Portrayal of waterpipe (shisha, hookah, nargile) smoking on Twitter: a qualitative exploration

ABSTRACT

Objectives: To describe and characterise social media content in relation to waterpipe smoking using qualitative methods.

Study Design: Exploratory qualitative design

Methods: A representative sample of pre-existing social media content from Twitter relating to waterpipe smoking and written in the English language were collected during a one week period in July 2014. 9,671 tweets were collected; duplicates and retweets were removed leaving 4,439 unique tweets. Data were analysed semiotically (positive, negative, positive and negative, no sentiment, unclassifiable) and thematically. Photographs attached to tweets written by individual users indexed using #hookah (n=299) were subjected to content analysis.

Results: Over half of all tweets were positive about waterpipe smoking (59%), with 3% negative, 21% lacking sentiment and 17% unclassifiable. However, there were variations by likely author of tweet, with 91% of tweets from individual users classified as positive. Twitter users focused on their emotional experience, location, other
products they were consuming alongside waterpipe smoking, and who they were with.

Analysis of photographs highlighted a high degree of synergy between text and visual representations of waterpipe smoking, and two thirds of photographs contained at least part of a waterpipe.

**Conclusions:** Waterpipe smoking may be normalised as an enjoyable activity in this online environment, posing a challenge for public health.

**Keywords:** Twitter, social media, social media analytics, waterpipe, tobacco, smoking, youth smoking

**HIGHLIGHTS**

- Waterpipe smoking is increasing among adolescents
- We characterised tweets from Twitter posts relating to waterpipe smoking for a one week period in 2014
- Over half (59%) of all tweets were positive about waterpipe smoking
- Twitter users described a largely positive experience of waterpipe smoking, where they smoked waterpipe, other products they consumed alongside waterpipe smoking and who they were with
- The majority of photographs individuals posted alongside “#hookah” had been taken inside and contained nobody or one person. Two thirds of images showed at least part of a shisha pipe.
INTRODUCTION

Waterpipes are single or multi-stemmed apparatus used to smoke tobacco which facilitate the passing of smoke through water (or some other liquid of the user’s choice) prior to inhalation. Systematic reviews suggest that waterpipe use results in similar negative health consequences to cigarette smoking, including increased risk of Chronic Obstructive Pulmonary Disorder (COPD), bronchitis, mouth cancer and lung cancer. Until the 1990s, waterpipe smoking had been in decline for many years and was largely prevalent in older men from the Middle East. In recent years, however, waterpipe smoking has been marketed to adolescents as a ‘cool’ behaviour to engage in across the UK and other developed countries, and 9.5% of US college students are regular (monthly) users, although rates of waterpipe smoking among adults are lower in national surveys. Rates of adolescent waterpipe smoking have been increasing or stable in many countries over the past decade, in contrast to cigarette smoking which has been stable or in decline.

School children and university students report that they began waterpipe smoking for a range of factors including peer pressure, curiosity and because it was fashionable. Moreover, there is a general belief among users that waterpipe smoking is less harmful for health, and less addictive, than cigarette smoking. Research with adults
highlighted that it was considered to be a more pleasurable and sociable activity than cigarette smoking, and most began smoking waterpipe with their friends aged under 25 years. Moreover, students have reported that users look ‘cool’, and that waterpipe smoking is a ‘fun’ and cheaper alternative to a night out in an alcohol-serving-premises. Positive beliefs regarding waterpipe smoking have been found to correlate with current waterpipe smoking. This is particularly concerning as waterpipe smoking leads to nicotine dependence in some users and may act as a gateway for adolescents into cigarette smoking.

Over the past decade, the use of social media has become a mainstream activity for many adolescents and has been growing annually among both adolescents and adults. Access may result in the viewing inappropriate or harmful content. Conversely, social media can be used to share health promotion messages, including through the use of images. A survey with a small sample of American adolescents found an association between some health risk behaviours, but not tobacco use, and social media use, although there is a need for further research in this area. More generally, research into how social media is used to discuss tobacco use and other health risk behaviours is in its infancy, but research has been carried out into marijuana and alcohol use. To date, one paper has examined the way in which waterpipe smoking is portrayed on Twitter, identifying that 90% of tweets from those
who were ‘popular’ on Twitter contained pro-waterpipe smoking views \(^\text{23}\). In our research, we seek to extend this understanding of how waterpipe smoking is discussed on Twitter by presenting a more interpretative thematic analysis alongside semiotic commentary and content analysis of linked photographs. More specifically, we aimed to:

