Readability, Coherence, and Content Analysis of Online Hidradenitis Suppurativa Patient Resources

by

Barry Michael McGrath

9224041

A thesis submitted for module EL6013 as part of the M.A. in Technical Communication and E-Learning to the University of Limerick

September 2018
Abstract

Increasingly, people are seeking health-related information online, and patient-oriented websites offer tremendous potential to educate and empower people. However, concerns have been raised about the quality of online patient resources as the internet is largely unregulated. This study assessed the text readability, coherence, and content of online resources for those with the chronic skin condition hidradenitis suppurativa (HS). This research evaluated online interest in HS using search engine analytics. The popular HS websites were assessed for text readability and coherence. The word counts, proportion of polysyllabic terms, average sentence length, and use of images and videos were also analysed. Recently, there has been a large increase in HS-related online search activity and in the volume of online HS content. The 21 websites analysed were all written above recommended readability levels. One readability formula significantly overestimated readability. Most of the websites’ coherence levels were appropriate. There was variation in the relevance, timeliness and presentation of the information on the HS sites. The information on many top-ranking HS websites may be too difficult for the intended audience to understand. Content creators should consult published guidelines on improving the comprehensibility of online health information for HS website revisions and future sites.
Table of Contents

Abstract ................................................................................................................................. ii
Dedication ............................................................................................................................ vi
List of Figures & Tables ....................................................................................................... vii
Chapter 1. Introduction ...................................................................................................... 1
  1.1 Subject of this dissertation ...................................................................................... 1
  1.2 Purpose of the study ............................................................................................... 3
  1.3 Research hypothesis and questions ......................................................................... 3
    1.3.1 Hypothesis ......................................................................................................... 3
    1.3.2 Research questions ........................................................................................... 3
  1.4 Scope of the dissertation ......................................................................................... 4
  1.5 Layout of the dissertation ....................................................................................... 4
Chapter 2. Hidradenitis Suppurativa ............................................................................. 5
  2.1 Definition and diagnostic delay ............................................................................... 5
  2.2 Stages ....................................................................................................................... 5
  2.3 Epidemiology .......................................................................................................... 7
  2.4 Causes and risk factors ............................................................................................ 8
  2.5 Treatments ................................................................................................................ 9
    2.5.1 Lifestyle modifications .................................................................................... 10
    2.5.2 Medical management ...................................................................................... 10
    2.5.3 Surgery and/or laser intervention ................................................................... 11
    2.5.4 Psychological support ..................................................................................... 11
  2.6 Impact ....................................................................................................................... 12
Chapter 3. Literature review .......................................................................................... 14
  3.1 Text comprehension .................................................................................................. 14
    3.1.1 Text readability .................................................................................................. 14
    3.1.2 Text coherence .................................................................................................. 16
    3.1.3 Text content analysis ....................................................................................... 17
  3.2 Assessing online patient education websites .......................................................... 17
  3.3 Online hidradenitis suppurativa resources ............................................................... 20
Chapter 4. Methodology .................................................................................................. 21
  4.1 Research design ........................................................................................................ 21
  4.2 Data gathering .......................................................................................................... 21
“Illness is the most heeded of doctors: to kindness and wisdom we make promises only; pain we obey.”

Marcel Proust, Sodom and Gomorrah
Dedication

This work is dedicated to all those living with hidradenitis suppurativa.
List of Figures & Tables

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hurley stages of hidradenitis suppurativa lesions</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>Hidradenitis suppurativa papers by decade</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>Hidradenitis suppurativa papers since 2015</td>
<td>29</td>
</tr>
<tr>
<td>4</td>
<td>Online Interest in hidradenitis suppurativa since 2016</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>New Dale–Chall formula readability scores for hidradenitis suppurativa websites</td>
<td>31</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Hurley hidradenitis suppurativa classification scheme</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>Readability scores and their equivalent US grade level and ease of reading rating</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>New Dale–Chall scores with their equivalent US grade level and ease of reading rating</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>Readability scores of the hidradenitis suppurativa websites analysed in this study</td>
<td>33</td>
</tr>
<tr>
<td>5</td>
<td>Coherence scores of the hidradenitis suppurativa websites analysed in this study</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>Associations between readability formula scores and proposition density</td>
<td>36</td>
</tr>
<tr>
<td>7</td>
<td>Content analysis of the top Google-ranking hidradenitis suppurativa (HS) websites</td>
<td>37</td>
</tr>
</tbody>
</table>

Appendix

1 List of hidradenitis suppurativa websites with their Google ranking, inclusion/exclusion status and category type | 71
Chapter 1. Introduction

Information can empower people to reflect and make informed decisions to change and improve their situations in wealth, health, and overall well-being. With the recent advent of the internet, vast amounts of information are now rapidly available to many people worldwide. The internet has profoundly impacted many aspects of modern life, and health is one such area to have experienced this impact. This study looks at the availability and comprehensibility of health information for those living with a chronic skin condition. In this chapter I outline the contents of this dissertation, focusing on the subject, purpose, scope and limitations, and finally the layout.

1.1 Subject of this dissertation

The patient–physician relationship is changing in this digital age. While traditionally health consumers sought health-related information from their physician and other healthcare professionals, the internet is now the primary source of health information for many (Tan and Goonawardene 2017) and ~80% of people start health-related searches by going online first (Fox and Duggan 2013). Health-related topics are common web-based searches and with the growth of the internet, numerous patient information websites have been published in recent years. However, unlike traditional forms of media the internet is largely unregulated, and concerns have been raised about the quality of online health-related information (Zhang et al 2015). Health consumers can now access online health-related information safely, quickly and privately with tremendous potential to educate people to make informed lifestyle choices that may benefit their health and society at large. To realise this potential, however, it is crucial that the information is understandable by the intended audience.

Communicating health-related information is an important but challenging task. Science, medical and other professional writers are tasked with producing scientifically accurate and stimulating texts at the right level for their target audiences. However, one of the challenges in conveying health-related information is to find common ground between professional and medical writers, who typically have advanced college degrees, and the target audiences, many of whom may have limited or little scientific knowledge.

Improving health literacy levels in the general population has attracted increasing attention worldwide in recent years due to the societal, political, economic and health benefits that enhancing health literacy can bring. To understand the importance of health literacy, it is
important to firstly explain the term. Kickbusch and Maag (2008, p. 208) define health literacy as “the ability of citizens to make sound decisions concerning health in daily life—at home, at work, in health care, at the market place and in the political arena. It is a critical empowerment strategy to increase people’s control over their health, their ability to seek out information, and their ability to take responsibility”. The European Centre for Disease Prevention and Control (2018) describe health literacy as “the capacity that an individual has to access and effectively use health-related information, in order to promote and maintain good health”. The World Health Organization (WHO 2018), explain health literacy as “the cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand and use information in ways which promote and maintain good health”. For the current study, I use the term health literacy to mean that people can find appropriate health information, be able to understand what they find, and to act appropriately to improve their health. Improving people's access to health information and their capacity to use this information effectively is critical to patient empowerment, leading to personal and social benefit (WHO 2018).

Low health literacy costs society and is associated with poor health outcomes and low use of health care services (Berkman et al 2013). Recent studies indicate that low health literacy is prevalent globally. Sørensen et al (2015) report that among people surveyed in eight European Union countries, almost one in two adults had limited (insufficient or problematic) health literacy. Duong et al (2017) found similar results in their study on six Asian countries. Almost one third of Israeli adults had inadequate or problematic health literacy (Levin-Zamir et al 2016). Four of five Egyptian adults surveyed had limited comprehensive health literacy (Almaleh et al 2016). Over half of a Brazilian population surveyed had problematic or insufficient health literacy (Carvalho et al 2016). In the USA, almost one in three adults have basic or below basic health literacy (Gill 2013) and low health literacy is associated with hundreds of billions of US dollars in additional healthcare costs and poor health outcomes (WHO, 2013; Lecuk 2018). In response to these low health literacy levels, associations in the USA, including the US Department of Health and Human Services (USDHSS) and The American Medical Association (AMA), have published recommendations for writing health information for patients to make them accessible for those with low health literacy. The USA is one of the few countries to publish such guidelines.
1.2 Purpose of the study

This study aims to examine the readability, coherence and the information content of some popular websites for those living with the chronic skin condition hidradenitis suppurativa (HS). As also seen for many other health conditions, there has been a large increase in the number of online HS patient resources in recent years. Currently there are few studies examining the information content on HS websites, with just one report assessing HS information readability, and none specifically addressing the information coherence and content on the popular HS websites. The current study seeks to address this gap in knowledge.

The purpose of this study is to perform readability, coherence, and content analyses of the popular online resources for those living with the chronic skin condition, HS. This work explores if the information found on the top-ranking HS websites is written and presented in ways easily understandable by the intended audience.

1.3 Research hypothesis and questions

I aimed to assess some of the more popular HS patient information resources with the following hypothesis and research questions.

1.3.1 Hypothesis

Current online HS patient information resources have poor readability and coherence, lack current and relevant information, and are presented in ways that much of the intended audience will have difficulty understanding.

1.3.2 Research questions

Based on the foregoing hypothesis, I aimed to address the following research questions:

- Research Question 1. To what extent are the readability levels of the popular online HS information resources in accordance with published guidelines on patient health information?
- Research Question 2. To what extent are the coherence levels of the popular online HS information resources appropriate for the intended audience?
• Research Question 3. Is the information on the top-ranking HS websites current and relevant to HS patients?
• Research Question 4. Does the presentation of HS-related content adhere to standard information design principles?

1.4 Scope of the dissertation
This dissertation reports on my analysis of some of the more popular online HS patient resources. The research focuses on the information readability and coherence and other information features such as overall website word count, average sentence length, use of polysyllabic words, use of pictures, images and videos.

The work is limited in that it only explores the text readability, coherence and some information design features of the websites. Other measures of comprehensibility are not examined. Furthermore, while five readability tools are used, only one coherence tool is applied, limiting the conclusions from this part of the research.

1.5 Layout of the dissertation
Chapter 2 describes the subject matter of this research. Chapter 3 reports the main findings from the literature review. Chapter 4 covers the methodology that I used in this research and the rationale for using this methodology. Chapter 5 presents the main results and Chapter 6 discusses these results. Finally, Chapter 7 reports the main conclusions, lists some recommendations and suggests possibilities for future research.
Chapter 2. Hidradenitis Suppurativa

In this chapter I describe the chronic skin condition HS. I begin by outlining what HS is and explain the most common system used to classify disease severity. I then summarise the characteristics of HS and the disease prevalence. Subsequently, I outline what is currently known about causes and risk factors. I then give an overview of the main HS treatments used currently. Finally, I report on some of the ways HS impacts on the patient.

2.1 Definition and diagnostic delay

HS is a chronic, relapsing and often debilitating skin disease. Clinically, HS is defined as “a chronic, inflammatory, recurrent, debilitating, skin disease of the hair follicle that usually presents after puberty with painful deep-seated, inflamed lesions in the apocrine gland-bearing areas of the body, most commonly, the axillary, inguinal, and anogenital regions” (Zouboulis et al 2015, p. 619). HS typically appears as painful, fluid-filled lesions that can affect the armpits, breasts (in females), buttocks, groin, genitals, and thighs, but can strike anywhere there is hair on the skin.

As HS primarily affects the intimate areas of the body, patients typically delay seeking medical help due to embarrassment and globally there is a reported average delay of over two years between disease onset and seeking medical help (Saunte et al 2015). The condition is not widely recognised in medicine and there is an average diagnostic delay of over seven years (Saunte et al 2015; Loget et al 2018). Furthermore, it is frequently misdiagnosed (Saunte and Jemec 2017). Micheletti (2015) reports that the diagnostic delay can have significant consequences for patient well-being.

2.2 Stages

Disease severity has been classified using various schemes (reviewed in Zouboulis et al 2015) with the Hurley classification system (Hurley 1989; Figure 1 and Table 1) being the most widely used in dermatology clinics. The Hurley system stratifies HS into three stages, I to III, reflecting the severity of the condition.
Stage I HS is most common, accounting for approximately three-quarters of patients, followed by stage II (approximately one-quarter) and stage III accounts for ~ 1% of patients (Scheinfeld 2013). The hallmark feature of stages II and III are sinus tracts, which are channels or tunnels just under the skin surface that can extend from the lesions to the skin surface, and which can discharge pus onto the surrounding skin.
Table 1 The Hurley hidradenitis suppurativa (HS) classification scheme

<table>
<thead>
<tr>
<th>HS stage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Abscess formation, single or multiple, without sinus tracts and cicatrisation.</td>
</tr>
<tr>
<td>II</td>
<td>Recurrent abscesses with tract formation and cicatrisation, single or multiple, widely separated lesions.</td>
</tr>
<tr>
<td>III</td>
<td>Diffuse or near-diffuse involvement, or multiple interconnected tracts and abscesses across the entire area.</td>
</tr>
</tbody>
</table>


The Hurley classification scheme is widely used in clinics to inform treatment strategies (Zouboulis et al 2015; Saunte and Jemec 2017). Different parts of the body can be at different Hurley disease stages where patients have lesions at more than one body location.

