# JMIR Dermatology

All topics related to diseases of the skin, hair, and nails, with special emphasis on technologies for information exchange, education, and clinical care
Volume 3 (2020), Issue 1    ISSN 2562-0959    Editor in Chief:  Robert Dellavalle, MD, PhD, MSPH

## Contents

### Original Papers

## Viewpoint

Original Paper

# Evaluation of the Level of Evidence Supporting the Recommendations Constituting the American Academy of Dermatology Clinical Practice Guidelines: Cross-Sectional Analysis

Courtney Cook[1], DO; Ryan Ottwell[1], BSc; Taylor Rogers[1], BSc; Jake Checketts[1], DO; Sanjeev Musuvathy[1]; Matt Vassar[1], PhD

[1] Oklahoma State University Center for Health Sciences, Tulsa, OK, United States

**Corresponding Author:**
Courtney Cook, DO
Oklahoma State University Center for Health Sciences
1111 W 17th St
Tulsa, OK, 74017
United States
Phone: 1 918 582 1972
Email: courco@okstate.edu

## Abstract

**Background:** Clinical practice guidelines are evidence-based recommendations used by physicians to improve patient care. These guidelines provide the physician with an assessment of the benefits and harms of a treatment and its alternatives. Therefore, it is essential that the clinical practice guidelines be based on the strongest available evidence. Numerous studies in a variety of different fields of medicine have demonstrated that recommendations supported by weak evidence are a common theme in clinical practice guidelines. A clinical guideline based solely on weak evidence has the capability to reduce the quality of care provided by physicians.

**Objective:** Our primary objective is to evaluate the levels of evidence supporting the recommendations constituting the American Academy of Dermatology clinical practice guidelines.

**Methods:** Using a cross-sectional study design, authors SM and RO located all current clinical practice guidelines on the American Academy of Dermatology website on June 10, 2017, and December 11, 2019. Each recommendation and its corresponding evidence rating were extracted in a duplicate and blinded fashion. A consensus meeting was planned a priori to resolve disagreements in extractions or stratifications.

**Results:** In total, 6 clinical guidelines and their subsections were screened and 899 recommendations were identified. Our final data set included 841 recommendations, as 58 recommendations contained no level of evidence and were excluded from calculations. Many recommendations were supported by a moderate level of evidence and therefore received a B rating (346/841, 41.1%). Roughly one-third of the recommendations were supported by a strong level of evidence and were given an A rating (n=307, 36.5%). The clinical practice guideline with the highest overall strength of evidence was regarding the treatment of acne, which had 17 of 35 (48.6%) recommendations supported by strong evidence and only 2 (5.7%) supported by weak evidence. The clinical practice guideline with the fewest recommendations supported by strong evidence was melanoma (13/63, 20.6%).

**Conclusions:** Clinical practice guidelines that lack strong supporting evidence could negatively affect patient care, and dermatologists should be mindful that not all recommendations are supported by the strongest level of evidence. An increased quantity of quality research needs to be performed in the field of dermatology to improve the evidence supporting the American Academy of Dermatology clinical practice guidelines.

***(JMIR Dermatol 2020;3(1):e17370)*** doi:10.2196/17370

**KEYWORDS**

clinical practice guidelines; level of evidence; dermatology; cross-sectional analysis

## Introduction

Clinical practice guidelines are defined as “statements that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options” [1]. Clinical practice guidelines are used by dermatologists to provide evidence-based treatment decisions to patients on a host of dermatologic conditions [2-4]. Evidence suggests widespread support among dermatologists regarding the use of clinical practice guidelines to improve the quality of patient care [5]. Since many dermatologists rely on clinical practice guidelines in the clinical setting [6], it is essential that the recommendations contained within clinical practice guidelines be supported by the best available evidence.

To date, the American Academy of Dermatology has released 6 clinical practice guidelines [7]. The American Academy of Dermatology uses the Strength of Recommendation Taxonomy (SORT) scale to assess the strength and quality of evidence used in the development of their guideline recommendations [8]. One advantage of the SORT scale is that it provides clinicians with a simple stratification system (A: strong supporting evidence; B: moderate supporting evidence; C: weak supporting evidence) for recommendations [8].

Guidelines based solely on expert opinion or low-quality evidence are at a significant risk of bias [9-12], a reality that can negatively affect the quality of care provided in the dermatology setting [3]. Collectively, low-quality recommendations are a common theme in clinical practice. Shaneyfelt and Centor [13] evaluated the strength of evidence in cardiology clinical practice guidelines and found 48% of the guidelines were based on low-quality evidence. Similar studies in dermatologic oncology [3], comorbidities in chronic conditions [14], orthopedic surgery [15], fibromyalgia [16], pediatric obesity [17], and wound care [18] have all demonstrated a need for improvement in the strength of evidence for clinical practice guidelines. In this study, we evaluate the strength of the evidence supporting the recommendations constituting the American Academy of Dermatology clinical practice guidelines.

