Health communication and adolescents: what do their emails tell us?

Kevin Harvey\textsuperscript{a}, Dick Churchill\textsuperscript{b}, Paul Crawford\textsuperscript{c}, Brian Brown\textsuperscript{d}, Louise Mullany\textsuperscript{a}, Aidan Macfarlane\textsuperscript{e} and Ann McPherson\textsuperscript{f}


**Background.** It is widely known that barriers exist in communication between adolescents and health professionals. However, little is known about the actual language used by young people articulating such difficulties and whether email might allow them to overcome these problems.

**Objectives.** The aims of this study were to investigate concerns and difficulties relating to communication among adolescents seeking online health advice.

**Methods.** The study design was a corpus linguistic analysis of a million-word adolescent health email database based on 62,794 emails from young people requesting health advice from a prominent UK-hosted and doctor-led website.

**Results.** Young people reported various concerns about their health. They described numerous difficulties in disclosing such concerns to other people, in particular to parents and doctors. However, they readily expressed their concerns by email, displaying elevated levels of directness, particularly in relation to potentially sensitive or embarrassing topics.

**Conclusion.** Email has the potential to facilitate and supplement face-to-face consultations with health professionals. Increased adoption of email by health providers may be an efficient means of engaging with a generation often reluctant to access more traditional health care services and thus encourage them to enter the primary care setting more readily.

**Keywords.** Adolescents, communication, corpus linguistics, doctor–patient relationship, email, Internet, primary care.

**Introduction**

A number of researchers have described the difficulties that young people have in communicating health concerns to professionals\textsuperscript{1,2}, including GPs\textsuperscript{3}.

These difficulties in communication relate to both personal and structural barriers. For example, owing to a poor understanding of confidentiality issues and the consequent fear of potential breaches in confidence by service providers\textsuperscript{4,5}, teenagers are often reluctant to disclose their health concerns to and request personal health advice from doctors\textsuperscript{6}. The structure of the medical consultation itself is also a potential barrier to more open and sustained interaction, with young people reporting that the brevity of exchanges with doctors does not allow for the satisfactory delivery of care\textsuperscript{4}. When consulting with doctors, adolescents have reported feeling marginalized and being unable to contribute more fully during the practitioner–patient exchange\textsuperscript{2}.

With adolescence being a time of physical, emotional and social change with distinctive health needs\textsuperscript{7}, such communicative difficulties may impede adolescents seeking appropriate face-to-face advice from health care providers. Despite their reluctance to consult with doctors, many young people wish to discuss health concerns with their GPs but are only likely to do so if they feel comfortable with the doctor they are seeing\textsuperscript{8}.
In view of these difficulties, adolescent health specialists have suggested the need for specific health provision tailored to the needs of young people. In particular, there have been calls for alternative information services for young people, including peer support, helpline and online services. The Internet affords increasing opportunities for obtaining health advice and information, constituting a convenient and anonymous route to services and products that previously required direct contact with health care providers. Adolescents are prime and early users of electronic resources, using the Internet for a range of recreational, occupational and personal activities, as well as for seeking health advice and information. There has also been an increase in the number of reputable, doctor-regulated Internet sites that have become a popular source of health advice and information for teenagers, with the electronic gateway offering confidential advice and information from trained professionals that otherwise might be difficult or compromising to obtain. The interactivity afforded by adolescent health websites offers a potential opportunity for gaining a rich and potentially deeper understanding of the barriers, and conversely the promoting factors, in adolescent health communication.

Corpus linguistics is an approach that is becoming increasingly popular in the linguistic analysis of health communication. This technique involves interrogating large data sets or 'corpora' of language and applies both quantitative and qualitative techniques to the analysis of language and is thus able to overcome some of the potential drawbacks of these approaches used in isolation. For example, although qualitative studies in health communication have afforded penetrating descriptions of particular interactions, they are often limited to small data sets. Consequently, their findings will not necessarily represent the wider field of language employed in that particular setting. Conversely, quantitative approaches, although able to ground generalizations on more substantial and representative samples of language, yield data which is often deprived of context and so are unlikely to be sufficient for providing a situational understanding of language use. Corpus-based studies involve analysing extensive data sets, consisting of thousands or millions of words so as to discover patterns of use which may be then subjected to more interpretive analyses.