2.3 Epidemiology

There are conflicting reports in the literature regarding the prevalence of the disease in the general population, with some estimating HS prevalence at 0.05% (Cosmatos et al 2013) and others reporting HS affects more than 1% of the global population (Dufour et al 2014; Pascoe and Kimball 2014). Zoubolis et al (2015) recently reviewed the literature and estimate the true incidence of the disease to be 0.08%–0.20%.

HS tends to start in the second or third decades of life, but HS in children has also been reported (Liy-Wong et al 2015; Braunberger et al 2018), while HS is rare in the elderly. HS is significantly more common in females than in males, with a female to male ratio of at least 3:1 (Jemec 2012). Dufour et al (2014) report that males have a tendency for more severe forms of the condition. HS is reported to be more common in African-American populations than in Caucasians (Reeder et al 2014) but further research is needed to explore this.
2.4 Causes and risk factors
The exact causes of HS are still not known but due to the increase in HS research in recent years (Hessam et al 2017), much progress has been made in understanding the condition. HS is a disease of the hair follicle in which blockage of the hair follicle appears to be the primary event. Several factors discussed subsequently may act together to play considerable roles in HS development.

Genetic susceptibility appears to be a significant contributor to HS. Approximately one in three of those with HS have a family relative with the disease (Jemec 2012, Zouboulis et al 2015). Mutations in certain genes have been linked to the condition in some cases with a family history (Wang et al 2010; Pink et al 2012, 2013) and genetics is one area of ongoing HS research (Pink et al 2018).

An abnormal immune response is also speculated to play a role in HS development. Initial blockage of the hair follicle may result from deregulation of the local immune system (Hotz et al 2016) attracting immunoregulatory molecules and immune cells to the blocked site (van der Zee et al 2011; Kelly et al 2015). This blockage triggers an inflammation response, which is a characteristic of HS development. As inflammation progresses, tissue destruction releases follicular content into the surrounding tissues causing the inflammatory process to spread. As healing from the inflammatory process occurs, tunnelling and tissue scarring occur (Saunte and Jemec 2017). The efficacy of some immunosuppressive agents in treating HS supports this abnormal immunological response as playing a role in HS development.

The role of bacteria in HS development is controversial, but it is believed that bacteria play some role in disease development (Nikolakis et al 2015). Early HS lesions have normal bacterial flora for the skin region, suggesting that bacterial infection is secondary to the underlying inflammatory process and that the inflammation is not caused by an infection (Ring et al 2017).

Hormonal factors have also been identified as influencing HS, given the observed HS prevalence among women, the fact that HS typically occurs after puberty, and that the condition tends to disappear in many post-menopausal women (Harris et al 2016). Many females report that symptom intensity fluctuates with menstrual cycle and pregnancy and antiandrogen therapy has been successful at alleviating symptoms in some (Theut Riis et al 2016). Further research and larger studies are needed to evaluate antiandrogen therapy in HS.
Beyond the non-patient controllable biological factors, some patient lifestyle choices have been linked to HS. Smoking and obesity are widely considered to be significant risk factors in developing the disease, and diet has also been linked to HS. Being overweight and obese are linked to the likelihood of developing HS, with between 43% and 77% of those with HS classified as being overweight or clinically obese (Pink et al. 2018). Furthermore, Kromann et al. (2014) and Fabbrocini et al. (2016) report that being overweight or obese may affect disease progression and severity.

There are numerous studies linking HS to tobacco smoking and the relationship is well established (reviewed in Jemec 2012; Dufour et al. 2014; Zouboulis et al. 2015), with rates of current or former smokers in HS groups being reported to be as high as 90% (Happle and König 2011). In a recent study of almost 4 million tobacco smokers in the USA, Garg et al. (2018a) found the HS incidence to be doubled amongst tobacco smokers. Greater remission has been reported in those who quit smoking (Kromann et al. 2014), while smokers have a reported lower rate of response to certain HS pharmacological therapies than non-smokers (Denny and Anadkat 2017). In addition, Scheinfeld (2013) reports that tobacco smoking undermines HS lesion healing and significantly worsens HS surgical outcomes.

There have been a few reports linking diet to HS. Cannistrà et al. (2013) report a possible link between HS and brewer’s yeast in a small study group (12 participants) while Danby (2015) found a possible link to dairy foods in a group comprising 47 HS patients. However, as both reports consisted of small study groups, further research is needed to explore the potential connection between diet and HS. Brocard et al. (2007) reported some complete remissions (eight cases) and partial remissions (14 cases) in a small study group (22 participants) after three months’ dietary supplementation with zinc salts. However, there have not been any follow-up studies on this single report. Vitamin D levels have also been linked to HS (Guillet et al. 2013) but further work is needed on this potential link.

2.5 Treatments

There is no cure for HS and treatments aim to try and manage the condition and prevent progression to more severe forms of the disease. Furthermore, there are no gold standard treatments as what works well in one patient does not necessarily translate to others. Effective management of HS requires a combination of some or all of lifestyle modifications, medical
therapies, surgery and/or laser intervention, and psychological supports (Zouboulis et al 2015; Harris et al 2016; Saunte and Jemec 2017).

### 2.5.1 Lifestyle modifications

Lifestyle modifications that can help alleviate disease severity include weight reduction and smoking cessation and the general medical opinion is that smoking and being overweight or obese must be avoided for effective disease management (Scheinfeld 2013; Zouboulis et al 2015; Harris et al 2016). Smoking cessation is even more important when a patient is considering surgery because smoking undermines healing and significantly worsens surgical outcomes (Gill et al 2013). Scheinfeld (2013), reports that weight loss is the single biggest factor that can help reduce HS severity, based on the author’s experience of treating hundreds of HS patients over many years. Some recent case reports of disease remission after substantial weight loss would appear to support the link between bodyweight and HS (Kromann et al 2014; Thomas et al 2014; Boer 2016; Gallagher et al 2017). There is some evidence that dietary modifications can help alleviate disease severity, but further research is needed.

### 2.5.2 Medical management

Medical management depends on disease severity and various drug therapies are used, depending on the Hurley stage(s) (reviewed in Zouboulis et al 2015; Saunte and Jemec 2017).

- Antibiotics are the first-line therapy used in Hurley stage I–III HS and are applied to the affected areas on the skin as lotions in mild lesions or are taken as pills or administered intravenously in more severe lesions.
- Corticosteroids can be taken orally or injected directly into lesions to rapidly reduce inflammation associated with acute flares.
- Retinoids (vitamin-A derivatives) are taken orally for Hurley stage I–III HS.
- Hormones, especially antiandrogens in females, are a first-line therapy in those with Hurley stage I–II HS. Hormones can be administered orally or intravenously.
- Biologics, a class of modern drugs which suppress the immune system, are increasingly being used to treat stage II–III HS. Most biologics are administered by self-injection or intravenously.
In 2015 the European (European Medicines Agency 2018) and US regulatory authorities (US Food and Drug Administration 2018) both approved the biologic drug Humira® (adalimumab) for the treatment of mild to moderate forms of HS. Humira is the first and currently the only licensed therapy for HS in Europe and the USA. Other similar biologics drugs are currently being evaluated in clinical trials to assess their efficacy and safety as treatment therapies for HS (Saunte and Jemec 2017; Theu Riis et al 2018).

2.5.3 Surgery and/or laser intervention
For severe Hurley stage II–III forms of HS that have not responded to pharmacological therapies, surgery is advocated, and Saunte and Jemec 2017 and Scuderi et al (2018) review the various surgical procedures that are recommended depending on severity and body location(s) affected:

- Incision and drainage is a useful procedure when patients present extremely painful lesions; it is beneficial in relieving acute pain and can be conducted in a physician’s surgery but is discouraged for long-term practice because of the high recurrence rates.
- Deroofing is a simple and minimally invasive procedure whereby the “roof” of a HS lesion is removed surgically or by using a laser.
- Radical wide excision, the treatment of choice in severe stage III HS, is a procedure in which the affected skin and some surrounding unaffected skin is removed and usually requires extended hospitalisation and prolonged healing times.
- Laser therapy is a promising treatment increasingly being used to vaporise all affected skin leaving healthy skin in between.

However, surgery has been reported to have unsatisfactory outcomes, with results unacceptable to patients. Harris et al (2016) report that the most common complications of surgery are high disease recurrence rates, risk of surgical wound infections, significant scarring, and skin graft failure in the case of radical wide excision surgery.

2.5.4 Psychological support
Due to the various ways HS can affect the patient (discussed in section 2.6), treating physicians are increasingly recognising the importance of psychological supports in disease management.
(Saunte and Jemec 2017; Matusiak 2018). In addition, peer-to-peer patient support groups are advocated as important adjunct treatment measures in disease management (Esmann and Jemec 2011; Woodruff et al 2015; Saunte and Jemec 2017).

2.6 Impact
Living with HS can impact upon patients’ lives on several levels. Physically, the pain associated with typical HS lesions can be very severe, and the persistent pain is often cited by those with HS as the single most important factor impairing quality of life (Smith et al 2010; Ring et al 2016; Matusiak et al 2018). Studies have shown that the pain can negatively impact on the patients’ ability to work as HS typically affects individuals during their most productive years i.e., 18–55 years old (Zouboulis et al 2015). Matusiak et al (2010) and Kluger et al (2017) report that both unemployment and work absenteeism are significantly higher among those with HS compared to unaffected individuals.

Emotionally, active HS lesions can be associated with malodorous discharge that can stain clothing, causing social embarrassment and stigmatisation (Matusiak et al 2010; Esmann and Jemec 2011). HS can affect patients’ sense of self-worth and ability to form interpersonal and intimate relationships due to the ugly, foul-smelling physical symptoms and lack of control over the disease (Zouboulis et al 2015; Schneider-Burrus et al 2018). Consequently, social isolation is often adopted by those with HS as a coping mechanism (Dufour et al 2014), and depression and anxiety levels among HS patients are reportedly high (reviewed in Matusiak 2018). Furthermore, Thorlacius et al (2018) identified those with HS as being at a significantly increased risk of suicide compared with the general population. Tiri et al (2018) report that this elevated suicide risk is more pronounced in females.

As HS can affect the intimate parts of the body, many (approximately two-thirds) HS patients have a significantly higher impairment of sexual health when compared to controls (Kurek et al 2012; Janse et al 2017; Alavi et al 2018; Matusiak 2018; Slyper et al 2018). Considering the chronic pain and significant physical, emotional, psychological and sexual impact that HS can have, Garg et al (2018b) recently reported that those with HS are twice as likely of substance use disorder. Zouboulis et al (2015) conclude that HS is a highly distressing disease for many patients, one of the worst that has been analysed and evaluated in dermatology.
Given the many impacts that HS can have, there is clear motivation for those with HS, or with suspected HS, to seek information about their clinical symptoms and condition online, which provides the opportunity to obtain disease-related information safely, quickly and privately (Hessam et al 2017). In addition, much progress has been made in HS research and in understanding the condition in recent years and the impact that patient lifestyle changes can have on disease severity; it is important that this information is communicated to those living with the condition. Disseminating this information online affords a convenient and rapid means to do so. This research assesses the quality of the disease-related information presented on some of the popular HS websites. In the next chapter, I examine the relevant recent literature for studies that examine the quality of online health-related information.
Chapter 3. Literature review

There are multiple factors linked to HS, and recent research (outlined in section 2.4) indicates that certain patient-controllable lifestyle choices can affect disease severity, and this information needs to be communicated clearly to HS patients. Numerous factors can affect the comprehension of information, and this study examines some features affecting the understanding of online HS information. In this chapter I outline the theoretical background to this study on information comprehension and introduce two main measures of text comprehension: readability and coherence. I begin this chapter by introducing the theory of readability and readability formulas; I then present propositional idea analysis as a method to analyse text coherence and conclude this section with a brief description of some other website features that can affect comprehension. In the second section I outline how these analytical tools have been applied to patient information websites. I conclude the chapter by examining comprehension assessments of online HS resources.

3.1 Text comprehension

Three aspects of text, readability, coherence and text content analysis, that are used to measure text comprehension, are the focus of this section of the chapter.

3.1.1 Text readability

One measurement of text comprehension is its ‘readability’. Most readability definitions (reviewed in DuBay 2004), refer to the ease of comprehension of information. For the current study on health-related information I refer to the meaning used by Zhang et al (2015, p. 2076) who define readability as “whether content is understandable for general consumers without a medical background”. According to the OECD (2016), the average adult in OECD countries has a few years of secondary level education, which usually consists of some scientific instruction. HS patients come from all sections of society, and most will not have a medical background. Therefore, ‘the ease with which someone with limited scientific knowledge can read and understand health-related information’ is the working definition I use for readability in this study.

For the past ~100 years, educators, researchers and other professional communicators have investigated the readability of information (DuBay 2004). Numerous readability formulas
have been developed in that time which indicate the reading difficulty associated with written information. Readability scores typically indicate how many years of education are required to understand the analysed information and most readability formulas predict text complexity by analysing vocabulary using word length counts. Many of the ‘classic’ pre-1980s formulas have proven popular and have been widely used to assess textual information, while the ‘new’ post-1980s readability formulas, which consider recent developments in education, linguistics, cognitive science, psychology, and computer science research, are increasingly being used (Benjamin 2012).