## Methods

Authors SM and RO located all current clinical practice guidelines from the American Academy of Dermatology website on June 10, 2017, and December 11, 2019. All 6 published guidelines and their subsections are included in this study (Table 1). We evaluated the guidelines provided by the American Academy of Dermatology as they were the first organization to develop and publish guidelines for the clinical management of various cutaneous diseases [19]. We did not include Companion Consensus Statements, which are recommendations that return no evidence. We also excluded Appropriate Use Criteria, since these documents identify areas where sound data are not available or do not provide sufficient evidence of applying to the full range of patients seen in clinical practice [20].

The guidelines provide a list of included studies as well as an assessment of the methodological quality of these studies. Based on the results of these assessments, a strength of evidence rating is assigned to each recommendation. These ratings are presented in each clinical practice guidelines Summary of Recommendations section. The SORT scale and definitions are included in Table 2.

SM and RO independently extracted each recommendation and its corresponding evidence rating from the clinical practice guidelines. Data extraction was conducted in duplicate fashion with each investigator masked to the other’s responses. Authors next stratified each recommendation by quality of evidence and clinical practice guideline. A consensus meeting was planned a priori to resolve disagreements in extractions or stratifications. If an agreement could not be reached, a third party arbitrator, MV, was available to resolve any disagreements. Author CC is a dermatology resident, JC is an orthopedic surgery resident, RO and TR are fourth-year medical students, SM is a student that participated in a summer research program through Oklahoma State College Center for Health Sciences, and MV is a clinical assistant professor of psychiatry and behavioral sciences.

**Table 1.** Characteristics of the American Academy of Dermatology clinical practice guidelines (N=841).

| Guidelines and sections | Publication date | Number of recommendations |
|---|---|---|
| Acne (n=35) | 2016 | 35 |
| **Atopic dermatitis (n=66)** | | |
| Diagnosis and assessment | 2014 | 8 |
| Topical therapy | 2014 | 26 |
| Phototherapy and systemic agents | 2014 | 16 |
| Disease flares and adjunctive therapy | 2014 | 16 |
| Melanoma (n=62) | 2019 | 62 |
| **Nonmelanoma skin cancer (n=97)** | | |
| Cutaneous squamous cell carcinoma | 2018 | 54 |
| Basal cell carcinoma | 2018 | 43 |
| Office-based surgery (n=38) | 2016 | 38 |
| **Psoriasis (n=543)** | | |
| Biologics | 2008/2019[a] | 138 |
| Comorbidity | 2019 | 51 |
| Psoriatic arthritis | 2008 | 14 |
| Topical therapy | 2009 | 81 |
| Systemic agents | 2009 | 139 |
| Phototherapy and photochemotherapy | 2010/2019[a] | 104 |
| Case-based review | 2011 | 16 |

[a] This section includes new additional guidelines.

**Table 2.** Strength of Recommendation Taxonomy (SORT) rating scale and definitions.

| Strength of recommendation | Evidence supporting recommendation | Description of recommendation |
|---|---|---|
| A | Strong | The recommendation is based on consistent and good-quality patient-oriented evidence. |
| B | Moderate | The recommendation is based on inconsistent or limited-quality patient-oriented evidence. |
| C | Weak | The recommendation is based on consensus, usual practice, opinion, disease-oriented evidence, or case series for studies of diagnosis, treatment, prevention, or screening. |

## Results

In total, 6 guidelines consisting of 899 individual recommendations were identified. Of these, 58 recommendations were not given an evidence level and were excluded from calculations, leaving 841 recommendations in our data set (Table 1). A large proportion of recommendations received a B rating, indicating moderate evidence to support the recommendation (346/841, 41.1%). Just over one-third of the recommendations were supported by strong evidence (n=307, 36.5%), while less than a quarter of the recommendations were supported by weak evidence (n=188, 22.4%; Table 3).