In order to find out about young people’s concerns relating to the communication of health-related issues, this paper describes the use of corpus linguistic techniques to explore a database of text derived from emails sent by young people to GPs on one particular teenage health website.

Methods

The data source
We analysed the content of a one-million-word corpus of emails sent to the adolescent health website, Teenage Health Freak (http://www.teenagehealthfreak.org). Operated by UK-based doctors specializing in adolescent health, the Teenage Health Freak website has been running and continuously updated on a weekly basis since its launch in 2000. The site is designed to be interactive, confidential and evidence based, providing adolescents with accessible advice and information pertaining to a broad range of health issues.

The interactive feature of the website allows teenagers to email their health questions in confidence to the online GP, Dr Ann. Given the large influx of messages that the site receives on a daily basis, it is not possible for the website doctors to respond to all the email requests for help. Consequently, only a small number of messages are answered and these are published on the website. Although we here use the term emails to refer to the messages posted by the adolescents, these messages do not constitute emails in the traditional sense. Rather they constitute one-off postings, which may or may not receive replies. Our analysis focuses solely on these adolescent communications rather than the professionals’ comparatively infrequent returns, although in future research we intend to follow up analysis into the doctors’ responses.

We were given permission by the Teenage Health Freak operators to collect and analyse the emails sent to the website between January 2004 and December 2005. Comprising 62 794 messages, this provides a substantial snapshot of the health concerns communicated on a daily basis by teenage contributors. The Teenage Health Freak website possesses a privacy policy informing contributors that their requests may be used for research purposes and that, in using the website to transmit such information, they consent to the collection and use of data which they provide.

Keyword, collocational and concordance analyses
We employed the WordSmith Tools software program to first create a list of keywords that appear in the teenage health messages. Keywords are words that are unusually frequent in comparison with general everyday English. Keywords can be described as words that reveal the ‘aboutness’ or content of a text or texts. Unlike pure frequency lists, which calculate the frequencies of every word appearing in a corpus,
keywords provide a clearer picture of the salient themes in a text and hence are a more useful technique for identifying words that warrant further thematic exploration. Thus, a keyword analysis constitutes a reliable means for best defining a particular language variety and, in the case of this study, characterizing the health language of teenagers.

We compared the corpus of teenage health emails with a one-million-word collection of general spoken English from the British National Corpus, which yielded 1160 keywords. Spoken language is an appropriate language variety with which to compare the adolescent emails since, although the messages are written, they are in an informal register so speech is their more appropriate comparator, rather than the more literary or technical styles typically found in written English.

The keywords so extracted from the email corpus related to a broad range of health concerns and topics, the most dominant area concerning sexual health, which we have addressed elsewhere. Here, we are concerned with adolescents’ concerns and difficulties relating to communicating health concerns and requesting help from other people. Consequently, we identified keywords surrounding the subject of communication, including verbs that expressly related to verbal interaction and keywords associated with the activity of advice and information seeking.

Once the keywords were established, we then conducted a collocational analysis of these signal items in order to explore the meanings and associations between them. Collocation is the linguistic phenomenon whereby two words typically appear together; it is the tendency of one word to attract another. Collocation can be measured informally (for example, by observing that, say, ‘blond’ is much more likely to collocate with ‘hair’ than it is with ‘paint’) or, more reliably, by computationally using a statistical measure. We used WordSmith to calculate the mutual information (MI) scores of these words (collocates) that were commonly associated with each keyword in the adolescent health emails. Derived from information theory, MI is a measure of the strength of collocation, calculating the extent to which words appear together compared with chance. The higher the MI score, the stronger the collocation (i.e. the firmer the link between co-occurring words). Conventions in corpus linguistics take an MI rating of $>3$ as significant, that is as being indicative of a strong collocation. Accordingly, we only recorded collocates that were either equal to or above this value of significance.