Friedman and Hoffman-Goetz (2006) and Zamanian and Heydari (2012) have reviewed the many available readability formulas and concluded that readability formulas:

- Give the writer relevant information to reach their target audience;
- Can allow the reader to know ahead of time if they will understand the material;
- Are easy to use;
- Are easily applied by common software packages;
- Can help writers convert their material into plain language.

However, the main disadvantages of readability formulas are that they:

- Cannot tell how well the target audience understands the text;
- Can give wide variation in results when different formulas are applied to the same text;
- Do not measure the prior knowledge, interest level, concept difficulty, or text coherence;
- Cannot assess the impact of audio–visual aids in text.

While most available readability formulas have been widely used for written information, these formulas may not be equally useful in the context of online health information. Zhou et al (2017) recently showed that some readability tools may be inappropriate for short (< 500 words) pieces of text. Consequently, most readability assessments of online health information tend to use several formulas simultaneously to improve the range of results (Friedman and Hoffman-Goetz 2006; Beaunoyer et al 2017).

There is no singly-accepted readability formula and using more than one is widely recommended to improve result validity when assessing information (Friedman and Hoffman-Goetz 2006; Badarudeen and Sabharwal 2010; Beaunoyer et al 2017). While useful, readability
tools are limited in that they only assess certain aspects of text difficulty. Schriver (2017) highlights that readability formulas are deficient in that they fail to assess information design, the visual and verbal content that has been designed for clarity and accessibility. In addition, Schriver (2000) argues that readability formulas do not consider text design characteristics associated with clarity, such as the use of bulleted lists, titles, multiple levels of headings, tables, figures and illustrations (Schriver 2000). Furthermore, readability formulas do not assess typographic signalling, the use of white space and other design features (Redish 2000). Consequently, some (Smith et al 2011; Benjamin 2012) have called for additional assessment tools to be developed to gain a comprehensive understanding of text difficulty.

3.1.2 Text coherence

Text coherence is another factor affecting text comprehensibility. In this section I present some background information on text coherence and describe a tool to assess text coherence, propositional density analysis.

In linguistics, ‘coherence’ refers to the extent to which the ideas in a text are connected. Text coherence may be defined as the extent to which the relationships between the ideas in the text are made explicit (McNamara 2001). If a text is not coherent, inferences must be made to supply the missing cohesive links, leading to increased processing demands and comprehension difficulties (Irwin 1980). Coherence has been shown to be a major factor in text comprehension (McNamara et al 1996; McNamara and Kintsch 1996). Text coherence can be measured using propositional idea analysis, where a proposition is an idea unit; the content of a sentence that makes a claim and is capable of being true or false (Brown et al 2008). Models of comprehension (e.g. Kintsch 1998) assume that the proposition is the fundamental unit of comprehension. Brown et al (2008) define propositional density as the number of propositions divided by the number of words. Propositional density analysis examines text complexity, something which readability formulas do not address and has also been linked to text understanding and retention (DeFrancesco and Perkins 2012). The more propositions text has, the greater the processing difficulty, and consequently, the memory performance of the reader is reduced (Kintsch and Keenan 1973). While none of the available readability formulas examine propositional density, DeFrancesco and Perkins (2012) suggest that using propositional density analysis may be useful for cross validation studies with readability analysis.
3.1.3 Text content analysis

Comprehension of online information is dependent on factors other than readability and coherence. While a comprehensive review of all other factors affecting text comprehension is beyond the scope of this dissertation, I focus on some of the variables analysed later on in this study, namely overall word count, the use of polysyllabic words, average sentence length and the use of images, pictures and videos.

How online health information is presented can impact on comprehension. Too much text will not engage readers as readers prefer text that is brief and to the point (Office of Disease Prevention and Health Promotion 2018). The US Office of Disease Prevention and Health Promotion guidelines also recommend minimising the number of polysyllabic words used in the health information. Sentence length can also affect comprehension; the US Library of Medicine and National Institute of Health (2018) recommend limiting sentence length to 10–15 words per sentence when presenting health information. Some health literacy guidelines recommend that online health information is supported by the inclusion of simple, relevant images to improve understanding and engagement (Office of Disease Prevention and Health Promotion 2018). Online video content can also engage the reader and enhance the reader experience (World Wide Web Consortium 2018) and this enhancement is especially true in presenting health information when the video content features other patients talking about disease self-management techniques (Chapman et al 2017).

3.2 Assessing online patient education websites

With the large increase in the number of health-related patient education websites published over the last decade or so, many studies have assessed the readability of such resources. Walsh and Volsko (2008) assessed over 100 websites associated with the leading medical-related causes of mortality in the USA. Using three classic readability tools, they found that most of the sites exceeded the recommended reading grade levels and ranked as difficult to read on the USDHSS scale. Fitzsimmons et al (2010) applied two readability tools to 100 Parkinson’s disease websites and found none of the sites complied with readability guidelines. Cherla et al (2012) used three readability tools to examine 31 endoscopic surgery sites and found most of the sites were written above recommended levels. Eloy et al (2012) employed 10 assessment tools to study 262
otolaryngology sites and found that all sites exceeded the recommended readability levels. Sanghvi et al (2012) used four assessment tools to assess 41 patient facial fractures information sites and found only 5% of sites were written at the suggested level. Edmunds et al (2013; 2014) applied four readability formulas to >200 sites for three medical conditions and report that none of the sites had readability scores within published guidelines. This trend is not just confined to these conditions as similar findings have been reported across the spectrum of medical conditions. These include, among others, stroke (Sharma et al 2014), mammography (AlKhalili et al 2015), parathyroid surgery (Patel et al 2015), lymphedema (Seth et al 2016), ophthalmology (Williams et al 2016), and heart failure websites (Kher et al 2017).

Recent work has shown that despite longstanding recommendations to the contrary, this trend of presenting website information at levels beyond the comprehension of many seen over the past decade persists to the present. Akinleye et al (2017) examined 100 patient websites for the 10 hand conditions most commonly treated by hand surgeons using just one readability formula. They found that the most frequently accessed materials for common maladies of the hand exceed the recommended readability limits, and the average reading ability of most US adults. Kapoor et al (2017) examined 454 materials provided to patients by the American College of Cardiology and the American Heart Association and found that articles were written at a reading level markedly higher than the recommended levels. Manchaiah et al (2017) used three readability formulas to assess 134 tinnitus websites. They found that, on average, only people with at least 10–12 years of education could read and understand the information for tinnitus in websites. Almost all the websites exceeded the reading level recommended for health information. Stewart et al (2017) analysed the readability of text on 2,731 urogynaecologic patient information sites and found that most (85%) of the websites had information written above the US average reading ability. Akinleye et al (2018) examined 50 websites for five of the most commonly accessed arthroscopy-related online patient education materials using one commonly-used readability formula. They report that a minority (26%) of the websites were at or below the US national average reading level. Minoughan et al (2018) evaluated the readability of 114 sports injury and prevention patient education materials provided by the American Academy of Orthopaedic Surgeons using six readability formulas. They found that patient education materials were written at a readability level too high for most patients to understand. On average, patient materials were written at least 2.5 grade levels higher than national recommendations and
only a small minority (7%) of the 114 articles had readability scores in line with recommendations.

While propositional analysis has been used extensively in linguistics and neuroscience research, reports on its use in evaluating online patient educational materials are limited, in contrast to the numerous readability studies that have been done in recent years. Ta-Min et al (2007) used propositional analysis with three classic readability formulas to assess some cancer websites and concluded that propositional analysis is an independent marker of text difficulty for web-based information as low readability did not imply high coherence and vice versa. They concluded that both readability and coherence should be used in health information content analysis as both are important factors in comprehensibility. Tulsieram et al (2016) applied propositional analysis to measure the coherence of the Canadian Ministry for Health human papilloma virus patient websites and found that the information coherence levels were not appropriate for the intended audience. Their study also used two readability formulas and concluded that ~ 40% of the population may not be able to understand the information on the websites.

In terms of website content analysis, a few studies have examined word counts on patient websites with figures ranging from an average ~1,000 words from 50 websites about Graves’ disease and thyroid-associated ophthalmopathy described by Edmunds et al (2014), to ~1,500 words on 100 Parkinson’s disease patient websites, reported by Fitzsimmon et al (2010), to an average ~2,200 words on 50 dermatology sites in the Tulbert et al (2011) study. Some researchers have combined readability studies with text content analysis of patient information sites in recent years. Tulbert et al (2011) used standardised research tools to comparatively assess the readability, length in words and use of images of several online dermatological patient education materials and to identify the strengths and weaknesses of some widely used patient resources. They report that no single source of commonly used internet patient-education material demonstrates optimal features in readability, length, and presence of photographic illustrations. Chapman et al (2017) examined 49 osteoarthritis patient websites for readability and text content analysis and report that only five of the websites met the recommended reading level for health education literature. Half of the websites included informative images to support the written information, and only a small minority of the sites included relevant videos.
3.3 Online hidradenitis suppurativa resources

Just one study to date has examined online HS resources. Hessam et al (2017) assessed 39 popular HS websites using three readability formulas: the automated readability index (ARI), the Flesch Reading Ease score, and the Flesch–Kincaid Grade Level (FKGL); they rated most of the websites as difficult to read and found that, on average, seven to nine years of school education were needed to understand the content. However, Hessam et al did not use any of the readability formulas specifically recommended for assessing clinical information or any of the newer readability formulas, and one of the formulas used (FKGL) has been reported to significantly underestimate the reading difficulty of clinical information in other studies (Ta-Min et al, 2007; Sharma et al 2014). In addition, while Hessam et al (2017) assessed the HS websites’ readability and some other quality characteristics, they did not examine the information coherence.

In this study I build on this earlier work and evaluate the readability of the popular HS websites using both classic and newer readability formulas. I also assess the coherence using propositional analysis, which has not yet been used widely to assess online information on HS or other patient websites. Furthermore, I examine the content of the top-ranking HS web resources and assess their timeliness and usefulness. The next chapter describes the methodology used to analyse the information on the popular HS websites.
Chapter 4. Methodology

In this work, I examine if the readability levels of the top-ranking HS patient information sites are suitable for most HS patients (research question 1). In addition, I assess the coherence levels of the popular HS sites (research question 2). I also evaluate if the information on these patient websites is current and relevant (research question 3) and presented in a user-friendly way (research question 4) to the HS patient. In this chapter I report on the methodology that I used to address these research questions. I begin by outlining the research design and then describe the data acquisition, generation and processing tools that I used. Subsequently, I explain the information analysis tools, and conclude with a description of the statistical analyses that I performed.

4.1 Research design
I mainly assessed two aspects of information comprehensibility, readability and coherence, of some online patient resources for those living with the chronic skin condition HS. I also assessed some other general website features of the online HS information.

I assessed the readability of the HS websites using five popular formulas and compared these scores to recommended guidelines, to answer research question 1.

I evaluated the coherence of the HS websites using available propositional density analysis software. As there are currently no published guidelines on text coherence levels for patient education resources, I generated coherence data representative of the general lay population and compared this data to the mean coherence scores of the HS websites to answer research question 2.

I assessed the timeliness, and any mention of patient support groups to answer research question 3. I also examined several general features of each website to answer research question 4. Specifically, I analysed the HS-related content of each site for the use of pictures, visual aids and videos, and for the overall word count, average sentence length, and proportion of polysyllabic words and compared these to recommended figures.

4.2 Data gathering
I describe the data gathering instruments, and the rationale for using these tools in this section.
4.2.1 Online interest in hidradenitis suppurativa

To determine academic interest in HS, I used the PubMed database (PubMed 2018) and queried the database using the search term “hidradenitis suppurativa”, filtering the results using the date options. I used Google Trends (Google Trends 2018) from Google Inc. to evaluate online public interest in HS. I performed a Google Trends search using the search term “hidradenitis suppurativa” to examine online interest in HS during 02 June 2009–02 June 2018. Google Trends collects internet search query data over time and across regions and the search activity is expressed as relative search volume. The week in the chosen period with the highest overall number of search queries scores 100 points, and all other weeks are viewed relative to that week.

4.2.2 Google search strategy

I used Google.com for searching for HS information resources as Google is the most used internet search engine globally, with 73% of the markets across desktop/laptop, 93% of mobile, and 87% of tablet platforms (NetMarketShare 2018). I queried Google.com with the search term “hidradenitis suppurativa” on 02 June 2018. Results from the first five Google pages were included, as the first 50 search results cover ~99% of click-throughs of internet search engine users (Chitika 2013). Although known by other names (such as “acne inversa” and others), “hidradenitis suppurativa” is the most common term in English. I restricted the search to English-language sites as ~53%) of websites are published in English (W3Techs 2018).

I excluded websites from further analysis if they met any of the following ten criteria:

- pages that were scientific articles;
- video content only sites (e.g. YouTube sites);
- commercial company sites;
- social media sites;
- sites with content exclusively for medical professionals;
- sites that required subscriptions/sign-ups to read the content;
- pages that were links-only pages;
- wiki sites;
- news articles;
- pages that did not meet the word count requirements of the readability software used (ReadabilityFormulas.com 2018a).
4.2.3 Data processing
I copied the text from each website into Microsoft Office Word. I deleted text not directly related to patient education, and removed website links, quotations, numbers, decimals, bullets, abbreviations, paragraph breaks, colons, semicolons, and dashes within a sentence as including these can skew readability results (Friedman and Hoffman-Goetz 2006; Beaunoyer et al 2017).