1. characterise who was tweeting about waterpipe smoking
2. describe the content of tweets about waterpipe smoking in terms of whether messages were positive, negative or lacked sentiment
3. understand the content of messages circulated on Twitter regarding waterpipe smoking in terms of both (i) the written text and (ii) attached links

**METHODS**

The use of pre-existing content from social media as data for social research is in its infancy. To date, Twitter has been the most investigated platform due to its open access nature. Much of the existing research uses data mining techniques based on computational algorithms\(^\text{24, 25}\). However recent research has focused on the content of
images\textsuperscript{19} and videos\textsuperscript{26} that are shared. Accordingly, we wished to make fuller use of the potential of Twitter data, and as such we also adopted qualitative techniques including interpretative analysis of tweets and content analysis of linked photographs.

**Data collection**

Data collection was facilitated by the open access software Twitter Archiving Google Spreadsheet (TAGS) v5.1.\textsuperscript{27}, which used Twitter’s Application Programming Interface (API) to allow us access to a representative sample of 1\% of all tweets. Following discussions with public health practitioners and third sector organisations, we selected eight pre-specified keywords with an aim of having international reach in data collection. The keywords (hookah, shisha, sheesha, nargile, narghile, waterpipe, hubblebubble, arghila) were collected during a one week period 00.01hrs 10\textsuperscript{th} July – 00.00 17\textsuperscript{th} July 2014. 9,671 tweets were collected over the seven day period, with 58\% (n=5648) containing the keyword “hookah”; 31\% (n=3032) “shisha”; 6\% (n=610) “nargile”; and 3\% (n=253) “sheesha”. All of the four other key words contained 1\% or less of the total data, amounting to a total of 128 tweets. Data collection errors, resulted in data loss on three occasions, and a total of 7 minutes of data was missing from #hookah (total data loss 0.07\% of the time period in which #hookah was
collected), and <1 minute of data from #nargile (total potential data loss 0.01% of #nargile) over the one week period.

As is often the case with pre-existing online data, the Chair of the University ethics committee stated that the project did not need ethical approval as all tweets were in the public domain. Subsequently, we discovered that Twitter prohibits the reproduction of original content from Twitter unless at least opt-out consent has been secured. Due to the delay between data collection and publication, we have not retrospectively sought consent. In order to provide illustrative examples, however, we have anonymized individual Tweets, through the use of correcting misspellings, substitution of individual words for alternative words with a very similar meaning, or removing individual words and replacing these with a descriptor in brackets, for example, when a brand of alcohol is mentioned, it is referred to as (type of alcohol).

Analysis

Data were imported into the qualitative analysis software NVivo 10 for analysis. First, data were coded for a ‘likely author’, to determine if the tweet should be fully analysed (n=4,439), or excluded from the semiotic and thematic analysis (n=5,232). Tweets which were excluded from the full analysis were retweets (symbolised by RT at
the beginning of the tweet, \( n=3343 \), duplicates \( n=1504 \), or tweets that were related to a song called “Hookah” by the artist Tyga \( n=385 \). The remaining tweets \( n=4,439 \) were divided into mutually exclusive categories to show which Tweets appeared to have originated from individual users, businesses, health agencies, media, geographical bloggers and those which were not possible to classify (unclassifiable) (see Table 1). Reasons for tweets being coded as unclassifiable was because they were not written in English or were mainly or entirely comprised of hashtags with very limited narrative.

Second, data were coded semiotically, that is, the text was used to determine if tweets were positive, negative, lacking sentiment or unclassifiable, as has occurred in other research using Twitter as a data source \(^{23, 24}\). Third, text within the tweets were analysed using interpretative thematic analysis \(^{29}\). Finally, the links to external content contained in a subset of the #hookah data set were followed. All links contained within the first 3,000 tweets in this dataset which had been coded as having been written by waterpipe users were followed, and those leading to photographs were subjected to content analysis, using the principles of documentary analysis \(^{30}\). Analysis was undertaken by (author 2) with regular analysis meetings between (author 2 and author 1) in which coding was reviewed with the aid of ‘outputs’ from NVivo, which showed all tweets which had been assigned a particular code. Agreement on
appropriate coding was high and wherever there was initial disagreement, it was resolved through discussion.

RESULTS

Results are presented in relation to the likely author of the tweet, before going on to describe the findings of the semiotic and thematic analysis of the textual data, before finally considering the content of attached photographs.

Likely author of tweets

The likely author of tweets can be seen in Table 1. Over half of tweets (52%) were found to be unclassifiable. Around one quarter of the unique tweets (n=999; 23%) were either partially (n=442; 10%) or completely (n=557; 13%) written in a language which was not English, and these accounted for over one third (39%) of the ‘unclassifiable’ tweets. The other reason for Tweets being found unclassifiable was because they were mainly or entirely comprised with hashtags with very limited narrative.