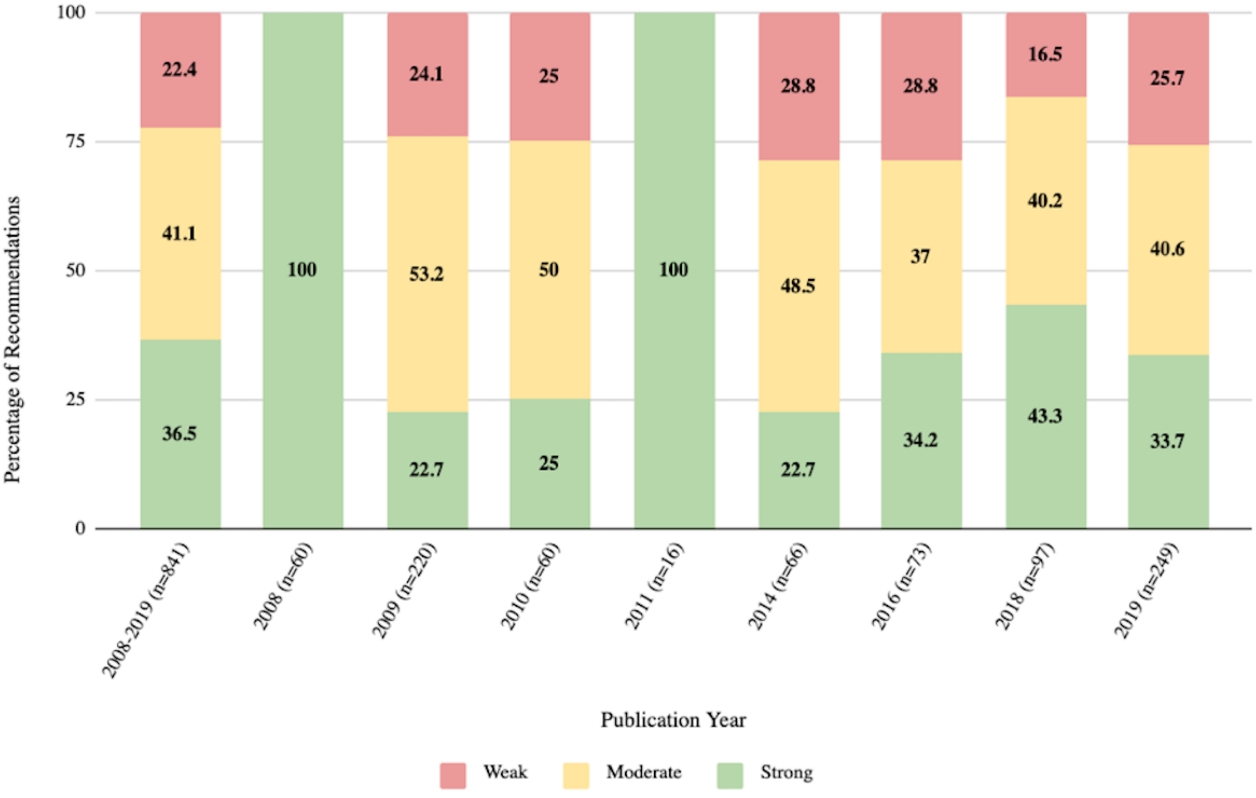
Recommendations were stratified by clinical practice guideline (Table 3). The clinical practice guideline with the highest overall strength of evidence was regarding the treatment of acne, which had 17 of 35 (48.6%) recommendations supported by strong evidence and only 2 (5.7%) recommendations supported by weak evidence. The nonmelanoma skin cancer guideline fared second-best, with 42 of its 97 (43.3%) recommendations supported by strong evidence and 16 (16.5%) of the recommendations supported by weak evidence. The clinical practice guideline with the fewest recommendations supported by strong evidence was melanoma, which only had 13 of its 63 (20.6%) recommendations supported by strong evidence. The guideline with the most recommendations supported by weak evidence is office-based surgery, with half of the recommendations supported by weak evidence (19/38, 50.0%).

The years 2008 (60/60, 100%) and 2011 (16/16, 100%) had the highest percentage of recommendations supported by strong evidence (Figure 1). The years 2014 (19/66, 28.8%) and 2016 (21/73, 28.8%) had the greatest percentage of recommendations supported by weak evidence. The newest guidelines, published in 2019, had a 9.6% decrease in the amount of recommendations supported by strong evidence and a 9.2% increase in recommendations supported by weak evidence. There is no consistent pattern suggesting that the level of evidence supporting the recommendations is improving or worsening.

**Table 3.** Strength of evidence for each American Academy of Dermatology clinical practice guideline.

| Guideline recommendations | Support of recommendation | | |
|---|---|---|---|
| | A (Strong), n (%) | B (Moderate), n (%) | C (Weak), n (%) |
| Acne (n=35) | 17 (48.6) | 16 (45.7) | 2 (5.7) |
| Atopic dermatitis (n=66) | 15 (22.7) | 32 (48.5) | 19 (28.8) |
| Melanoma (n=63) | 13 (20.6) | 20 (31.8) | 30 (47.6) |
| Nonmelanoma skin cancer (n=97) | 42 (43.3) | 39 (40.2) | 16 (16.5) |
| Office-based surgery (n=38) | 8 (21.1) | 11 (28.9) | 19 (50) |
| Psoriasis (n=542) | 212 (39.1) | 228 (42.1) | 102 (18.8) |
| Total (n=841) | 307 (36.5) | 346 (41.1) | 188 (22.4) |

**Figure 1.** Level of evidence supporting the American Academy of Dermatology clinical practice guideline recommendations, stratified by year of publication.



## Discussion

### Overview

The American Academy of Dermatology clinical practice guidelines aim to compile and summarize the best available evidence to guide dermatologists in their practice of evidence-based medicine. These guidelines are seen as essential by the majority of dermatologists, who believe they increase the quality of patient care [5]. However essential these guidelines may be to dermatologists, our study found that only 36.5% of the American Academy of Dermatology published guidelines are supported by strong evidence. Furthermore, there were inconsistencies in the strength of the recommendations between the 6 guidelines, with the range of recommendations supported by strong evidence spanning from nearly 20% to 50%. The newly published 2019 guidelines actually had a lower number of recommendations supported by strong evidence compared to the guidelines published in 2016 and 2018. Based on these results, we suggest that dermatologists be mindful that not all clinical practice guidelines are based on the strongest evidence and that dermatologists should regularly review guidelines with special attention to their level of evidence.