4.3 Data analysis
I begin this section by outlining the data analysis tools I used to determine readability, and then describe the text coherence assessment tool used. Finally, I describe the statistical analysis tools I used to analyse the data.

4.3.1 Readability assessments
There is no gold standard readability formula and Friedman and Hoffman-Goetz (2006) and Beaunoyer et al (2017) recommend using more than one formula to assess any given piece of text. Beaunoyer et al (2017) also report that most readability studies on internet-based health information use several formulas simultaneously to broaden the range of results. I copied the processed text (section 4.2.3) into the following five readability formulas available online (ReadabilityFormulas 2018a):

- The ARI (Kincaid et al 1975) assesses characters, words and sentences and is derived from ratios that represent word and sentence difficulty;
- The FKGL (Kincaid et al 1975) analyses the number of syllables, words, and sentences;
- The Gunning FOG Index (GFI, Gunning 1952) examines average sentence length and number of hard words;
- The New Dale–Chall Index (NDC, Chall and Dale 1995) uses a list of thousands of words that most fourth-grade US students understand and considers words not on that list to be difficult;
- The Simple Measure of Gobbledygook (SMOG) formula (McLaughlin 1969) multiplies word length and sentence length. SMOG has been widely used for assessing health information and is recommended in healthcare education (Friedman and Hoffman-Goetz 2006).
I selected ARI and FKGL to enable comparisons with extant readability assessments of some HS websites (Hessam et al 2017). I chose GFI and SMOG as they have been widely used for assessing health information and are recommended in healthcare education (Friedman and Hoffman-Goetz 2006; Beaunoyer et al 2017). I picked NDC as it was developed primarily for the assessment of health education materials (Friedman and Hoffman-Goetz 2006). Beaunoyer et al (2017) report that these five formulas are also the most commonly used for assessing the readability of health-related information.

These readability formulas output scores corresponding to the number of years of US high school education needed to understand the content i.e., a score of 5 requires at least a fifth-grade US high school level of education (Table 2, ReadabilityFormulas 2018a). The USA is one of the few countries to have published such readability recommendations. Many researchers use these guidelines as a starting point and extrapolate to their own country’s education system.

<table>
<thead>
<tr>
<th>GFI/ FKGL/ SMOG/ARI</th>
<th>US Grade Level</th>
<th>USDHHS Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;12</td>
<td>College</td>
<td></td>
</tr>
<tr>
<td>10–12</td>
<td>10th–12th</td>
<td>Difficult</td>
</tr>
<tr>
<td>9–10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8–9</td>
<td>8th–9th</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7th</td>
<td>Average</td>
</tr>
<tr>
<td>6</td>
<td>6th</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5th</td>
<td>Easy</td>
</tr>
</tbody>
</table>

Table 2: Readability scores with their equivalent US grade level and ease of reading rating

Abbreviations: ARI = Automated Readability Index; FKGL = Flesch-Kincaid Grade Level; GFI= Gunning Fog Index; FRE= Flesch Reading Ease Score; SMOG= Simple Measure of Gobbledygook; USDHHS= United States Department of Health and Human Services

I compared the readability data to the USDHSS patient information readability guidelines and to the AMA guidelines. The AMA (Weis 2003) recommend that patient information should
be written at or below the sixth-grade and these have been linked with the USDHHS rating difficulty scale (Table 2). The USDHHS recommends that patient information should be classed as “easy” to read (The National Work Group on Literacy and Health 1998; USDHHS 2000). Information is categorised as ‘easy’ if written at or below the sixth-grade level, ‘average’ if between seventh- and ninth-grade, and ‘difficult’ if higher. The USDHSS scale has been correlated with ARI, FKGL, FOG, and SMOG outputs (Table 2).

I tabulated the ARI, FKGL, GFI and SMOG outputs to enable comparisons between each formula. As NDC uses its own unique readability scale, I displayed this data graphically. The NDC scores have also been correlated with the US grade level (Table 3).

### Table 3: New Dale–Chall scores with their equivalent US grade level and ease of reading rating

<table>
<thead>
<tr>
<th>Final NDC Score</th>
<th>Grade Level</th>
<th>USDHSS Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>10+</td>
<td>16+ (college graduate)</td>
<td></td>
</tr>
<tr>
<td>9.0–9.9</td>
<td>13–15 (college)</td>
<td></td>
</tr>
<tr>
<td>8.0–8.9</td>
<td>11–12</td>
<td>Difficult</td>
</tr>
<tr>
<td>7.0–7.9</td>
<td>9–10</td>
<td></td>
</tr>
<tr>
<td>6.0–6.9</td>
<td>7–8</td>
<td>Average</td>
</tr>
<tr>
<td>5.0–5.9</td>
<td>5–6</td>
<td></td>
</tr>
<tr>
<td>≤ 4.9</td>
<td>4 and below</td>
<td>Easy</td>
</tr>
</tbody>
</table>

Abbreviations: NDC= New Dale–Chall; USDHHS= United States Department of Health and Human Services

¹= from ReadabilityFormulas (2018b)

4.3.2 Coherence assessments

I assessed the propositional density scores of each HS website with computerized propositional idea density rater software (CPIDR; Brown et al 2008). Text content was copied into CPIDR and the software outputs scores between 0 and 1; the closer the score is to 0, the more coherent the text is and the closer the score is to 1, the less coherent it is.
According to the Programme for the International Assessment of Adult Competencies, the average adult literacy level among 33 OECD countries is level 2 on the Programme for the International Assessment of Adult Competencies scale, corresponding to some secondary level education (OECD 2016). The Irish junior cycle is equivalent to one to three years of secondary education. I chose text excerpts from two junior cycle science books, Exploring Science (O’Callaghan et al 2016) and Investigating Science (Comiskey et al 2016) to generate text representative of the online lay population and applied CPIDR. I chose two science textbooks because scientific information is presented and organised like many health information websites.

4.3.3 Website content analysis

For each HS website that met the criteria for further analysis (section 4.2.2), I analysed the website text content for overall wordcount, average sentence length and percentage of polysyllabic words (defined as having three or more syllables) per sentence.

I noted the use of HS-specific illustrations, photographs, pictures and the use of informational videos on each site. I also searched the websites for mention of the Hurley HS severity classification scheme (section 2.2), and specifically for photos or clinical images of the three Hurley stages.

In 2015, adalimumab (tradename Humira®) became the first and currently the only approved medicine for HS within the Europe Union and USA (European Medicines Agency, 2018; US Food and Drug Administration, 2018). To assess how current the information was on the websites, I searched for mentions of “adalimumab” or “Humira” on the websites. Other similar drugs, classified as “biologics” are currently in development for HS (Saunte and Jemec, 2017; Theut Riis et al 2018). I also searched the websites for mention of “biologics” to assess the timeliness of the websites.

In recent years numerous online patient support groups have been established and research has shown that there are benefits of such groups, such as improved health outcomes and a reduction in patient feelings of loneliness, particularly for those living with chronic illness (Wright 2016). As an indicator of the usefulness of each web resource to the typical HS patient, I searched each site for mention of HS patient support groups, which are increasingly being considered as important measures in disease management (Esmann and Jemec 2011; Woodruff et
al 2015; Saunte and Jemec 2017). I also checked to see if appropriate functioning hyperlinks were supplied to any of the support groups that were mentioned.

4.3.4 Statistical analysis

I used single factor analysis of variance (ANOVA) to compare readability scores from each formula to determine if there were significant differences between the figures generated from the readability formulas. Further post-hoc testing was performed with the Tukey–Kramer test and the Student’s t-test. I used the Tukey–Kramer test to answer research question 1 and the Student’s t-test to answer research question 2. A $P < 0.1$ significance level is recommended for technical communications studies (Hughes and Hayhoe, 2008), but as most of the reports consulted for the current study used a $P < 0.05$ level, in all statistical tests $P$ values $< 0.05$ were considered significant. I calculated the Pearson correlation coefficients ($r$) to determine if the readability and coherence scores had a mutual association. Values of $r \leq -0.7$ or $r \geq 0.7$ were considered as strong correlations. All statistical analyses were performed with Microsoft Office Excel. In the next chapter I present the main findings of this research.
Chapter 5. Results

While there have been many studies assessing the readability of online health information resources, few have focused on online HS resources. In this chapter I present the results of my analysis of online HS resources. I begin with an analysis of the academic interest in HS and look at the online public interest in the disease. I then present the findings from my evaluations of the most popular HS online resources. Subsequently, I describe the readability score assessments of the most popular HS websites. Then, I explore the coherence of the top-ranking HS resources and conclude with an examination of potential associations between readability and coherence scores.

5.1 Academic interest in hidradenitis suppurativa

There has been an increase in the number of HS papers published in recent years. To quantify this increase, I performed a search of the PubMed database and the results are shown in Figure 2. This figure illustrates how dramatically academic interest in HS has increased over the past eight decades. More papers have been published on HS in the seven years from 2011 to 2017 (975 publications), than in the preceding 70 years (720 publications).

Figure 2: Hidradenitis suppurativa (HS) papers by decade. Chart showing the increase in HS research over the past ~80 years, measured by the number of papers published on HS in the PubMed database as a function of the range in years over the past eight decades. For the last (incomplete) decade shown (2011–2017), publications up to December 2017 were included.
Hessam et al (2017) reported an increase in academic interest in HS during 2006–2015. To determine if this reported trend continued in the intervening years, I performed a search of the PubMed database spanning the January 2015 to June 2018 period (Figure 3). The figure illustrates that the increase in academic interest in HS has continued in the years since 2015, with an average annual growth rate in HS publications of ~35%.

Number of HS papers published since 2015

![Chart showing the increase in HS research over the past few years, measured by the number of papers published in PubMed on HS as a function of years. Data were extrapolated for 2018 based on the number of HS papers published up to 02 June 2018.]

**Figure 3: Hidradenitis suppurativa (HS) papers since 2015.** Chart showing the increase in HS research over the past few years, measured by the number of papers published in PubMed on HS as a function of years. Data were extrapolated for 2018 based on the number of HS papers published up to 02 June 2018.

### 5.2 Online interest in hidradenitis suppurativa

A Google.com search conducted on 02 June 2018 for “hidradenitis suppurativa” yielded over a million results. Only the first 50 results (or five pages) were analysed further. The complete list of the 50 sites retrieved are shown in Appendix 1. Of these 50 sites, 29 met the exclusion criteria described in section 4.2.2, comprising:

- scientific articles (14/29 sites),
- sites without enough content to meet the requirements of the readability software (4/29),
- video only content (2/29),
- news articles (2/29),
- social media sites (2/29),
- commercial company (1/29),
- links only sites (1/29),
- wiki sites (1/29),
- sites requiring a subscription to access content (1/29), sites with content exclusively for medical professionals (1/29).

The 21 remaining sites were from countries or regions with English as a primary or official language as follows:
- the USA (12/21),
- the UK (5/21),
- Australia (1/21),
- Canada (1/21),
- Ireland (1/21),
- New Zealand (1/21),

To determine if there has been a global increase in online HS activity in recent years, I conducted a Google Trends search, and the results are displayed in Figure 4.

![Global Google searches for "hidradenitis suppurativa"](chart)

**Figure 4: Online interest in hidradenitis suppurativa (HS) since 2016.** Chart showing the international public interest in HS, as measured by the number of Google searches using the search term “hidradenitis suppurativa”, from 01 March 2016 to 02 June 2018. The data points on the trend line indicate the relative search volumes at the start, middle and end of the period analysed.
Figure 4 illustrates that online interest in HS in the parts of the world that use “hidradenitis suppurativa” to identify the disease (i.e. most English-speaking countries) has continued to increase from the beginning of 2016 (relative search volume of 51) through to June 2018 (relative search volume of 77). Within this period there was a ~1.5-fold increase in the relative search volume for HS using Google.

5.3 Readability assessments of the top hidradenitis suppurativa online resources

I initially used the NDC formula to assess the readability of the HS websites that met the inclusion criteria for this study and the results are shown in Figure 5. This graph presents the HS sites and their NDC readability scores, with their corresponding US grade levels and USDHSS ratings. Most (90%, 19/21) of the HS web resources scored grade 9–10 or above and ranked as ‘difficult’ to read on the USDHSS scale (cf. Table 3). NDC analysis also revealed that 10% (2/21) of the sites require a college graduate level education to comprehend the information. Only 10% of the websites were rated as USDHSS ‘average’ reading difficulty. None of the sites scored grade 6 or lower, the USDHSS recommended grade level for patient education materials.