The final stage of our study involved a concordance analysis of the keywords in order to establish their patterns of use in situ. Although examining collocates affords an immediate overview of the themes and topics surrounding a given keyword, a collocational analysis alone provides limited information concerning how words function in context. Therefore, in order to appreciate more subtle and detailed uses of meaning, a concordance analysis was necessary. Concordance analysis enables researchers to examine every occurrence of a particular word or phrase in context and thus identify its patterns of use. This enabled us to systematically describe the connotations of the keywords and so provide insights into the teenagers’ attitudes concerning the content of their messages. Ultimately, analysing the context of keywords enabled us to determine whether the adolescents had positive or negative experiences in communicating health concerns.

In order to appreciate the evidence for negative or positive evaluation, a large number of concordance lines are needed. The keywords we identified typically occur with high frequencies (for example, there are 2818 instances of the keyword ‘tell’ in the teenage health emails). However, such high occurrences make it impossible, or at the very least infinitely time consuming, to look closely at each individual concordance line. Accordingly, we adopted Sinclair’s procedure, well established in corpus linguistics research, of first randomly selecting 30 concordance lines, observing their patterns, then proceeding to another randomly chosen 30 lines, then another, until a saturation point was reached (i.e. where no new patterns are apparent).

### Results

When compared with general spoken English, we found that the adolescent health emails contained a number of keywords pertaining to communication. These included verbs relating expressly to face-to-face spoken interaction, as well as verbs and nouns associated with the specific activity of advice seeking (Table 1). Although these keywords individually constitute only small percentages of the total one-million-word count that makes up the adolescent email corpus, their frequencies in relation to general English were significantly elevated, indicating that a key factor for

<table>
<thead>
<tr>
<th>Word</th>
<th>Teenage Frequency</th>
<th>Teenage %</th>
<th>BNC spoken Frequency</th>
<th>BNC spoken %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tell</td>
<td>2818</td>
<td>0.28</td>
<td>777</td>
<td>0.08</td>
</tr>
<tr>
<td>Ask</td>
<td>1132</td>
<td>0.07</td>
<td>185</td>
<td>0.02</td>
</tr>
<tr>
<td>Talk</td>
<td>552</td>
<td>0.05</td>
<td>177</td>
<td>0.02</td>
</tr>
<tr>
<td>Answer</td>
<td>361</td>
<td>0.04</td>
<td>132</td>
<td>–</td>
</tr>
<tr>
<td>Advice</td>
<td>309</td>
<td>0.03</td>
<td>11</td>
<td>–</td>
</tr>
<tr>
<td>Question</td>
<td>293</td>
<td>0.03</td>
<td>135</td>
<td>–</td>
</tr>
<tr>
<td>Advise</td>
<td>77</td>
<td>–</td>
<td>6</td>
<td>–</td>
</tr>
<tr>
<td>Explain</td>
<td>75</td>
<td>–</td>
<td>17</td>
<td>–</td>
</tr>
</tbody>
</table>

BNC, British National Corpus.
teenagers concerning health issues involved the process of communication. Although the keywords revealed a preoccupation on the part of adolescents with communication, particularly verbal interaction, the words themselves did not disclose any negative or positive evaluation (the verb tell, for example, does not in isolation, removed from any context of use, convey any inherent attitude). However, collocational analysis presented clear evidence that the keywords were situated around negatively laden collocates that constructed communicating about health to be fraught with problems (Table 2). The most frequently appearing keywords pertaining to interaction (tell, talk, ask and answer) shared a number of identical and related collocates that intimated a negative viewpoint. For example, adjectives which described the participants as being ‘afraid’, ‘scared’, ‘worried’, ‘embarrassed’ and ‘stressed’ were found to recur with communication keywords.