To our knowledge, we are the first to explore the level of evidence supporting the American Academy of Dermatology guidelines to include the most recently published guidelines (2019). The results of our study further complement the paradigm demonstrated in studies of other medical specialties in which the strength of clinical practice guidelines have been examined. For example, in obstetrics and gynecology, Wright et al [11] found that only 30% of the recommendations were based on high-quality evidence. In the American College of Cardiologists and American Heart Association clinical practice guidelines, Tricoci et al [10] found that a mere 11% of the recommendations were based on high-quality evidence. Shah et al [21] found that 0% of the evidence underlying the American College of Chest Physicians clinical practice guidelines for venous thromboembolism were based on the highest quality ratings, and Poonacha et al [22] found that only 9% of the National Comprehensive Cancer Network clinical practice guidelines recommendations were based on high-quality evidence. Members of our own research team demonstrated similar results to these studies. In addition, Meyer et al [23] found that in the American College of Gastroenterology, only 15% of the guidelines were supported by high evidence.

These findings, as well as those of our own study, suggest that a greater emphasis be placed on conducting and completing good-quality randomized trials and other studies to improve the quality of evidence-based medicine in dermatology. For example, the discrepancy observed between the strength of evidence supporting the acne and melanoma guidelines highlights the need for more high-quality melanoma studies. The American Academy of Dermatology melanoma clinical practice guideline identifies several gaps in research, with one being a lack of randomized controlled trials, thus demonstrating the need for an interdisciplinary approach to increase the number of high-quality studies [24]. The Strength of Recommendation Taxonomy states that for a recommendation to receive an A rating, it needs to have one of the following: a Cochrane Review with a clear recommendation, a United States Preventive Services Task Force (USPSTF) Grade A recommendation, consistent findings from at least two good-quality randomized controlled trials or diagnostic cohort studies, validated clinical decision rule in a relevant population, or a clinical evidence rating of “Beneficial” [8]. Therefore, we recommend that further research be performed to evaluate the quality and quantity of the evidence supporting the clinical practice guidelines.

Though the American Academy of Dermatology clinical practice guideline recommendations were primarily not supported by a strong level of evidence, the recommendations (especially the acne and nonmelanoma skin cancer clinical practice guidelines) were of much higher quality than those of other fields of medicine [10,21,22]. Though significant work is still required to raise the quality and quantity of evidence in dermatology guidelines, dermatology seems poised to become a leader in the compilation of truly evidence-based clinical practice guidelines if this direction is prioritized.

The strengths of our study include the following: data extraction was completed in a duplicate and blinded fashion, following the recommendations set forth by the Cochrane Handbook of Systematic Reviews of Interventions [25]; data extraction occurred over a 3-year period to include the most recent 2019 guidelines; and the results of our study were consistent with previously published articles in numerous fields of medicine. However, our study is not without limitations. Our study evaluated only clinical practice guidelines published by the American Academy of Dermatology and therefore is not generalizable outside the field of dermatology. Additionally, our study is not generalizable toward other American Academy of Dermatology quality measures such as Appropriate Use Criteria or other published literature. As some of the guidelines were published prior to the current year, they may not accurately reflect the current levels of evidence in dermatologic literature, and therefore our study may underestimate the current research quality in dermatology. Although we evaluated the overall levels of evidence of research underpinning guideline recommendations, we did not perform risk of bias assessments on each individual study. This may be seen as a perceived limitation of this study or as an opportunity for follow-up investigations. Furthermore, it should also be noted that just because a recommendation is not underpinned by strong quality evidence does not imply that it should be omitted from clinical practice guidelines. For certain recommendations, it may be difficult or impossible to achieve strong recommendations if they are based on high-risk populations or are of a direction unlikely to receive the attention of a randomized controlled trial.

## Conclusion

We demonstrated that the American Academy of Dermatology clinical practice guideline recommendations are primarily supported by moderate levels of evidence and only about one-third of the recommendations are supported by strong evidence. Although the American Academy of Dermatology clinical practice guidelines are supported by stronger evidence than several other guidelines, there is still a need for improvements in the quality and quantity of research in the field of dermatology.

## Conflicts of Interest

None declared.

## References

1. Institute of Medicine (US) Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. In: Graham R, Mancher M, Miller Wolman D, Greenfield S, Steinberg E, editors. Clinical Practice Guidelines We Can Trust. Washington, DC: National Academies Press (US); 2014.