---

**Figure 5. New Dale–Chall (NDC) formula readability scores for hidradenitis suppurative (HS) websites.** AMA= American Medical Association; US= United States; USDHSS = United States Department of Health and Human Services.
The ARI, FKGL, GFI and SMOG readability assessments of the top online HS resources that met the inclusion criteria for this study are displayed in Table 4, which lists the HS websites in descending order of popularity, as measured by their Google ranking (cf. Appendix 1). Shown also in Table 4 are the readability scores of each HS website when analysed with the ARI, FKGL, GFI and SMOG formulas, along with the mean readability score for each HS website. Notably, none of the 21 websites assessed in this study ranked as ‘easy’ to read on the USDHSS scale (equivalent to a grade level score of ≤ 6; cf. Table 2), the recommended reading level for patient education materials. Just 14% (3/21) of HS sites have average readability scores of ≤9, equivalent to an ‘average’ USDHSS reading rating. The remaining 86% (19/22) of websites ranked as ‘difficult’ to read on the USDHSS scale; of these, seven sites scored ≥12, requiring a college level education to comprehend the content. The average readability from the four formulas was 10.97, meaning that a grade 10–11 education level (equivalent to a USDHSS rating of ‘difficult’) is required to read the average online HS resource. These figures exceed the AMA recommended grade 6 or lower level (USDHSS ‘easy’ to read rating) recommended for online patient education materials.
Table 4: Readability scores of the hidradenitis suppurativa websites analysed in this study. The websites are listed in descending order of popularity, as measured by their Google ranking (cf. Appendix 1).

<table>
<thead>
<tr>
<th>Website provider/location</th>
<th>Readability Scores</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARI</td>
<td>FKGL</td>
</tr>
<tr>
<td>Mayo Clinic, USA</td>
<td>9.9</td>
<td>10.2</td>
</tr>
<tr>
<td>Irish Skin Foundation</td>
<td>9.2</td>
<td>9.7</td>
</tr>
<tr>
<td>American Association of Dermatology</td>
<td>7.8</td>
<td>7.8</td>
</tr>
<tr>
<td>DermNet New Zealand</td>
<td>13.1</td>
<td>14.1</td>
</tr>
<tr>
<td>WebMD, USA</td>
<td>5.6</td>
<td>5.8</td>
</tr>
<tr>
<td>National Health Service, UK</td>
<td>12.6</td>
<td>12.9</td>
</tr>
<tr>
<td>Healthline, USA</td>
<td>13.0</td>
<td>12.5</td>
</tr>
<tr>
<td>HSOOnline Australia</td>
<td>11.5</td>
<td>11.9</td>
</tr>
<tr>
<td>Canadian Dermatology Association</td>
<td>10.4</td>
<td>10.5</td>
</tr>
<tr>
<td>Patient Info, UK</td>
<td>8.8</td>
<td>9.0</td>
</tr>
<tr>
<td>Rare Diseases, USA</td>
<td>10.7</td>
<td>10.9</td>
</tr>
<tr>
<td>University of Rochester Medical Center, USA</td>
<td>9.0</td>
<td>9.2</td>
</tr>
<tr>
<td>NewsMedical. Net, UK</td>
<td>12.9</td>
<td>13.0</td>
</tr>
<tr>
<td>No BS About HS, USA</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>British Skin Foundation</td>
<td>9.0</td>
<td>9.7</td>
</tr>
<tr>
<td>MSD Manual, USA</td>
<td>12.3</td>
<td>12.4</td>
</tr>
<tr>
<td>Hidradenitis Suppurativa Foundation, USA</td>
<td>13.4</td>
<td>13.9</td>
</tr>
<tr>
<td>SkinSight, USA</td>
<td>10.9</td>
<td>11.1</td>
</tr>
<tr>
<td>Medical News Today, USA</td>
<td>8.4</td>
<td>9.4</td>
</tr>
<tr>
<td>British Association of Dermatologists</td>
<td>11.3</td>
<td>11.6</td>
</tr>
<tr>
<td>Florida Westcoast Skin &amp; Cancer Center</td>
<td>11.3</td>
<td>11.3</td>
</tr>
</tbody>
</table>

Abbreviations used: ARI = Automated Readability Index; FKGL = Flesch–Kincaid Grade Level; GFI= Gunning Fog Index; NDC= New Dale-Chall; $P =$ probability; $r =$ correlation coefficient; SMOG= Simple Measure of Gobbledygook.
I used ANOVA to test if there was any significant between-group differences in the readability scores generated using ARI, FKGL, GFI and SMOG. ANOVA indicated that there was at least one formula generating scores significantly different to the others ($P < 0.05$, see Appendix 2). Further post-hoc testing using the Tukey–Kramer test indicated that the GFI formula readability outputs were significantly different from ARI, FKGL, and SMOG scores (all $P < 0.05$, Appendix 2, 3). On average, GFI overestimated readability by at least two grade levels. There were no significant differences between the ARI, FKGL and SMOG readability scores (Appendix 3).

5.4 Coherence assessment of the websites’ content

I used CPIDR software to assess the coherence of the top-ranking HS websites by comparing them to excerpts from two Irish secondary level education science textbooks, reflective of the OECD average adult literacy, and the results are shown in Table 5. The coherence scores ranged from 0.4653 to 0.5660 with a mean coherence score for all the websites of 0.4991. Most (86%, 18/21) of the sites had coherence scores that were not significantly different to the samples from the science textbooks ($P > 0.05$, Table 5) Just 14% (3/21) of the sites analysed had coherence scores significantly different to the science textbook samples representative of the lay population ($P < 0.05$, Table 5). Therefore, it can be inferred that most of the popular HS websites have coherence scores appropriate for the lay population.
Table 5: Coherence scores of the hidradenitis suppurativa websites analysed in this study (after applying the exclusion criteria in section 4.2.2). The websites are listed in descending order of popularity, as measured by their Google ranking (cf. Appendix 1).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mayo Clinic</td>
<td>0.5027</td>
<td>0.1182</td>
<td>0.2491</td>
</tr>
<tr>
<td>Irish Skin Foundation</td>
<td>0.4737</td>
<td>0.4648</td>
<td>0.3188</td>
</tr>
<tr>
<td>American Association of Dermatology</td>
<td>0.4960</td>
<td>0.1121</td>
<td>0.2587</td>
</tr>
<tr>
<td>DermNet</td>
<td>0.4727</td>
<td>0.4719</td>
<td>0.2482</td>
</tr>
<tr>
<td>WebMD</td>
<td>0.5413</td>
<td>0.0734</td>
<td>0.1068</td>
</tr>
<tr>
<td>National Health Service</td>
<td>0.4657</td>
<td>0.4278</td>
<td>0.2267</td>
</tr>
<tr>
<td>Healthline</td>
<td>0.5227</td>
<td>0.1581</td>
<td>0.2420</td>
</tr>
<tr>
<td>HSOnline</td>
<td>0.5660</td>
<td><strong>0.0042</strong></td>
<td><strong>0.0079</strong></td>
</tr>
<tr>
<td>Canadian Dermatology Association</td>
<td>0.4820</td>
<td>0.3618</td>
<td>0.4284</td>
</tr>
<tr>
<td>Patient Info</td>
<td>0.4933</td>
<td>0.2477</td>
<td>0.4199</td>
</tr>
<tr>
<td>Rare Diseases</td>
<td>0.4907</td>
<td>0.1615</td>
<td>0.3979</td>
</tr>
<tr>
<td>University of Rochester Medical Center</td>
<td>0.5467</td>
<td><strong>0.0132</strong></td>
<td><strong>0.0255</strong></td>
</tr>
<tr>
<td>New Medical Net</td>
<td>0.4793</td>
<td>0.3404</td>
<td>0.3155</td>
</tr>
<tr>
<td>NoBSAboutHS</td>
<td>0.5020</td>
<td>0.1654</td>
<td>0.3031</td>
</tr>
<tr>
<td>British Skin Foundation</td>
<td>0.5017</td>
<td>0.0952</td>
<td>0.2161</td>
</tr>
<tr>
<td>MSD Manual</td>
<td>0.4673</td>
<td>0.4459</td>
<td>0.2224</td>
</tr>
<tr>
<td>Hidradenitis Suppurativa Foundation</td>
<td>0.4647</td>
<td>0.4339</td>
<td>0.2722</td>
</tr>
<tr>
<td>Skinsight</td>
<td>0.4963</td>
<td>0.1523</td>
<td>0.3298</td>
</tr>
<tr>
<td>Medical News Today</td>
<td>0.5530</td>
<td><strong>0.0195</strong></td>
<td><strong>0.0381</strong></td>
</tr>
<tr>
<td>British Association of Dermatologists</td>
<td>0.4653</td>
<td>0.4032</td>
<td>0.1550</td>
</tr>
<tr>
<td>Florida Westcoast Skin &amp; Cancer Center</td>
<td>0.4993</td>
<td>0.1866</td>
<td>0.3368</td>
</tr>
</tbody>
</table>

The Student’s t-test was applied to the mean coherence scores of the websites and the mean coherence scores of text excerpts from two Irish junior cycle textbooks, Comiskey et al (2016) and O’Callaghan et al (2016). Significant P values (< 0.05) are highlighted in bold.
I performed correlation analysis between the CPIDR scores and each of the five readability formulas used in this study to determine if there was any correlation between text coherence and readability and the results are displayed in Table 6.

**Table 6: Associations between readability formula scores and proposition density.**

<table>
<thead>
<tr>
<th>Readability formula</th>
<th>ARI</th>
<th>FKGL</th>
<th>GFI</th>
<th>NDC</th>
<th>SMOG</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>0.4604</td>
<td>0.4550</td>
<td>0.4456</td>
<td>0.3234</td>
<td>0.4109</td>
</tr>
<tr>
<td>$P$</td>
<td>0.0357</td>
<td>0.0382</td>
<td>0.0429</td>
<td>0.1527</td>
<td>0.0642</td>
</tr>
</tbody>
</table>

Abbreviations used: ARI = Automated Readability Index; FKGL = Flesch–Kincaid Grade Level; GFI = Gunning Fog Index; NDC = New Dale-Chall; $P$ = probability; $r$ = correlation coefficient; SMOG = Simple Measure of Gobbledygook.

I did not find any evidence of significantly strong correlations ($r \leq -0.7$ or $r \geq 0.7$) between the CPIDR scores and any of the five readability formulas used in this study. A weak to moderate correlation was detected between proposition density and ARI, FKGL and GFI, indicating that readability and coherence are independent measures of comprehension.

### 5.5 Content analysis of websites

I analysed the content of the HS websites to determine their overall word count, average sentence length and percentage of polysyllabic words per sentence. I also evaluated the use of HS-specific illustrations, pictures, photographs and videos and furthermore examined the timeliness of the websites. The results are shown in Table 7.

There was an 8.55-fold difference in the overall HS-related information word count of the websites, with an average of 1,283 words per website. It would take the reader 2–3 times longer to read the longest articles than the average article. The average sentence length in the HS-specific information on the websites was ~16 words. The average percentage of polysyllabic words per website was 18%.
Table 7: Content analysis of the top Google-ranking hidradenitis suppurativa (HS) websites analysed in this study (after applying the exclusion criteria in Section 4.2.2). The websites are listed in descending order of popularity, as measured by their Google ranking (cf. Appendix 1).

<table>
<thead>
<tr>
<th>Website</th>
<th>HS-specific word count</th>
<th>Average sentence length</th>
<th>Polysyllabic words / sentence(%)</th>
<th>HS photos / illustrations</th>
<th>Hurley stage photos</th>
<th>Video content</th>
<th>Biologics mentioned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mayo Clinic</td>
<td>726</td>
<td>12</td>
<td>19</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Irish Skin Foundation</td>
<td>260</td>
<td>16</td>
<td>18</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>American Association of Dermatology</td>
<td>2222</td>
<td>13</td>
<td>12</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>DermNet New Zealand</td>
<td>790</td>
<td>10</td>
<td>33</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>WebMD</td>
<td>1675</td>
<td>12</td>
<td>9</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>National Health Service</td>
<td>1120</td>
<td>18</td>
<td>21</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Healthline</td>
<td>1101</td>
<td>18</td>
<td>19</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>HSOnline Australia</td>
<td>2819</td>
<td>20</td>
<td>18</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Canadian Association of Dermatology</td>
<td>1094</td>
<td>18</td>
<td>16</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Patient Info</td>
<td>2820</td>
<td>13</td>
<td>17</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Rare Diseases</td>
<td>819</td>
<td>15</td>
<td>21</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>UMRC Health Encyclopaedia</td>
<td>766</td>
<td>14</td>
<td>17</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>NewsMedical. Net</td>
<td>559</td>
<td>22</td>
<td>18</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>No BS About HS</td>
<td>3354</td>
<td>14</td>
<td>11</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>British Skin Foundation</td>
<td>431</td>
<td>14</td>
<td>16</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>MSD Manual</td>
<td>585</td>
<td>20</td>
<td>18</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Hidradenitis Suppurativa</td>
<td>2606</td>
<td>21</td>
<td>22</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>SkinSight</td>
<td>631</td>
<td>16</td>
<td>17</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Medical News Today</td>
<td>392</td>
<td>15</td>
<td>15</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>British Association of Dermatologists</td>
<td>1142</td>
<td>18</td>
<td>19</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Florida Westcoast Skin &amp; Cancer Center</td>
<td>1032</td>
<td>16</td>
<td>20</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>
Two-thirds (14/21) of the websites featured a HS-specific illustration, picture or photograph. Only 24% (5/21) of the websites contained photographs of the Hurley stages of HS used to classify patients. Just 14% (3/21) of the top HS websites featured video content. In each of these, videos featuring HS patients talking about disease self-management were included on the sites. Furthermore, 52% (11/21) websites mentioned the use of modern biologics drugs to treat HS. Just a single website out of the 21 analysed (5%) featured HS-specific photographs, pictures, illustrations or informational videos, included images of the Hurley stages of the disease, and referred to biologics as treatment options. Almost one-quarter (24%, 5/21) of the websites failed to feature any photographs, pictures, illustrations or videos, or mention biologics (Table 7).