A second recurring pattern involving keyword collocates concerned terms referring to potential participants in communication. In particular, of the eight keywords, five were found to collocate significantly with the word ‘Dr’ as well as related forms such as ‘doctor’, ‘doctors’ and ‘doctor’ and ‘GP’. Other common collocates describing communication sources included ‘parents’, ‘family’, ‘mum’, ‘dad’ and ‘friends’. Further contextual examination of the keywords via concordance lines described an emerging picture of the problems experienced by adolescents communicating about health matters, reinforcing the negative assessment provided by the range of collocates.

In particular, concordance analysis revealed a fraught picture of communication in which doctors, parents, family and friends were deemed to be problematic sources of health advice (Fig. 1). Although adolescents expressly signalled a pressing need to discuss health concerns and obtain professional and familial advice, their emails reported troubles in confiding with others (Fig. 1: lines 3–5), dilemmas over whether it is right to disclose problems to other people (Fig. 1: line 9) and issues concerning their ability to express their complaints owing to embarrassment (Fig. 1: lines 1 and 2). Often the reasons for being reluctant or unable to discuss health issues with others were left unelaborated. However, there were times when the adolescents provided reasons for their communicative difficulties. These were brought clearly into relief when we examined in context the collocates referring to participants in communication, namely health professionals (e.g. GP, doctor and doctors), family (parents, mum and dad) and friends. For instance, Figure 2 shows extended concordances (i.e. greater textual context around the collocates) for both GP and parents.

The concordance lines revealed that adolescents articulate various reasons for their difficulties in sharing problems with health professionals (as exemplified by the references to GPs in Fig. 2). The most common explanation related to concerns about confidentiality, a finding that has been extensively identified elsewhere in the health communication literature.7,8 Specifically, the adolescents expressed concern over their parents finding out about a health problem they might have (Fig. 2: lines 2 and 3) and fears that practitioners might divulge their health complaints to parents (Fig. 2: lines 1 and 19–21). Another reason for the adolescents’ reluctance to consult involved the communication of intimate, sensitive issues, particularly sexual health concerns (Fig. 2: lines 4–7 and 13–17).

At a more fundamental level what many of these explanations disclose is uncertainty about health care services, in particular the role of the doctor and consultation protocol: for instance, the belief that GPs are liable to disclose information concerning their health to parents. Similar uncertainty about and misunderstanding of the remit of health professionals and general practice was further evident in the adolescent emails. For instance, complaints pertaining to sexual health were sometimes construed by adolescents as being problems outside the concern of general practice, with uncertainty being communicated as to whether GPs were able to respond to such problems (‘can u go to the doctor about an STI’s check up being under 16’, ‘can

---

**Table 2** Content words (with MI scores of >3) that collocate with keywords concerning communication

<table>
<thead>
<tr>
<th>Tell</th>
<th>Abortion, afraid, alone, boyfriend, dad, doc, doctor, doctors, eating, embarrassed, friends, family, gay, go, GP, leave, lose, mates, mum, need, nipples, normal, older, parent, parents, people, person, pill, police, raped, really, risk, said, scared, see, shy, stressed, teacher, things, think, trust, try, wondering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ask</td>
<td>Advice, afraid, boy, boys, courage, dad, doctor, embarrassed, find, friend, friends, girl, girls, help, mum, parents, people, question, really, scared, stupid, sorry, thought, wanted</td>
</tr>
<tr>
<td>Talk</td>
<td>Afraid, boyfriend, dad, doctor, family, feel, find, friend, friends, girls, hard, mates, mum, parents, phone, people, scared, school, shy, stressed, tell, time, told, wants</td>
</tr>
<tr>
<td>Answer</td>
<td>Dr, email, getting, give, know, need, really, please, question, questions, worried, wrong</td>
</tr>
<tr>
<td>Advice</td>
<td>Ask, get, give, good, like, need, please, really, talk, thanks</td>
</tr>
<tr>
<td>Question</td>
<td>Ask, asking, answered, doctor, Dr, health, help, know, name, need, really, reply, sent, stupid, worried</td>
</tr>
<tr>
<td>Advise</td>
<td>Give, help, need, please</td>
</tr>
<tr>
<td>Explain</td>
<td>Know, please, hard</td>
</tr>
</tbody>
</table>
my gp help me if i have an sti?). These concerns echo the findings of previous research that has described how adolescents are often unsure of where to go for help for reproductive health (such as sexually transmitted disease treatment) and mental health services.25,26