2. Emmert B, Hallier E, Schön MP, Emmert S. Disease management guidelines in dermatology: implementation, potentials and limitations exemplified by the guidelines for the management of hand eczema [in German]. Hautarzt 2011 Apr 23;62(4):308-314. [doi: 10.1007/s00105-011-2128-7] [Medline: 21424892]
3. de Haas E, de Vijlder H, van Reesema WS, van Everdingen J, Neumann H. Quality of clinical practice guidelines in dermatological oncology. J Eur Acad Dermatol Venereol 2007 Oct 09;21(9):1193-1198. [doi: 10.1111/j.1468-3083.2007.02216.x] [Medline: 17894704]
4. Radtke MA, Reich K, Spehr C, Augustin M. Treatment goals in psoriasis routine care. Arch Dermatol Res 2015 Jul 31;307(5):445-449. [doi: 10.1007/s00403-014-1534-y] [Medline: 25549606]
5. Wakkee M, Lugtenberg M, Spuls P, de Jong EM, Thio H, Westert G, et al. Knowledge, attitudes and use of the guidelines for the treatment of moderate to severe plaque psoriasis among Dutch dermatologists. Br J Dermatol 2008 Aug;159(2):426-432. [doi: 10.1111/j.1365-2133.2008.08692.x] [Medline: 18616791]
6. Taba P, Rosenthal M, Habicht J, Tarien H, Mathiesen M, Hill S, et al. Barriers and facilitators to the implementation of clinical practice guidelines: a cross-sectional survey among physicians in Estonia. BMC Health Serv Res 2012 Dec 13;12(1):455 [FREE Full text] [doi: 10.1186/1472-6963-12-455] [Medline: 23234504]
7. Journal of the American Academy of Dermatology. Clinical Guidelines. URL: https://www.jaad.org/content/collection-clinical-guidelines [accessed 2019-12-22]
8. Ebell MH, Siwek J, Weiss BD, Woolf SH, Susman J, Ewigman B, et al. Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature. J Am Board Fam Pract 2004 Jan 01;17(1):59-67 [FREE Full text] [doi: 10.3122/jabfm.17.1.59] [Medline: 15014055]
9. Hirsh J, Guyatt G. Clinical experts or methodologists to write clinical guidelines? The Lancet 2009 Jul;374(9686):273-275. [doi: 10.1016/s0140-6736(09)60787-x]
10. Tricoci P, Allen JM, Kramer JM, Califf RM, Smith SC. Scientific evidence underlying the ACC/AHA clinical practice guidelines. JAMA 2009 Feb 25;301(8):831-841. [doi: 10.1001/jama.2009.205] [Medline: 19244190]
11. Wright JD, Pawar N, Gonzalez JSR, Lewin SN, Burke WM, Simpson LL, et al. Scientific Evidence Underlying the American College of Obstetricians and Gynecologists Practice Bulletins. Obstetrics & Gynecology 2011;118(3):505-512. [doi: 10.1097/aog.0b013e3182267f43]
12. Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill S, GRADE Working Group. Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches The GRADE Working Group. BMC Health Serv Res 2004 Dec 22;4(1):38 [FREE Full text] [doi: 10.1186/1472-6963-4-38] [Medline: 15615589]
13. Shaneyfelt TM, Centor RM. Reassessment of clinical practice guidelines: go gently into that good night. JAMA 2009 Feb 25;301(8):868-869. [doi: 10.1001/jama.2009.225] [Medline: 19244197]
14. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. JAMA 2005 Aug 10;294(6):716-724. [doi: 10.1001/jama.294.6.716] [Medline: 16091574]
15. Kahle JT, Highsmith MJ, Kenney J, Ruth T, Lunseth PA, Ertl J. The effectiveness of the bone bridge transtibial amputation technique: A systematic review of high-quality evidence. Prosthet Orthot Int 2017 Jun 20;41(3):219-226. [doi: 10.1177/0309364616679318] [Medline: 27913784]
16. Thieme K, Mathys M, Turk DC. Evidenced-Based Guidelines on the Treatment of Fibromyalgia Patients: Are They Consistent and If Not, Why Not? Have Effective Psychological Treatments Been Overlooked? J Pain 2017 Jul;18(7):747-756. [doi: 10.1016/j.jpain.2016.12.006] [Medline: 28034828]
17. Nissen T, Wayant C, Wahlstrom A, Sinnett P, Fugate C, Herrington J, et al. Methodological quality, completeness of reporting and use of systematic reviews as evidence in clinical practice guidelines for paediatric overweight and obesity. Clin Obes 2017 Feb 23;7(1):34-45. [doi: 10.1111/cob.12174] [Medline: 28112500]
18. Morton LM, Bolton LL, Corbett LQ, Girolami S, Phillips TJ. An Evaluation of the Association for the Advancement of Wound Care Venous Ulcer Guideline and Recommendations for Further Research. Advances in Skin & Wound Care 2013;26(12):553-561. [doi: 10.1097/01.asw.0000434057.81199.6c]
19. Drake LA, Yale KP, Lowery BJ, Dunbar AL, Gillies R. American Academy of Dermatology guidelines of care. Development and process. Arch Dermatol 1997 Nov;133(11):1369-1374. [Medline: 9371018]
20. Journal of the American Academy of Dermatology. Appropriate use criteria. URL: https://www.aad.org/practicecenter/quality/appropriate-use-criteria [accessed 2017-07-26]
21. Shah S, Khan AR, Khan S, Assaly R. Evidence Underlying the American College of Chest Physicians (ACCP) Clinical Practice Guidelines for the Prevention, Diagnosis, and Management of Venous Thromboembolism. Chest 2014 Oct;146(4):821A. [doi: 10.1378/chest.1995242]
22. Poonacha TK, Go RS. Level of Scientific Evidence Underlying Recommendations Arising From the National Comprehensive Cancer Network Clinical Practice Guidelines. JCO 2011 Jan 10;29(2):186-191. [doi: 10.1200/jco.2010.31.6414]
23. Meyer C, Bowers A, Wayant C, Checketts J, Scott J, Musuvathy S, et al. Scientific evidence underlying the American College of Gastroenterology's clinical practice guidelines. PLoS One 2018 Oct 3;13(10):e0204720 [FREE Full text] [doi: 10.1371/journal.pone.0204720] [Medline: 30281671]

