which food products entice us at every turn, as these jokes do.

Some jokes are specific to diabetic patients: humour helps to build camaraderie between diabetes sufferers, who bond by joking about what they have to go through on a daily basis to survive. Irony and sarcasm, in particular, are used in this case as people who have diabetes build a community through shared experience.

There were also some jokes stigmatising people with diabetes. Stigmatisation of obesity is well-documented (Brewis 2014), and although it was not as common as jokes about food and diabetes in general, similar stigmatisation of people with diabetes did occur in our data set. When it did occur, people with T1D were quick to highlight that theirs was not a form of diabetes caused by being irresponsible consumer-citizens.

Conclusion

Our investigation of diabetes on Twitter clearly illustrates that this social media platform is much more than a news source. It is a site for social interaction and support. It is a site where collective action is organised and coordinated, especially around raising awareness about a particular cause or event. It is also a site where social power relations and hierarchies are evident.

Digital platforms are said to make disruption of the social order possible, and to democratise knowledge and power. For example, digital platforms arguably permit users to choose where to direct their attention, and to collectively give more authority to some voices over others. However, several findings in our analysis suggest that this may not be always the case.

In the digital environment, it is difficult to tell where the real powers lie. Health agencies often invoke number of followers as an indicator of influence, but this does not take into account the impact of content posted. Calculating authority and hub scores is a more sophisticated approach. The top 1,000 authority users extracted from our analysis are similar to many of the authorities that we would expect to see based on observations in the non-virtual world, including health authorities, lifestyle coaches, pharmaceutical firms and celebrity chefs (Beguerisse-Díaz et al. 2015). However, there is a notable exception: aside from retailer Tesco, the food industry is poorly represented. The absence of the food industry may be related to our use of hub and authority scores for this analysis. Both of these rely on a response from the wider public (i.e., retweets) in order to achieve a high authority score. The absence of the food industry in this case could be interpreted as users choosing to direct their attention elsewhere, which supports the claims that digital platforms like Twitter democratise knowledge and power.

However, this interpretation assumes that influence is primarily exerted in direct and obvious ways. Looking more closely, the food industry is present in our data in other ways. Some advertising appears in the blogger profiles that we investigate, and some authorities receive funding and sponsorship from companies (for example, Tesco sponsors Diabetes UK (Tesco PLC 2014) and soda companies sponsor health organisations (Aaron & Siegel 2016)). Sponsorship, corporate philanthropy and lobbying expenditures are largely undocumented and unregulated by nation states (Aaron &

Siegel 2016), so the extent to which corporate powers also exert power in the landscape of social media is unknown.

Food brands appear frequently in many users' tweets, especially jokes about foods around them being likely to give them diabetes. Coca Cola, Hershey's, McDonalds, Burger King and Oreos are mentioned, to name a few. Users' jokes about these brands imply a sense of powerlessness in the real world food environment. Humorous tweets that tend to be considered irrelevant to public health researchers and policy makers suggest a public resigned to the dominance of food products in everyday life. Users appear to jokingly set their everyday food environment against health advice about how to avoid diabetes. The humorous and flippant quips simultaneously convey a deeper ironic observation about the disjuncture between the food environment and dietary advice.

George Orwell observed that "*a thing is funny when – in some way that is not actually offensive or frightening – it upsets the established order. Every joke is a tiny revolution*" (Orwell 1945). Orwell pointed out that jokes can highlight the relative weakness of established powers, and use of humour can disrupt the established hierarchy. He also observed, that *"whatever destroys dignity, and brings down the mighty from their seats, preferably with a bump, is funny. And the bigger they fall, the bigger the joke"* (Orwell 1945). If the size of a joke is crudely measured by the amount of times it is posted, or the number of retweets it obtains, then the biggest jokes on Twitter where diabetes is concerned bring down the mighty health authorities, and remind them that the world of food and health is much bigger than their narrow interpretations of it. Everyday citizens joke about which foods are going to give them diabetes, and in doing so, highlight an uneven balance of power between citizens, governments and organisations that advocate for healthy diets, and the world around us.

Acknowledgements

MBD acknowledges support from the James S. McDonnell Foundation Postdoctoral Program in Complexity Science/Complex Systems Fellowship Award (#220020349-CS/PD Fellow), and the Oxford-Emirates Data Science Lab. All authors would like to thank Guillermo Garduño of Sinnia for his assistance with the data collection, and Mauricio Barahona for his contribution to an earlier manuscript.

References

Aaron, D.G. & Siegel, M.B., 2016. Sponsorship of National Health Organizations by Two Major Soda Companies. *American Journal of Preventive Medicine*, pp.1–11. Available at: http://dx.doi.org/10.1016/j.amepre.2016.08.010.

Beguerisse-Díaz, M. et al., 2014. Communities, roles, and informational organigrams in directed networks: the Twitter network of the UK riots. *Journal of the Royal Society Interface*, 11(1), p.20140940.

Beguerisse-Díaz, M. et al., 2015. The "who" and "what" of #diabetes on Twitter. , 2, pp.1–23. Available at: http://arxiv.org/abs/1508.05764.

Bennett, W.L., 2012. The Personalization of Politics: Political Identity, Social Media, and Changing Patterns of Participation. *Annals of the American Academy of Political and Social Science*, 644, pp.20–39.

Brewis, A.A., 2014. Stigma and the perpetuation of obesity. *Social Science & Medicine*, 118, pp.152–8.

Carty, J. & Musharbash, Y., 2008. You've got to be joking: asserting the analytical value of humour and laughter in contemporary anthropology. *Anthropological Forum*, 18(789296667), pp.209–217.