The desire for secrecy and concerns over confidentiality were equally characteristic of the concordance lines referring to ‘parents’. Yet other reasons for the non-disclosure of troubles were further evident. These included potential rebuke from parents after disclosure, as well as doubts that parents would be able to provide appropriate help and support (Fig. 2: lines 8–12, 18–21 and 29–31). Interestingly, however, despite their inability to communicate with parents about their health complaints, the adolescents occasionally intimated that they wanted to consult with their parents if only they could overcome the difficulty of first broaching their troubles (Fig. 1: line 15; Fig. 2: line 25).

Conclusions

This is the first published study to describe the use of corpus linguistic techniques to explore a corpus of
email texts from young people concerning their attitudes to communication with health professionals. The corpus of texts on which this analysis is based is extensive, including one million words from 62,794 email messages. The Teenage Health Freak website itself is very popular with young people, having received 60 million visits between February 2000 and January 2007, an average of 52,864 hits a day. Thus, although we cannot make precise demographic claims about the nature of the population visiting the website, it is likely to be accessed by a wide range of young people. Unlike the data from other interactive health sites which often standardize messages sent by advice seekers, the corpus data interrogated in this study consist of the entire, unedited queries and concerns posted to the site doctors. Consequently, this afforded access to the original word choice of the email requests, which retained all the nuances of individual expression and, arguably, the original communicative intentions of the correspondents.

Corpus linguistic techniques are a powerful method of analysing large databases of naturally occurring text: the initial quantitative approach of keyword identification defines the important areas for subsequent qualitative collocational and concordance analysis in order to understand the context and relevance of the language used. Our research was not generated by the ‘outsider perspective’, that is it was not prompted by what researchers deemed to be the issues but rather sought to explore what young people themselves saw as pertinent. In contrast to much research on health care communication, which is typically based on patients’ responses to professional-initiated actions, this paper is unique in that it focuses on what are patient-initiated actions, placing the responsibility and the main say in the hands of young people. The fraught issue of communication with health care practitioners and parents was signalled by the adolescents themselves.

Adolescent experiences of health communication commonly involve descriptions of a reluctance or inability to consult with doctors, parents and friends. A key finding of this study is that teenagers in their emails to online health professionals regularly express difficulty in first approaching adults face to face for advice and information about their health concerns. In particular, the recurring use of verbs such as ‘tell’ and ‘talk’, which emphasize verbal interaction, and their negatively loaded collocates, point specifically to problems directly engaging in face-to-face exchanges.

Our results show that adolescents are liable to have health problems which they may be reluctant to share through established institutional routes and traditional face-to-face support networks including parents and doctors. For some young people, access to primary care is problematic—as is the option of approaching their parents or friends to broach a health concern. Consequently, access to Internet sources of health provision may be their prime or, in some cases, their only recourse when in need of medical advice—the only source of help they feel able to utilize in times of turmoil. The anonymity afforded by the Internet, specifically its facility of providing advice without the risk of parents finding out about their concerns, make it a viable source of health information for adolescents.