[Insert table 1 about here]
Within tweets where it was possible to determine a likely author, the majority (n=1496) appeared to originate from individual users, who described their activities and thoughts relating to waterpipe smoking. Tweets originating from businesses (n=551) were sub-divided into those attempting to sell products relating to waterpipe smoking for individuals to consume off the premises (n=141), advertising a premises where it was possible to smoke waterpipe (n=339), or other business (n=71).

**Semiotic analysis**

Data were subjected to semiotic analysis; that is tweets were coded as positive, negative, or neutral. Overall, the vast majority of tweets were positive (n=2,604; 59%), with 85% of these Tweets positive specifically about waterpipe smoking. The second most popular category was those which lacked sentiment (n=922, 21%), with 523 (12%) containing narrative content but no sentiment, and 399 (9%) neither narrative nor sentiment content. Only 3% (n=114) of tweets contained negative content, with only half of this relating directly to waterpipe smoking. The vast majority (99%) of the 747 (17%) unclassifiable tweets were written wholly or totally in a language that was not English and did not contain emoji’s (graphic symbols to represent faces, people or objects) or smiley-face emoticons (typed representations of a smiling face) which could be used to determine sentiment.
Positive tweets from individuals emphasised the enjoyable experience of smoking waterpipe or another activity that they were simultaneously engaged in, whilst businesses highlighted the potential to have a positive experience whilst smoking waterpipe. Negative tweets, although a small minority of the sample, ranged from slightly negative tweets alluding to addiction to very negative tweets highlighting concerns with the prevalence or health impacts of waterpipe smoking. There were considerable differences in the semiotic classification of tweets by likely author of tweet. For example, 91% of individual user tweets, 83% of business tweets and 73% of geographical blogger tweets were positive, whilst only a minority of tweets from the small group of health agency (3%) and media were positive (14%). Few businesses (0.4%) and individual users (1%) highlighted the potential for waterpipe use to be negative in their tweets.

**Thematic analysis**

The 4439 tweets which were subjected to semiotic analysis were also assessed for thematic content. 2558 unique tweets had sufficient content to enable thematic analysis, and a mean of 2.0 codes per tweet were applied. Five main themes were
inductively generated and used to code tweets during analysis: experience of using waterpipe, location, consumed alongside, activities and relationships. The most prominent themes and sub-themes can be seen in Table 2. Other themes are not reported due to low numbers within the data set.

[Insert table 2 around here]

Over one quarter of all codes assigned related to the experience of using waterpipe (n=1,255), with tweeters stating that smoking waterpipe was sociable (n=432), relaxing (n=223) or fun (n=107). However, a minority of tweets also stated the potential for addiction alongside these largely positive themes, showing some awareness of the potential for harm. Only seven tweets explicitly mentioned quitting waterpipe smoking. A further quarter of codes used to describe the data related to a geographical location in which tweeters reported that they had or planned to smoke waterpipe. Tweets relating to premises in which waterpipe was served and geographical locations were from a mixture of individual users and businesses.

The third most popular theme was discussing the consumption of other products alongside waterpipe. The most common accompaniment was alcohol (n=337), but food (n=171), non-alcoholic drinks (N=63) and cannabis (N=41) were also discussed. Again, tweets in these categories were from both individual users and businesses.
Alongside this, a small minority of tweets discussed e-cigarettes (n=82) or the use of e-shisha (n=71).

A minority of tweets (n=414) reported undertaking other activities alongside their discussion of waterpipe smoking, including watching football, listening to music or partying. Tweets relating to football focused on the FIFA 2014 World Cup, Brazil, which was in progress at the time of the research, and were largely generated by individuals. By contrast, tweets relating to music were mostly from businesses who played music, but also had waterpipes available.

Finally, in 15% (n=374) of tweets, other people were mentioned, including friends, family and partners. A very small minority (n=12) explicitly reported waterpipe use alone, and this was sometimes viewed positively, associated with relaxing, whilst other times boredom was reported.