Desilver, D., 2016. 5 facts about Twitter at age 10. *Pew Research Center*. Available at: http://www.pewresearch.org/fact-tank/2016/03/18/5-facts-about-twitter-at-age-10/ [Accessed September 13, 2016].

Fast, I. et al., 2015. Social media for public Health: An Exploratory policy analysis. *European Journal of Public Health*, 25(1), pp.162–166.

Ferzacca, S., 2004. Lived food and judgments of taste at a time of disease. *Medical Anthropology*, 23(1), pp.41–67. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14754667 [Accessed July 21, 2011].

Fox, S. & Purcell, K., 2010. Chronic disease and the internet. *Chronic disease and the internet*. Available at: http://www.pewinternet.org/2010/03/24/chronic-disease-and-the-internet/ [Accessed September 11, 2016].

Geismar, H., 2012. Museum + Digital = ? In H. A. Horst & D. Miller, eds. *Digital Anthropology*. London: Berg, pp. 266–287.

González-Bailón, S. et al., 2011. The dynamics of protest recruitment through an online network. *Scientific Reports*, 1, p.197.

Harris, J. et al., 2013. Local health department use of twitter to disseminate diabetes information, United States. *Preventing Chronic Disease*, 10(3), p.E70. Available at: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L369195002%5Cnhttp://limo.libis.be/resolver?&sid=EMBASE&issn=15451151&id=doi:&atitle=Local+health+department+use+of+twitter+to+disseminate+diabetes+information,+United+States.&stitle=.

Hawn, C., 2009. Take two aspirin and Tweet me in the morning: how Twitter, Facebook, and other social media are reshaping health care. *Health Affairs*, 28(2), pp.361–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19275991 [Accessed November 1, 2013].

Hayashi, K. et al., 2003. Laughter lowered the increase in postprandial blood glucose. *Diabetes Care*, 26(5), pp.1651–1652.

International Diabetes Federation, 2015. *IDF Diabetes Atlas (Seventh Edition)*, Brussels: International Diabetes Federation.

Kelly, B. et al., 2015. New Media but Same Old Tricks: Food Marketing to Children in the Digital Age. *Current obesity reports*, 4(1), pp.37–45. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/26627088.

Kleinberg, J.M., 1999. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5), pp.604–632.

Kwak, H. et al., 2010. What is Twitter, a social network or a news media? In *Proceedings of the 19th international conference on world wide web (ACM)*. pp. 591–600.

McCreaddie, M. & Wiggins, S., 2008. The purpose and function of humour in health, health care and nursing: A narrative review. *Journal of Advanced Nursing*, 61(6), pp.584–595.

McLennan, A.K. & Ulijaszek, S.J., 2015. An anthropological insight into the Pacific Island diabetes crisis and its clinical implications. *Diabetes Management*, In press.

Newman, M., 2010. *Networks: an introduction*, Oxford: Oxford University Press.

Orwell, G., 1945. Funny, but not vulgar. *Leader*, July28. Available at: http://orwell.ru/library/articles/funny/english/e_funny [Accessed July 17, 2016].

Paul, M.J. & Dredze, M., 2011. You are what you Tweet: Analyzing Twitter for public health. *International Conference on Weblogs and Social Media (ICWSM)*.

Pew Research Center, 2013. Health fact sheet. Available at: http://www.pewinternet.org/fact-sheets/health-fact-sheet/ [Accessed September 11, 2016].

Porter, M.A., Onnela, J.-P. & Mucha, P.J., 2009. Communities in networks. *Notices of the American Mathematical Society*, 56, p.1082.

Public Health England, 2014. Public Health England Marketing Strategy: 2014 to 2017. , p.25. Available at: http://cogprints.org/280/1/getreal.htm.

Scanfeld, D., Scanfeld, V. & Larson, E.L., 2010. Dissemination of health information through social networks: Twitter and antibiotics. *American Journal of Infection Control*, 38(3), pp.182–188.

Shirky, C., 2008. *Here comes everybody: the power of organising without organisations*, London: Allen Lane. Available at: https://vpngate1.hrz.tu-chemnitz.de/+CSCO+0h756767633A2F2F767272726B63796265722E767272722E626574++/stamp/stamp.jsp?tp=&arnumber=4607925.

Simonsen, M., Skipper, L. & Skipper, N., 2015. Price sensitivity of demand for prescription drugs: exploiting a regression kink design. *Journal of Applied Econometrics*. Available at: http://dx.doi.org/10.1002/.

Sornette, D. et al., 2004. Endogenous Versus Exogenous Shocks in Complex Networks: An Empirical Test Using Book Sale Rankings. *Physical Review Letters, American Physical Society*, 93, p.228701.

Tesco PLC, 2014. Major new partnership announced between Tesco, Diabetes UK and the British Heart Foundation. Available at: https://www.tescoplc.com/news/news-releases/2014/major-new-partnership-announced-between-tesco-diabetes-uk-and-the-british-heart-foundation/ [Accessed October 15, 2016].

Twitter Inc., 2016. Getting started with Twitter. *Twitter Support and FAQs*. Available at: https://support.twitter.com/articles/215585 [Accessed July 31, 2016].

Ulijaszek, S.J. & Mclennan, A.K., 2016. Framing obesity in UK policy from the Blair years, 1997-2015: The persistence of individualistic approaches despite overwhelming evidence of societal and economic factors, and the need for collective responsibility. *Obesity Reviews*, 17(5), pp.397–411.

Wasilewska, E., 2013. *Anthropology of humour and laughter* E. Wasilewska, ed., San Diego: Cognella.