Unlike traditional health care services, where the emphasis is on the client attending an unfamiliar environment, typically at a time dictated by the professional, the Internet provides convenient unconstrained access to health services. In terms of communicating with professionals online, email allows young people to formulate their problems in their own terms, space and time, affording them a platform from which to ask awkward, sensitive or detailed questions without the fear of being judged or stigmatized—a negative outcome commonly identified in our health email corpus.

The results show that young people are adept at articulating their health concerns electronically (concerns that they otherwise might not have communicated to others), doing so with high degrees of candour and directness. Given the regularity with which electronic forms of communication are used by young people, with, for example, 75% of today’s youth having used the Internet to seek health advice, there is arguably greater scope for more extensive and systematic use of email in health care. As the popularity of interactive websites specializing in adolescent health testifies, email and online messaging have the potential to reach out to and connect with young people who might be reluctant to engage with more traditional face-to-face health services.

The advantages and disadvantages of email consultations between patients and health professionals have been widely debated. With concerns of consultation quality, liability and workload, some practitioners are understandably sceptical about its utility. However, the benefits of the medium are well documented. Email communication is ideal for short questions, brief updates and follow-ups and as a first point of contact. It is less suitable for complex issues, which need to be addressed face to face or for situations where physical examinations are necessary for a diagnosis. Email therefore can never be a substitute for face-to-face consultations. However, we see the medium as a means of supplementing the traditional consultation—a communicative resource through which young people would first have the facility to communicate their health concerns with a practitioner anonymously before following up with a face-to-face consultation if required. For young people reluctant to visit their GP, electronic contact with a health professional in the first instance might well encourage subsequent face-to-face visits, helping them to enter the primary care setting more readily and thus maintain continuity with their doctors.
Despite health services being slow to use email for doctor–patient communication, it is important for health care professionals to understand that, in the digital age, their role is changing in that ‘they increasingly have to act as health guides, mediators, and information brokers for patients and consumers’. Professional resistance to the application of Internet and email services in health care remains a substantial obstacle to the realization of the new technologies’ potential. Part of this potential lies in the opportunities to connect with people who may have health needs but find other kinds of communication embarrassing or difficult to initiate.

It is therefore crucial for health professionals to understand and respond to not just how new information services are affecting the delivery of health care but also how digital technologies are rapidly shaping and changing contemporary communicative practices. Awareness of these impacts is particularly important in relation to adolescent health since today’s teenagers, the so-called Millennial or net generation, have grown up with digital technologies in a world mediated by digital texts—distinctly different from adults’ ideas about literacies. Adolescents feel that their communicative practices and literacies are not being recognized and accommodated by mainstream educational institutions. Such lack of accommodation parallels the complaints that adolescents have in relation to the lack of specialized health services and the lack of recognition of them as a distinct group with specific needs and as individuals. In utilizing their preferred medium of communication, online health advice may promote better engagement with this population.

There is a need for further research into how the provision of email communication between young people and primary health care professionals could influence consultation behaviour and alter health outcomes. It will also be important to identify the extent to which young people are able to share health concerns by email with known health professionals in contrast with teenage health websites since it may be the anonymity of the latter that encourages maximum disclosure of their concerns. This is also likely to be influenced by the perceived confidentiality of such a service as well as how it is marketed. Cost-effectiveness and remuneration issues would need to be considered if such services were to become a routine part of primary care.

We have demonstrated that exploration of data sets of this kind can provide a valuable source of insight into the vocabularies used, the concerns raised and issues attended to by the people whom the website addresses. As we have shown here, these data can be used to examine participants’ own accounts of the difficulties which beleaguer their attempts to succeed in face-to-face encounters. Thus, to anyone concerned with the health care of adolescents, such a data set is a valuable resource to get to grips effectively with this client group. For as Pasteur put it in a different context, the likelihood of a successful health care encounter ‘favours the mind prepared’.

**Declaration**

**Funding:** None.

**Ethical approval:** Not applicable.

**Conflicts of interest:** None.

**References**