**Content analysis of images**

In order to gain a detailed analysis of external content relating to hookah, the most commonly used word for waterpipe smoking within the dataset, further analysis occurred. Within a sub-sample of 3000 tweets within the #hookah dataset, 433 tweets
had been coded as having an individual waterpipe user as the likely author and also contained a URL linking to an external site. Of these, 136 links were excluded from analysis (see Table 3). Within the images subjected to analysis (n=299), the majority contained only a single image and were colour photographs. Two thirds of photographs contained a waterpipe, although the focus of the image was rarely on the entire pipe. It was common for images to contain beverages (n=69), with a mixture of alcoholic and non-alcoholic drinks. Images accompanying tweets were largely taken indoors (n=210), including bars and at home. Overall, images were most likely to contain one person or nobody rather than larger groups. There was a high level of synergy between the text in tweets and the photographs that accompanied them, with the content of most of the photographs relating to the text included in the text (98%). Moreover, the images containing people echoed the thematic and semiotic content of tweets in representing waterpipe smoking as sociable and/or enjoyable, and broadly positive.

DISCUSSION

This research has provided a unique insight into how waterpipe smokers who post on the social media site Twitter document and represent their waterpipe smoking, which
is largely related to socialising and having ‘fun’. In the vast majority of tweets relating to waterpipe use in one week of July 2014, waterpipe smoking was not viewed negatively; it was either explicitly represented as positive or unclassifiable. As previously identified in research, the definition of waterpipe smoking as enjoyable was a key element of the reported behaviour. The thematic analysis described the emotional and practical experience of using waterpipe and the location of waterpipe use. As has been previously identified, waterpipe smoking was sometimes identified as relaxed, but it was also reported to be fun. Our analysis of images accompanying the textual elements of tweets was consistent with these findings, showing waterpipe smoking as an enjoyable and positive experience.

The research also gives a preliminary understanding of how businesses are using Twitter to encourage their followers to smoke waterpipe. This includes a range of special offers on food, alcohol and waterpipe, and the promotion of the idea that smoking waterpipe will be a positive experience. Businesses allowing waterpipe smoking on the premises catered to both a relaxed and to a livelier clientele. As has been identified previously in relation to tobacco control laws, there is poor regulation of waterpipe smoking, and it would be challenging to enforce marketing regulations.
on social media sites. However, in order to reduce breaches of article 13 of the World Health Organization’s Framework Convention on Tobacco Control, which prohibits tobacco advertising, this should be addressed with some urgency. Public health bodies in countries with high waterpipe smoking prevalence should make fuller use of social media to disseminate public health messages relating to waterpipe smoking, but crucially should ensure the accuracy of their messages and seek engagement from young people in their campaigns to design acceptable messages. Alongside this, it is important to develop new methods to describe social media content and to evaluate online public health campaigns.

There are obvious limitations to research conducted using ‘Big Data’ and the use of social media as data, as data show only a snapshot in time and it is not possible for researchers to retrospectively seek clarification to determine meaning. Moreover, we collected data for a one week period during the month of Ramadan, where strict fasting occurs between sunrise and sunset, and tobacco usage may vary. Further methodological limitations include a limited selection of key words (notably excluding goza and narguile) and variations in spellings, which is likely to have reduced data capture from some cultures. Furthermore, the analysis of external content was undertaken on the hookah dataset alone, and this may have reduced the analysis of external content from the Middle East, where the term hookah is less commonly used.
Despite the limitations of the current sample, over 4,000 unique micro-blogs relating to both the experiences and marketing of waterpipe smoking were analysed. The sheer volume of participants is unlike that found in routinely conducted qualitative research which is time consuming (and costly) to conduct, but still allowed for qualitative analysis of tweets and associated photographs. Our classification of tweets into likely author based on the text of the tweet was incorrect in 1% of cases when linked images were viewed, showing high reliability in the coding of text alone. Moreover, our findings mirror those of the only other study of waterpipe discussions on Twitter, and extend their finding to a more general sample of tweets (as opposed to only high influence tweeters), and validate such analysis through the use of content analysis of linked photographs. Further research should be undertaken in order to understand if the online discussion of waterpipe smoking varies outside of Ramadan, and could focus on specific populations, including businesses who are promoting waterpipe smoking.

ACKNOWLEDGEMENTS: Many thanks to Dr Lucy Bennett who provided methodological support during the data collection process.
FUNDING: The research was funded by the Cardiff University Respiratory Bequest Fund. The Fund had no influence on the research design, procedures and dissemination.

COMPETING INTERESTS: None declared.

ETHICAL APPROVAL: Not required for this study, as confirmed by the Cardiff University School of Medicine research ethics committee, as all data were in the public domain.

REFERENCES

13. Soneji A, Sargent JD, Tanski SE, Primack BA. Associations Between Initial Water Pipe Tobacco Smoking and Snus Associations Between Initial Water Pipe Tobacco Smoking and Snus JAMA Paediatrics. 2014; Published online December 08 2014.