To determine the relevance of the websites to the typical HS patient, I analysed each site for the mention of and links to HS peer support groups. Table 7 shows that just 24% (5/21) of the HS websites referred to support groups. Of these, only three provided links to support groups. In the next chapter, I discuss the results presented here in the context of other findings in the literature.
Chapter 6. Discussion

In this chapter I discuss the relevance of the results generated during this study. I begin with an analysis of the online interest in HS, focusing on academic and public interest. I then scrutinise the readability score assessments of the top-ranking HS websites. Subsequently, I explore the coherence of the popular HS resources and look at potential associations between readability and coherence scores. Then I describe some additional content analysis of the HS websites, and list some of the limitations of this work. I conclude with a summary of this research.

6.1 Online interest in hidradenitis suppurativa

Hessam et al (2017) reported an increase in academic interest in HS during 2006–2015. The PubMed database analysis conducted in the current study indicates that this growth has continued in the intervening years with the number of HS papers almost doubling during 2016–2018, representing an average annual growth rate of ~35%. This rate exceeds the general growth rate in scientific publications, which is 8%–9% per annum, equating to a doubling approximately every nine years (Bornmann and Mutz, 2014). The HS publications’ growth rate also greatly exceeds the general growth rate in academic publishing, which is ~3.5% per annum (Forgues and Liarte 2013).

Mirroring the increased academic interest in HS, the Google search for “hidradenitis suppurativa” that I conducted in June 2018 yielded over a million results, almost twice the figure reported by Hessam et al (2017) who performed their Google search in March 2016. Notably, only five of the top ten ranking HS websites identified by Hessam et al (2017) also ranked in the top ten popular sites in the current study, illustrating how quickly the HS digital landscape is changing. The growth in HS digital content reflects a general growth trend on the internet, where the total number of websites has approximately doubled since 2016 (Internet Live Stats, 2018). Using Google Trends, I also found an increase in online search activity for “hidradenitis suppurativa” among the public, with HS-related online activity increasing ~1.5-fold in the same period. These increases in online HS-related search activity may be due to the increasing numbers of people worldwide with internet access. At the beginning of 2016, there were approximately 3.4 billion internet users worldwide. By the end of 2017, there were approximately 4.2 billion users (Internet World Stats, 2018). A further explanation for the rise in
HS search activity is that over the past decade or so, there has been an increase in the incidence of HS (Garg et al. 2017). In addition, there have been some recent online campaigns to raise HS awareness in English-speaking countries that use the term “hidradenitis suppurativa” such as the USA (Hope for HS, 2017), Canada (Canadian Hidradenitis Suppurativa Foundation 2018), the UK (Hidradenitis Suppurativa Trust, 2018) and Ireland (Irish Skin Foundation, 2018), which may have prompted this increased online HS search activity.

The increase in academic interest in HS is a welcome development that will likely lead to enhanced understanding of the condition, and ultimately to more effective HS treatments and therapies. The increased online interest in HS will undoubtedly also help raise disease awareness and understanding among physicians, patients and the public. The increased online interest also offers a great opportunity to educate those with HS or suspected HS so that they can have informed consultations with their physician. As patient self-management is advocated by many physicians as an adjunct treatment strategy for HS (Zouboulis et al. 2015; Jemec 2017), patients will only benefit from these online materials if the content is easily read and understandable and presented in accordance with relevant guidelines and recommendations. Subsequent sections examine some of these features on the popular online HS resources in further detail.

6.2 Readability assessments of online hidradenitis suppurativa resources

This research searched for the 50 top-ranking HS digital resources determined with the popular search engine Google. Of these 50 websites, 29 were excluded from further analysis. Almost half (14/29) of these resources were academic research papers, which I excluded as recent research has shown that the readability of scientific publications has decreased over time (Plavén-Sigray et al. 2017). This decreased readability is indicative of a growing use of general scientific jargon in publications, with ~22% of publications calculated to have a readability beyond college graduate level English, thereby severely limiting their accessibility to the lay population (Plavén-Sigray et al. 2017).

The USDHSS and AMA guidelines recommend that patient education materials should be written at an easy-to-read level, equivalent to at or below the sixth-grade education level. My analysis revealed that all 21 HS websites analysed in the current study have readability scores above the recommended guidelines, regardless of what readability formula was used. The average HS website readability score was 10.97 (when using ARI, FKGL, GFI and SMOG),
several grade levels above the recommended AMA level. Most (90%) of the HS websites scored a grade 9–10 or above and were ‘difficult’ to read on the USDHSS scale when using the NDC formula. While the readability assessments agreed in terms of most HS websites being at a ‘difficult’ USDHSS reading level, there was some variation in readability across the five formulas used in this study. For instance, ARI, FKGL, GFI and SMOG rated 38% (8/21) sites as requiring college level education, while NDC scored just 10% (2/21) of the websites in the same category. The ARI, FKGL, GFI and SMOG formulas do not incorporate word familiarity along with word length in determining readability. Thus, long words common on many HS websites such as ‘hidradenitis’, ‘abscesses’ and ‘inflammation’, which would be familiar to most people with HS, may have artificially increased the readability score of a website. The NDC formula addresses this familiarity with longer words which can skew readability results in other formulas (Misra et al 2013). NDC uses sentence length and number of familiar words in calculating readability, where familiar words are 3000 common words that have been deemed to be comprehensible to most fourth-grade students (Chall and Dale 1995).

I found that the GFI formula significantly overestimated the reading difficulty of the popular HS web resources relative to the other four readability formulas used. Others have also report GFI producing noticeable overestimates of reading difficulty relative to other commonly used formulas (Edmunds et al 2013; Patel et al 2013, 2015; Mehta et al 2018), although it is not known if the differences were significant as statistical testing was not applied in these studies. GFI calculates readability based on two characteristics of the given text: the average sentence length and the number of polysyllabic (i.e., containing three or more syllables) words. With regards to polysyllabic content, Friedman and Hoffman-Goetz (2006) contend that GFI does not account for variation in difficulty level across polysyllabic words and that GFI may overestimate the reading level required to comprehend some words that are lengthy but commonly understood. In the context of the current study, the 10-syllable “hi-dra-den-i-tis supp-ur-a-ti-va” likely inflates the GFI scores. As most people with a HS diagnosis will at least be familiar with the term “hidradenitis suppurativa”, which appears frequently on all the analysed HS websites, GFI will likely overestimate the reading difficulty. Furthermore, GFI is the only formula used in this study to calculate the text complexity using the polysyllabic word percentage, rather than word count, potentially making it more sensitive to texts with higher proportions of complex words (Cook et al 2017). The results of the current study highlight the importance of using more than
one formula when assessing readability (such as relying on GFI alone), as widely recommended (DuBay, 2004; Friedman and Hoffman-Goetz, 2006; Beaunoyer et al 2017).

Hessam et al (2017) analysed 39 of the top-ranking HS websites in a 2016 search with three readability formulas and most of the resources analysed in that study also had readability scores above the recommended USDHSS and AMA levels. However, as the authors did not quantify the proportion of resources that scored above the recommended levels, a direct comparison between their results and the current study could not be done. Hessam et al (2017) also found that, on average, seven to nine years of education were needed to comprehend the content on the HS websites. The current study found an average grade 10–11 level score for the HS websites (which equates to 10–11 years’ education), indicating that the readability of the newer HS digital content is more difficult to understand than pre-existing content. In a wider context, my readability analysis is consistent with recent reports on online patient education materials being written beyond the comprehension of the intended audiences (as detailed in section 3.2). Within the dermatology discipline, an assessment of online dermatological patient resources by Tulbert et al (2011) found that the readabilities of several major dermatology websites were above the recommended reading grade levels. Following the Tulbert et al (2011) report, John et al (2016) reviewed the readability of 260 dermatological patient education materials using 10 readability formulas and found the resources were written at a mean grade level of 11.13, with 66% of articles written above a tenth-grade level and no articles written at or below the AMA recommended sixth-grade level. Taking a broader view encompassing various medical disciplines, Daraz et al (2018) performed a systematic review to evaluate the readability of online health information in the USA and Canada. They included 157 cross-sectional studies evaluating 7,891 websites using 13 readability formulas and found that the mean readability grade level across websites ranged from grade 10 to 15 based on the different scales. Stratification by specialty, health condition, and type of organisation producing information revealed the same findings. Daraz et al (2018) concluded that online health information in North America has a readability level that is inappropriate for public use.

To improve the readability and comprehension of patient education materials, the US Library of Medicine and National Institute of Health (2018) guidelines also recommend limiting each paragraph to one idea, breaking up dense paragraphs into bullet points, finding alternatives for complex words, and avoiding abbreviations, acronyms and jargon. These guidelines and
other similar resources (such as Weis 2003; Markel 2012, pp. 227–259; Kasabwala et al 2013; Matthews and Matthews 2014, pp. 128–131) should be considered when writing future versions of HS-related websites.

Readability assessments, while useful, have their inherent flaws (as described in section 3.1.1) and only tell part of the understandability story. Other features of a given piece of text should also be assessed to gain a more comprehensive view on the comprehensibility of written materials. Subsequent sections look at some other aspects of information quality on the popular HS patient websites.

6.3 Coherence assessments

As there is a lack of HS website proposition density data, results from this study could not be compared with other HS studies. The data reported herein indicate that most of the popular HS websites are written at appropriate coherency levels for the lay population as the sites had similar coherent scores to excerpts from samples representative of the average literacy levels in OECD countries.

The average proposition density score (0.4991) for the popular HS websites is slightly higher than the figure reported by Tulsieram et al (2016) who found an average coherence scores of ~0.4702 for Canadian Ministry for Health information websites on human papilloma virus. Tulsieram et al (2016) found that the coherence levels on some of the information websites were significantly different than information representative of the Canadian lay population and concluded that the information on some of the websites were not appropriate for the intended audience. Ta-Min et al (2007) reported slightly lower proposition density scores (0.39–0.45) for their analysis of various cancer websites. Ta-min et al (2007) concluded that the information on the patient information websites was inadequate for the lay population. In contrast, the results of the current study indicate that most (18) of the 21 HS patient websites analysed are written at coherence levels appropriate for the intended audience.

To determine if there was any significant correlation between text coherence and readability, I performed correlation analysis. I did not find any significantly strong correlation between text readability and coherence, which is consistent with other reports (Ta-min et al 2007; Tulsieram et al 2016). This lack of correlation is not surprising given that both measures assess different aspects of text quality. Readability formulas assess a text’s comprehension by
analysing word and sentence length (DuBay 2004), while coherence evaluates the ability to process and store information (DeFrancesco and Perkins 2012). As they are independent measures, Ta-min et al (2007) and Tulsieram et al (2016) suggest that both text readability and coherence should be used when assessing information comprehension. The lack of a significantly strong correlation between readability and coherence found in the current study support the lack of an association between coherence and readability.

6.4 Content analysis

To assess the timeliness of the information on the websites, I analysed the proportion that mentioned the modern biologics drugs that are being increasingly used to treat HS (Saunte and Jemec, 2017; Theut Riis et al 2018). Some 52% (11/21) of sites referred to biologics as treatment options, which is almost double the 28% reported by Hessam et al (2017) when they analysed 39 of the top Google-ranking HS websites in 2016. While HS information is expanding at a rapid rate (as discussed in section 6.1), it is encouraging to see that an increasing number of the popular HS sites are incorporating this recent, relevant information.

To estimate one aspect of the usefulness of the HS websites, I examined if the sites mentioned patient support groups, which are increasingly being considered as important measures in HS management (Esmann and Jemec 2011; Woodruff et al 2015; Jemec 2017). Just 24% (5/21) referred to support groups and only three of these sites provided links to support groups. As social isolation and stigmatisation are frequent themes in the HS experience (Dufour et al 2014, Zouboulis et al 2015), online patient support groups have been advocated and several groups have been established in recent years. Such support structures enable those with HS to share common experiences and importantly, to realise that they are not alone, another common HS theme (Kouris et al 2016).

The readability software used in this study also generated summary statistics for the analysed text. This data revealed that the 21 HS websites analysed in this study have an average of 16 words per sentence across the HS-related content of the sites. However, the US Library of Medicine and National Institute of Health (2018) guidelines for improving the readability of patient information materials recommend using just 10–15 words per sentence. There was considerable variation in the number of polysyllabic words on the 21 HS websites. These ranged from one in eleven (9%) to one in three (33%), which were also deemed to be the most and least
readable, respectively, of the websites analysed. The US Library of Medicine and National Institute of Health (2018) guidelines recommend minimising the use of polysyllabic words in patient information to improve reader comprehension.