24. Swetter SM, Tsao H, Bichakjian CK, Curiel-Lewandrowski C, Elder DE, Gershenwald JE, et al. Guidelines of care for the management of primary cutaneous melanoma. J Am Acad Dermatol 2019 Jan;80(1):208-250. [doi: 10.1016/j.jaad.2018.08.055] [Medline: 30392755]
25. Higgins JG. editor. Cochrane Handbook for Systematic Reviews of Interventions. John Wiley & Sons 2008;2008. [doi: 10.1002/9780470712184]

## Abbreviations

**SORT:** Strength of Recommendation Taxonomy

---

*Edited by G Eysenbach; submitted 23.12.19; peer-reviewed by T Muto, F Kaliyadan; comments to author 01.05.20; revised version received 05.05.20; accepted 14.05.20; published 03.07.20.*

*Please cite as:*
*Cook C, Ottwell R, Rogers T, Checketts J, Musuvathy S, Vassar M*
*Evaluation of the Level of Evidence Supporting the Recommendations Constituting the American Academy of Dermatology Clinical Practice Guidelines: Cross-Sectional Analysis*
*JMIR Dermatol 2020;3(1):e17370*
*URL: http://derma.jmir.org/2020/1/e17370/*
*doi:10.2196/17370*
*PMID:*

Original Paper

# The Feasibility and Acceptability of Using a Wearable UV Radiation Exposure Monitoring Device in Adults and Children: Cross-Sectional Questionnaire Study

Elizabeth Sela Nagelhout[1*], MPH, PhD; Riley Lensink[2*], BS; Angela Zhu[2*], BS; Bridget G Parsons[2*], MSPH; Jakob D Jensen[3*], PhD; Yelena P Wu[2,4*], PhD

[1]Department of Family and Preventive Medicine, University of Utah, Salt Lake City, UT, United States

[2]Huntsman Cancer Institute, Salt Lake City, UT, United States

[3]Department of Communications, University of Utah, Salt Lake City, UT, United States

[4]Department of Dermatology, University of Utah, Salt Lake City, UT, United States

[*]all authors contributed equally

**Corresponding Author:**
Yelena P Wu, PhD
Huntsman Cancer Institute
2000 Circle of Hope
Salt Lake City, UT, 84112
United States
Phone: 1 8012135653
Email: yelena.wu@utah.edu

## Abstract

**Background:** In the United States, melanoma is the fifth most common type of cancer. Reducing UV radiation (UVR) exposure is essential for the prevention of melanoma. The assessment of individual-level UVR exposure using wearable technology is a promising method to monitor and reduce UVR exposure. However, the feasibility and acceptability of using wearable UVR monitoring devices have not been assessed.

**Objective:** This study aimed to assess the feasibility and acceptability of using a commercially available UVR monitoring wearable device in adults and children.

**Methods:** We recruited families (1 parent and 1 child) to test a new, commercially available UVR monitoring device (namely, Shade). Participants wore Shade for 2 weeks and completed questionnaires assessing the feasibility and acceptability of wearing the device. Qualitative analyses were conducted to summarize participants' open-ended responses regarding device feasibility.

**Results:** A total of 194 individuals (97 adults and 97 children) participated in this study. The participating children were aged, on average, 12.7 years. Overall, adults and children reported moderate satisfaction with wearing Shade. The feasibility of the use of Shade was adequate, with 73% (65/89) of adults and 61% (54/89) of children reporting that they wore the device "all of the time they were outside." Through open-ended responses, participants reported that the device was easy to use, was compact, and increased their awareness about their exposure to UVR.

**Conclusions:** Adults and children can feasibly use a wearable UVR monitoring device, and the use of the device was acceptable to participants. The device could be integrated into melanoma preventive interventions to increase individuals' and families' awareness of UVR exposure and to facilitate the use of recommended melanoma preventive strategies.