The word counts of each HS website were included as an important characteristic because text length impacts the readability of educational materials (Tulbert et al. 2011). Even motivated patients have a finite amount of time to read information and a limited attention span (Baker 1991). Considerable variation was seen in the word counts on each HS site, with an 8.55-fold difference between the HS websites with the lowest and highest word counts. Tulbert et al. (2011) recommend that online dermatology patient resources should ideally be ~1,200 words and the average word count for the 21 HS websites in this study was close to this recommended level at ~1,300 words. This average word count figure is also in line with content analyses of websites for other medical conditions such as Tulbert et al. (2011) who found an average ~2,200 words on 50 dermatology sites while Fitzsimmons et al. (2010) report an average of ~1,500 words on 100 Parkinson’s disease patient websites and Edmunds et al. (2014) report an average ~1,000 words from 50 websites about Graves’ disease and thyroid-associated ophthalmopathy. While HS is a complex and multifactorial disease, the essential disease information (i.e. description, causes, symptoms and treatments) can and should be communicated briefly and succinctly.

Regarding the use of jargon, most of the 21 HS sites mention the Hurley classification scheme for disease severity; of these, all sites quote verbatim the initial report (Hurley 1989) which appeared in a specialist medical textbook and uses much technical jargon. This jargon will likely not be understood by many lay readers. Substitution of specialised medical terms like “sinus tracts” with more common terms like “tunnels” would help improve the readability of these sections.

Visual aids were not used widely in the HS websites though they have been shown to enhance understanding of health information in those with low health literacy when used in combination with text (Houts et al. 2006; Williams et al. 2016). Even though clinical images can enhance patient comprehension and information recall and are recommended for online dermatology resources (Tulbert et al. 2011), just 24% of the HS websites in the current study contained pictures of the different Hurley HS stages. This number is a reduction on the figure from Hessam et al. (2017) who reported that 33% of the HS websites they analysed had HS-specific pictures. Hessam et al. (2017) recommend that HS patient websites should use
representative disease-specific photos, e.g. of all three Hurley stages, to enhance the educative message, especially against a background of low readability scores. As Hurley stages also inform treatment strategies (Zouboulis et al 2015; Saunte and Jemec 2017), it is important that there is not any ambiguity about what stage(s) the patient is at. One-third of the HS websites assessed in the current study failed to include any informative images, which have been shown to engage viewers and to support written text (World Wide Web Consortium 2018). Only a minority of websites included videos, which may further negatively impact on the comprehension of the online HS health information. Using videos featuring other people with similar health experiences can engage health information website users (Office of Disease Prevention and Health Promotion 2018).

6.5 Study limitations

There were some limitations to this study. While Google is the dominant search engine, other engines such as Yahoo and Bing are also used. A more comprehensive review would incorporate these and other search engines. In addition, other country-specific search engines are growing in popularity (e.g. Baidu in China and Yandex in Russia) and future work could incorporate these. This study looked at the term “hidradenitis suppurativa”, used mainly in English-speaking countries, and did not search for other names the disease is known by (e.g. “Akne inversa” in German-speaking regions, “La maladie de Verneuil” in French-speaking territories, “Inversa onlus” in Italy, and so on). Therefore, the results may not apply to non-English speaking regions. In addition, the readability formulas used in this study all have their inherent flaws (as discussed in section 6.2). Alternative readability scales catered specifically to scientific writing should be explored in future work. Some critics of readability formulas recommend usability/reader response tests as an alternative to readability assessments, and such an approach could be used to explore the comprehensibility of HS websites with patients.

Another study limitation is that only the text content was measured. The quality of the pictures, images and video content used was not assessed, all of which can influence the reader’s comprehension. Finally, this study offers a snapshot only of the digital HS landscape taken in June 2018. As illustrated by the results of this study (sections 5.1, 5.2), this digital environment is rapidly evolving and therefore may not accurately reflect each user’s online experience.
6.6 Summary

The results of this study show that there has been a large increase in online HS-related activity in recent years. However, this study shows that much of the popular online HS information intended for patients is written at levels far above the recommended levels. In contrast, the coherence of most of the HS sites is suitable for the intended audience. Furthermore, there is wide variation in how this online HS information is presented. Content creators should follow published guidelines on improving the readability of the HS-related content and effectively presenting this disease information online. Following these guidelines will increase the number of patients that this information can reach and will empower more HS patients to improve their health literacy.
Chapter 7. Conclusions and Recommendations

In this study, I analysed online interest in the chronic skin condition HS using Google search data and trends and found that there has been a large increase in online interest in HS in recent years. At the same time, the volume of HS-related digital content has also grown considerably. While these increases are welcome developments for those living with HS, I sought to assess the quality of information on the most popular HS websites by measuring the text readability (research question 1) and the coherence (research question 2). I also aimed to determine if the information on these online resources was up-to-date and useful to HS patients (research question 3) and if the information was presented in ways that engage the reader (research question 4).

I found, using five commonly used formulas, that the readability of all 21 HS sites analysed did not meet recommended readability guidelines. One-third of the sites are difficult to read, requiring a college level education to understand the content. Consequently, many people will struggle to comprehend the content. Four of the readability formulas used gave similar readability scores but the GFI formula significantly overestimated readability, highlighting the need to use more than one readability formula when conducting readability analysis. In contrast to the readability, I determined that the coherence of most of the HS sites to be appropriate for the intended audience. Readability and coherence are independent measures of information comprehensibility.

I assessed the timeliness of the websites and found that just over half of the HS websites mentioned the modern biologic drugs that are helping to alleviate disease severity and improve the quality of life in many HS patients. A quarter of the popular online HS resources referred to peer support groups, which are important to the typical HS patient as means of combatting the social isolation and loneliness commonly experienced by those living with the condition. I found large variations in the amount of disease-related information on the websites with an almost ten-fold difference in disease-related content between the websites with the lowest and highest wordcounts. Much of the information on the HS websites is presented in ways that will not engage readers. The use of visual aids in the popular HS patient resources varied widely. Incorporating more visual aids is a simple measure that will engage with more readers and facilitate deeper understanding of the information on the sites.
The main recommendations from this research are as follows:

- Patient-oriented HS websites provide an ideal opportunity to educate and empower patients and to rapidly communicate the latest research findings to the wider public;
- When assessing the readability of online content, it is important to use more than one readability formula;
- Content creators should consult published guidelines to improve the overall readability of the information;
- In addition to readability, other features such as coherence and website content should also be considered when assessing the quality of online content;
- The inclusion of informative clinical images and relevant videos to support written text may improve reader comprehension;
- HS patient websites should mention peer support groups to help those affected with HS realise that they are not alone.

My recommendations for future research are to:

- Assess the comprehensibility of online HS resources published in other languages, to determine if the findings of this study are limited to English;
- Suitably qualified healthcare professionals could assess the quality of visual information, specifically the use of clinical images, on HS websites;
- Incorporate usability studies into future comprehensibility assessments, to determine if patient-oriented HS sites engage readers, and to explore the overall comprehensibility of these sites with real users;
- Determine if there are significant differences in the readability of sites written by healthcare professionals when compared to sites written non-healthcare professionals
- Evaluate the overall suitability of HS patient sites using other information assessment tools such as the Suitability of Assessment Materials method
- Examine the quality of HS patient sites using other quality assessment tools such as HONCode compliance, a quality certification from the Health On the Net Foundation

In summary, there has been a large increase in online interest in HS in recent years. Simultaneously, there has been a sizeable rise in the volume of HS-related digital content and in
HS research. Communicating these research findings online provides a great opportunity for those living with HS, their family, friends and carers to find pertinent up-to-date disease-related information. However, much of this online HS information will be of limited use as it is written and presented in ways that many of the intended audience will struggle to understand. There are published recommendations on how to improve the comprehensibility of health-related content. Following these guidelines will widen the number of patients that this information can reach. In doing so, more HS patients will be able to find the appropriate information, be able to understand what they find, and to act appropriately to improve their health.
References


Canadian Hidradenitis Suppurativa Foundation (2018) [online], available at: 

1126–1130.

health literacy scale (HLS-EU-BR) applied in a Brazilian higher education population of Rio 
Grande do Sul (RS)’, in: Lavonen, J., Juuti, K., Lampiselkä, J., Uitto A. and Hahl K, eds, 
Helsinki, Finland: University of Helsinki. ISBN: 978-951-51-1541-6. [online], available at: 
https://repositorium.sdum.uminho.pt/bitstream/1822/41977/1/ESERA_HLS-EU-BR.pdf 
[accessed 02 Jan 2018].


self-management support websites for people with osteoarthritis: A text content analysis’, 
2018].

of Internet-based patient education materials related to endoscopic sinus surgery’, Laryngoscope, 
122(8), 1649–1654.

Chitika (2013), The value of Google result positioning, [online], available at: 


Irish Skin Foundation (2018), #five17eight, [online], available at: https://irishskin.ie/five17eight/ [accessed 16 June 2018].


treatments based on over 350 hidradenitis patients’, *Dermatology Online Journal*, Apr 15,19(4),1
Retrieved from: [http://escholarship.org/uc/item/5vw402nf](http://escholarship.org/uc/item/5vw402nf) [accessed 13 May 2018].

‘Association of hidradenitis suppurativa with body image’, *JAMA Dermatology*, 154(4), 447–
451.

Documentation*, 24(3), 138–140.

Transactions on Professional Communication*, 60(4), 343–383.

‘Medical and surgical treatment of hidradenitis suppurativa: a review’, *Skin Appendage
Disorders*, 3(2), 95–110.

information about lymphedema and its treatment’, *Plastic and Reconstructive Surgery*,137(2),
287e–295e.

stroke information’, *Journal of Stroke and Cerebrovascular Diseases*, 23(6), 1362–1367.

hidradenitis suppurativa: a population-based retrospective analysis’, *British Journal of


US Food and Drug Administration (FDA) website (2018), Drugs@FDA: FDA Approved Drug Products, [online], available at: [https://www.accessdata.fda.gov/scripts/cder/daf/#apphist](https://www.accessdata.fda.gov/scripts/cder/daf/#apphist) [accessed on 08 June 2018].


# Appendix 1: List of hidradenitis suppurativa websites with their Google ranking, inclusion/exclusion status and category type