***(JMIR Dermatol 2020;3(1):e15711)*** doi:10.2196/15711

**KEYWORDS**

ultraviolet radiation exposure; wearable device; melanoma; melanoma prevention

# Introduction

## Background

In the United States, melanoma is the fifth most common type of cancer [1]. Over the past three decades, overall melanoma incidence has increased by 1.5% each year; the incidence in pediatric populations is increasing, and these trends are expected to continue [1]. Individual UV radiation (UVR) exposure is the primary modifiable risk factor related to melanoma development [2]. The reduction of UVR exposure through the use of sun protection strategies (eg, wearing a sunscreen, wearing long sleeves, and avoiding peak UVR hours from 10 AM to 4 PM) is essential for the prevention of melanoma. Decreasing UVR exposure during childhood is especially critical because 25% of one's lifetime UVR exposure occurs during childhood, a key period of skin cell development and susceptibility [3,4]. Children and adolescents often rely on their parents to engage in sun-safe habits (eg, helping apply a sunscreen, providing long-sleeved shirts for children to wear, and buying a sunscreen for the household), thereby decreasing their UVR exposure, and are more likely to be successful in engaging in these behaviors if their parents model sun-safe habits [5-7]. However, many parents do not provide sufficient sun protection for their children [7]. Furthermore, children often do not use the recommended sun protection strategies to decrease UVR exposure, and they receive multiple sunburns, which can double their lifetime risk of melanoma [8,9].

Typically, the assessment of an individual's UVR exposure has relied on self-reported questionnaires and diaries of sun protection, time spent outdoors, and the number of sunburns [10-13]. However, self-reported measurements of UVR exposure can be upwardly biased because of inaccurate recall and social desirability effects [14-17]. There are few objective assessments of UVR exposure available. The most common method for assessing objectively measured UVR exposure has been through the use of personal electronic dosimeters, which are worn as wristwatches or fixed to clothing and collect individual UVR exposure at instantaneous readings (eg, every second) [18-24]. Previous studies using electronic dosimeters have been limited by the use of dosimeters that detect UV-B only, are not commercially available and thus have a lower likelihood for broader dissemination, have low sensitivity and specificity, and are unable to filter out visible light that could overestimate actual UVR measurements [25-27].

The Shade UVR sensor is a newly developed wearable radiometer, which records both instantaneous readings and accumulated UVR doses over time [26]. The device is capable of measuring both UV-B and UV-A, filtering out visible light, and is weighted according to the erythemal action spectra, which is standardized by the International Commission on Illumination and adopted by the World Health Organization [26,28,29]. Using a standardized unit of UVR exposure allows for the comparability of exposure measurements among study populations. In addition, Shade has documented excellent sensitivity and accuracy, including when compared with other UVR monitoring devices [26]. For example, the accuracy of Shade is roughly 80%, whereas the accuracy of the Band wearable device is roughly 20% [26].

## Objectives

Commercially available UVR monitoring devices have the potential to be integrated into melanoma preventive interventions aimed at reducing UVR exposure. For example, such devices could facilitate the self-monitoring of UVR exposure and increase an individual's awareness of their overall UVR exposure levels. In other areas of research and practice, including those focused on increasing physical activity to prevent or treat obesity, self-monitoring of health behaviors (eg, time spent engaging in physical activity) has been shown to be beneficial for positive health behavior changes [30,31]. However, previous studies have not yet assessed the feasibility and acceptability of using a commercially available UVR monitoring device, from participants' perspectives [19,21,32]. Establishing the adequate feasibility and acceptability of such devices is essential before being used in skin cancer preventive interventions aimed at decreasing UVR exposure. The purpose of this study was to examine the feasibility and acceptability of using a newly developed, commercially available UVR monitoring wearable device (ie, Shade) in children and adults.

# Methods

## Study Sample

All study participants were recruited within the state of Utah, which has the highest incidence of melanoma in the United States [33]. Adults were eligible to participate in the study if (1) they were at least aged 18 years, (2) were residents of Utah, (3) had at least one child aged 8 to 17 years who was willing to participate in the study, (4) did not have a pacemaker (because of the strong magnet in the UVR monitoring device), (5) had and were willing to use a Bluetooth or Wi-Fi–enabled mobile phone for research purposes (to communicate with the research staff and synchronized the device), (6) were willing to download and use a free mobile phone app that shared their UVR exposure data with the research team, and (7) were able to read and speak English. Children were eligible to participate if (1) they were aged 8 to 17 years, (2) lived with a primary caretaker in Utah, and (3) did not have a pacemaker. Children who did not have a Bluetooth or Wi-Fi–enabled device that allowed them to share their UVR exposure data with the research team were loaned one for the duration of the study.