<table>
<thead>
<tr>
<th>Rank</th>
<th>Website title</th>
<th>Website uniform resource locator</th>
<th>Include/Excluded</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mayo Clinic</td>
<td><a href="https://www.mayoclinic.org/diseases-conditions/hidradenitis-suppurativa/symptoms-causes/syc-20352306">https://www.mayoclinic.org/diseases-conditions/hidradenitis-suppurativa/symptoms-causes/syc-20352306</a></td>
<td>I</td>
<td>Health care company</td>
</tr>
<tr>
<td>2</td>
<td>Irish Skin Foundation</td>
<td><a href="https://irishskin.ie/hidradenitis-suppurativa/">https://irishskin.ie/hidradenitis-suppurativa/</a></td>
<td>I</td>
<td>NFP</td>
</tr>
<tr>
<td>3</td>
<td>American Association of Dermatology</td>
<td><a href="https://www.aad.org/public/diseases/painful-skin-joints/hidradenitis-suppurativa">https://www.aad.org/public/diseases/painful-skin-joints/hidradenitis-suppurativa</a></td>
<td>I</td>
<td>Professional</td>
</tr>
<tr>
<td>5</td>
<td>DermNet NZ</td>
<td><a href="https://www.dermnetnz.org/topics/hidradenitis-suppurativa/">https://www.dermnetnz.org/topics/hidradenitis-suppurativa/</a></td>
<td>I</td>
<td>Professional</td>
</tr>
<tr>
<td>6</td>
<td>WebMD</td>
<td><a href="https://www.webmd.com/skin-problems-and-treatments/hidradenitis-suppurativa#1">https://www.webmd.com/skin-problems-and-treatments/hidradenitis-suppurativa#1</a></td>
<td>I</td>
<td>Health portal</td>
</tr>
<tr>
<td>7</td>
<td>National Health Service</td>
<td><a href="https://www.nhs.uk/conditions/hidradenitis-suppurativa/">https://www.nhs.uk/conditions/hidradenitis-suppurativa/</a></td>
<td>I</td>
<td>Health service</td>
</tr>
<tr>
<td>8</td>
<td>Emedicine</td>
<td><a href="https://emedicine.medscape.com/article/1073117-overview">https://emedicine.medscape.com/article/1073117-overview</a></td>
<td>E</td>
<td>Scientific article</td>
</tr>
<tr>
<td>9</td>
<td>Medline</td>
<td><a href="https://medlineplus.gov/hidradenitissuppurativa.html">https://medlineplus.gov/hidradenitissuppurativa.html</a></td>
<td>E</td>
<td>Links page/basic information</td>
</tr>
<tr>
<td>10</td>
<td>MedicineNet</td>
<td><a href="https://www.medicinenet.com/image-collection/furuncle_s_aureus_picture/picture.htm">https://www.medicinenet.com/image-collection/furuncle_s_aureus_picture/picture.htm</a></td>
<td>E</td>
<td>Pictures only</td>
</tr>
<tr>
<td>11</td>
<td>HealthLine</td>
<td><a href="https://www.healthline.com/health/hidradenitis-suppurativa">https://www.healthline.com/health/hidradenitis-suppurativa</a></td>
<td>I</td>
<td>Health portal</td>
</tr>
<tr>
<td>12</td>
<td>JAMA Network</td>
<td><a href="https://jamanetwork.com/journals/jama/article-abstract/2664466">https://jamanetwork.com/journals/jama/article-abstract/2664466</a></td>
<td>E</td>
<td>Scientific article</td>
</tr>
<tr>
<td>13</td>
<td>JAMA Network</td>
<td><a href="https://jamanetwork.com/journals/jamadermatology/fullarticle/1731691">https://jamanetwork.com/journals/jamadermatology/fullarticle/1731691</a></td>
<td>E</td>
<td>Scientific article</td>
</tr>
<tr>
<td>14</td>
<td>Health Service</td>
<td><a href="https://www.hse.ie/eng/services/list/2/gp/antibiotic-prescribing/conditions-and-treatments/skin-soft-tissue/hidradenitis-suppurativa/">https://www.hse.ie/eng/services/list/2/gp/antibiotic-prescribing/conditions-and-treatments/skin-soft-tissue/hidradenitis-suppurativa/</a></td>
<td>I</td>
<td>Health service</td>
</tr>
<tr>
<td>16</td>
<td>Canadian Dermatology Association</td>
<td><a href="https://dermatology.ca/public-patients/skin/hidradenitis-suppurativa/">https://dermatology.ca/public-patients/skin/hidradenitis-suppurativa/</a></td>
<td>I</td>
<td>Professional</td>
</tr>
<tr>
<td>No</td>
<td>Source Description</td>
<td>URL</td>
<td>Type</td>
<td>Additional Info</td>
</tr>
<tr>
<td>-----</td>
<td>--------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>17</td>
<td>Patient Information</td>
<td><a href="https://patient.info/doctor/hidradenitis-suppurativa-pro">https://patient.info/doctor/hidradenitis-suppurativa-pro</a></td>
<td>E</td>
<td>For Medical professionals</td>
</tr>
<tr>
<td>18</td>
<td>Patient Information</td>
<td><a href="https://patient.info/health/hidradenitis-suppurativa-leaflet">https://patient.info/health/hidradenitis-suppurativa-leaflet</a></td>
<td>I</td>
<td>Health portal</td>
</tr>
<tr>
<td>19</td>
<td>Science Direct</td>
<td><a href="https://www.sciencedirect.com/science/article/pii/S1521693414001369">https://www.sciencedirect.com/science/article/pii/S1521693414001369</a></td>
<td>E</td>
<td>Scientific article</td>
</tr>
<tr>
<td>21</td>
<td>University of Rochester Medical Center</td>
<td><a href="https://www.urmc.rochester.edu/encyclopedia/content.aspx?ContentTypeID=134&amp;ContentID=163">https://www.urmc.rochester.edu/encyclopedia/content.aspx?ContentTypeID=134&amp;ContentID=163</a></td>
<td>I</td>
<td>Professional</td>
</tr>
<tr>
<td>22</td>
<td>British Medical Journal</td>
<td><a href="https://www.bmj.com/content/346/bmj.f2121">https://www.bmj.com/content/346/bmj.f2121</a></td>
<td>E</td>
<td>Scientific article</td>
</tr>
<tr>
<td>25</td>
<td>No BS about HS</td>
<td><a href="https://www.nbsabouths.com/what-is-hidradenitis-suppurativa/causes">https://www.nbsabouths.com/what-is-hidradenitis-suppurativa/causes</a></td>
<td>I</td>
<td>Health portal</td>
</tr>
<tr>
<td>26</td>
<td>British Skin Foundation</td>
<td><a href="http://www.britishskinfoundation.org.uk/SkinInformation/AtoZofSkinDisease/HidradenitisSuppurativa.aspx">http://www.britishskinfoundation.org.uk/SkinInformation/AtoZofSkinDisease/HidradenitisSuppurativa.aspx</a></td>
<td>I</td>
<td>NFP</td>
</tr>
<tr>
<td>27</td>
<td>eMedicine</td>
<td><a href="https://www.emedicinehealth.com/image-gallery/hidradenitis_suppurativa_picture/images.htm">https://www.emedicinehealth.com/image-gallery/hidradenitis_suppurativa_picture/images.htm</a></td>
<td>E</td>
<td>Pictures only</td>
</tr>
<tr>
<td>28</td>
<td>MSD</td>
<td><a href="https://www.msdmanuals.com/home/skin-disorders/acne-and-related-">https://www.msdmanuals.com/home/skin-disorders/acne-and-related-</a></td>
<td>I</td>
<td>Health portal</td>
</tr>
<tr>
<td>30</td>
<td>Guy’s and St. Thomas’ NHS Foundation Trust</td>
<td><a href="https://www.guysandstthomas.nhs.uk/our-services/dermatology/specialties/hidradenitis-suppurativa/Overview.aspx">https://www.guysandstthomas.nhs.uk/our-services/dermatology/specialties/hidradenitis-suppurativa/Overview.aspx</a></td>
<td>E</td>
<td>Hospital, not enough info</td>
</tr>
<tr>
<td>31</td>
<td>YouTube</td>
<td><a href="https://www.youtube.com/watch?v=zWlr5g-M198">https://www.youtube.com/watch?v=zWlr5g-M198</a></td>
<td>E</td>
<td>Video content</td>
</tr>
<tr>
<td>No.</td>
<td>Source</td>
<td>URL</td>
<td>Type</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>---------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>The Hidradenitis Suppurativa Foundation</td>
<td><a href="http://www.hs-foundation.org/">http://www.hs-foundation.org/</a></td>
<td>I</td>
<td>NFP</td>
</tr>
<tr>
<td>33</td>
<td>Mayo Clinic Proceedings</td>
<td><a href="https://www.mayoclinicproceedings.org/article/S0025-6196(15)00711-9/fulltext">https://www.mayoclinicproceedings.org/article/S0025-6196(15)00711-9/fulltext</a></td>
<td>E</td>
<td>Scientific article</td>
</tr>
<tr>
<td>34</td>
<td>SkinSight</td>
<td><a href="https://wwwskinsightcomskin-conditionsaadulthidradenitis-suppurativa">https://wwwskinsightcomskin-conditionsaadulthidradenitis-suppurativa</a></td>
<td>I</td>
<td>Health portal</td>
</tr>
<tr>
<td>35</td>
<td>The Mighty.com</td>
<td><a href="https://themightycomhidradenitis-suppurativa/">https://themightycomhidradenitis-suppurativa/</a></td>
<td>E</td>
<td>Blog</td>
</tr>
<tr>
<td>36</td>
<td>Medical News Today</td>
<td><a href="https://www.medicalnewstodaycomarticles75194php">https://www.medicalnewstodaycomarticles75194php</a></td>
<td>I</td>
<td>Health portal</td>
</tr>
<tr>
<td>37</td>
<td>UpToDate</td>
<td><a href="https://wwwuptodatecomcontentshidradenitis-suppurativa-treatment">https://wwwuptodatecomcontentshidradenitis-suppurativa-treatment</a></td>
<td>E</td>
<td>Health portal/ Subscription required</td>
</tr>
<tr>
<td>38</td>
<td>YouTube</td>
<td><a href="https://wwwyoutubecomwatch?v=fpeRpte_xM">https://wwwyoutubecomwatch?v=fpeRpte_xM</a></td>
<td>E</td>
<td>Video content</td>
</tr>
<tr>
<td>42</td>
<td>Humira.com</td>
<td><a href="https://wwwhumiracomhidradenitis-suppurativa">https://wwwhumiracomhidradenitis-suppurativa</a></td>
<td>E</td>
<td>Commercial site</td>
</tr>
<tr>
<td>43</td>
<td>Cochrane</td>
<td>http://wwwcochraneorgCD010081SKIN_treatments-hidradenitis-suppurativa</td>
<td>E</td>
<td>Scientific article</td>
</tr>
<tr>
<td>45</td>
<td>Thought Catalog</td>
<td><a href="https://thoughtcatalogcomI-a-wright201802what-gives-what-can-trigger-hidradenitis-suppurativa-and-how-to-avoid-these-triggers/">https://thoughtcatalogcomI-a-wright201802what-gives-what-can-trigger-hidradenitis-suppurativa-and-how-to-avoid-these-triggers/</a></td>
<td>E</td>
<td>Blog</td>
</tr>
<tr>
<td>46</td>
<td>Hidradenitis Suppurativa Info</td>
<td><a href="http://wwwhidradenitisinfocomedwhat-is-hidradenitis-suppurativa/">http://wwwhidradenitisinfocomedwhat-is-hidradenitis-suppurativa/</a></td>
<td>I</td>
<td>Hospital</td>
</tr>
<tr>
<td>47</td>
<td>American Family Physician</td>
<td><a href="https://wwwaafporgafp20051015p1547html">https://wwwaafporgafp20051015p1547html</a></td>
<td>E</td>
<td>Scientific article</td>
</tr>
<tr>
<td></td>
<td>Publication Title</td>
<td>URL</td>
<td>Type</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>------------------</td>
<td>-----------------------------------------------------------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Women’s Health</td>
<td><a href="https://www.womenshealthmag.com/health/a19941366/hidradenitis-suppurativa/">https://www.womenshealthmag.com/health/a19941366/hidradenitis-suppurativa/</a></td>
<td>E</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>F1000 Research</td>
<td><a href="https://f1000research.com/articles/6-1272/v1">https://f1000research.com/articles/6-1272/v1</a></td>
<td>E</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Canadian Family</td>
<td><a href="http://www.cfp.ca/content/63/2/114">http://www.cfp.ca/content/63/2/114</a></td>
<td>E</td>
<td></td>
</tr>
</tbody>
</table>

NFP: not-for profit
Appendix 2: Single factor analysis of variance output of Table 4 readability scores in Microsoft Office Excel.

<table>
<thead>
<tr>
<th>Groups</th>
<th>Count</th>
<th>Sum</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARI</td>
<td>21</td>
<td>218.6</td>
<td>10.40952</td>
<td>4.564905</td>
</tr>
<tr>
<td>FKGL</td>
<td>21</td>
<td>224.4</td>
<td>10.68571</td>
<td>4.632286</td>
</tr>
<tr>
<td>GFI</td>
<td>21</td>
<td>274.1</td>
<td>13.05238</td>
<td>4.815619</td>
</tr>
<tr>
<td>SMOG</td>
<td>21</td>
<td>203.2</td>
<td>9.67619</td>
<td>2.234905</td>
</tr>
</tbody>
</table>

ANOVA

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>134.4927</td>
<td>3</td>
<td>44.83091</td>
<td>11.03685</td>
<td>3.86E-06</td>
<td>2.718785</td>
</tr>
<tr>
<td>Within Groups</td>
<td>324.9543</td>
<td>80</td>
<td>4.061929</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>459.447</td>
<td>83</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ANOVA = analysis of variance; ARI = Automated Readability Index; FKGL = Flesch–Kincaid Grade Level; GFI= Gunning Fog Index; FRE= Flesch Reading Ease Score; SMOG= Simple Measure of Gobbledygook;
Appendix 3: Tukey–Kramer analysis of readability scores from Table 4 in Microsoft Office Excel.

<table>
<thead>
<tr>
<th></th>
<th>ARI</th>
<th>FKGL</th>
<th>GFI</th>
<th>SMOG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum</td>
<td>218.6</td>
<td>224.4</td>
<td>274.1</td>
<td>203.2</td>
</tr>
<tr>
<td>Mean</td>
<td>10.40952</td>
<td>10.68571</td>
<td>13.05238</td>
<td>9.67619</td>
</tr>
<tr>
<td>Variance</td>
<td>4.564905</td>
<td>4.632286</td>
<td>4.815619</td>
<td>2.234905</td>
</tr>
</tbody>
</table>

Factor levels 4
n 88
S^2_pooled 4.082911255
C 21

Qd 3.7107

<table>
<thead>
<tr>
<th>Tukey-Kramer procedure</th>
<th>Comparison</th>
<th>Absolute difference*</th>
<th>Critical Range **</th>
<th>Results***:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARI v FKGL</td>
<td>0.276190476</td>
<td>1.631970627</td>
<td>not significantly different</td>
</tr>
<tr>
<td></td>
<td>ARI v GFI</td>
<td>2.642857143</td>
<td>1.631970627</td>
<td>significantly different</td>
</tr>
<tr>
<td></td>
<td>ARI v SMOG</td>
<td>0.733333333</td>
<td>1.631970627</td>
<td>not significantly different</td>
</tr>
<tr>
<td></td>
<td>FKGL v GFI</td>
<td>2.366666667</td>
<td>1.631970627</td>
<td>significantly different</td>
</tr>
<tr>
<td></td>
<td>FKGL v SMOG</td>
<td>1.00952381</td>
<td>1.631970627</td>
<td>not significantly different</td>
</tr>
<tr>
<td></td>
<td>GFI v SMOG</td>
<td>3.376190476</td>
<td>1.631970627</td>
<td>significantly different</td>
</tr>
</tbody>
</table>

* Absolute difference: ABS (mean[sampleA]-mean[sampleB])

** Critical range:
Qd*[sqrt(S^2_pooled)/(c)]

*** Results: =IF(Absolute difference>critical range, means "significantly different", "not significantly different")

Abbreviations: ARI = Automated Readability Index; FKGL = Flesch–Kincaid Grade Level; GFI= Gunning Fog Index; FRE= Flesch Reading Ease Score; SMOG= Simple Measure of Gobbledygook;