A total of 194 participants (97 parent-child dyads) were enrolled in the study. Of the adults who completed eligibility screening, 34 were ineligible. The reasons for ineligibility included the parent not having children aged between 8 and 17 years (n=28), not having a mobile phone with Bluetooth and Wi-Fi (n=5), and being unable to read English (n=1). Of the 116 eligible adults, 7 decided not to participate and 12 were unable to participate because of a limited number of devices available for use within the context of this study. There were no parent-child dyads excluded because of child ineligibility after it was determined that the parent was eligible to participate. Data were collected between June 2018 and October 2018 to capture the experiences of adults and children wearing the device during the summer months and once school had started. An institutional review board approved all the study procedures.

### Study Procedures

Before their study enrollment, parental consent, parental permission, and child assent forms were emailed to participants, and Shade was mailed to participants. Once participants received the devices, they provided informed consent by phone. During the consent call, research assistants reviewed instructions on the use and care for the device. Parents and children were each asked to wear Shade clipped to the chest portion of their shirt for 14 days during waking hours. Shade collects the individual, time-stamped standard erythemal dose (SED), a standardized measure of UVR exposure, accumulated over the course of the day and reports readings of the SED every 6 min. The device is 1.58 inches in diameter and weighs 0.48 ounces and can be attached to clothing with a built-in magnet. The data collected by the device were synchronized using the Shade research app. Shade, which is commercially available, is accompanied by an app; however, for the purposes of our research project, we used the Shade research app. The research app did not provide participants any information on their UVR exposure, as we were focused on examining the feasibility and acceptability of wearing the device separately from the UVR feedback feature. The app was used to collect the UV data from the device and to notify participants when the device battery needed to be recharged. The participants were also instructed to avoid getting the device wet (as it is not waterproof), to synchronize the device to the Shade research app daily, and to recharge their device at least every 3 days. At the start of the 14-day study period, parents and children were asked to complete a baseline questionnaire that included items assessing demographic information. They were also asked to complete once-daily questionnaires assessing whether they had worn the device. Finally, parents and children were asked to complete an exit questionnaire assessing the feasibility and acceptability of wearing Shade. Parents and children reported on their own use of the device separately. Feasibility and acceptability were conceptually defined based on previous research [34,35]. Feasibility was defined as “Are participants capable of and willing to integrate wearing the Shade device into their daily lives?” Acceptability was defined as “To what extent wearing the Shade device was judged as suitable, satisfying, or attractive to study participants?”

### Feasibility

Feasibility was assessed through four multioption quantitative items and two open-ended questions. The quantitative items included the following: (1) “During days you wore the Shade device, how much of the time between 7 am and 7 pm did you wear the device when you were outside?” (eg, “wearing it ¼ of the time” and “wearing it all of the time”); (2) “When did you wear the Shade device?” (eg, “I put it on in the morning and wore it all day [except in the shower or while swimming], and took it off again in the evening” and “I wore it most of the time but took it off at special occasions”); (3) “When you wore the Shade device, did you pay attention to it?” (eg, “Yes, I was aware that it should be clipped to my clothing, exposed to the sun, but I didn’t change my behavior because of it” and “No, I didn’t pay attention to it at all”); and (4) “While you wore the Shade device, did people around you notice it and ask you what it was for?” (eg, “yes, many” and “no, no one”). Open-ended feasibility questions assessed perceptions of enjoyment and usability and included the following questions: “What did you like about the Shade device?” and “What were some of the challenges of wearing the Shade device?”

### Acceptability

The acceptability of using Shade was measured using a 17-item modified questionnaire that assessed the perceptions of comfort, enjoyment, privacy, and usefulness of the device [34]. The original questionnaire was modified to focus on the acceptability of wearing the device. Responses were measured on a 5-point Likert-type scale, scored 1 to 5, ranging from *strongly disagree* to *strongly agree*. For example, participants were asked to rate how much they agreed with items such as “I think the device was comfortable.” All questionnaires were administered using Research Electronic Data Capture (REDCap) [36,37].

### Analytic Plan

Descriptive statistics were calculated to summarize participant demographic characteristics and to report the proportions of participants’ responses to quantitative feasibility and acceptability questions. Means and standard deviations were calculated to summarize the number of days the device was worn. The qualitative feasibility questions were coded using content analysis to identify common themes describing what parents and children liked about the device (eg, ease of use and awareness of their UVR exposure) and challenges they experienced when wearing the device (eg, the device falling off) [38]. The codebook was created by three research team members based on the initial coding of 10% of parents’ and children’s responses for each of the two open-ended feasibility questions. Overall, 50% of all parents’ and children’s responses were coded by two independent coders (percent agreement=95%). All discrepancies in coding were discussed and resolved by the coders before commencing the analysis.

## *Results*

### Demographic Characteristics

A total of 97 adults and 97 children (N=194 individuals) participated in this study. Among adults, 87% (83/95) were non-Hispanic white and 5% (5/95) were Hispanic and 77% (73/95) were women (Table 1). The average age of participating children was 12.7 years (SD 2.6), 85% (81/95) were non-Hispanic white and 8% (8/95) were Hispanic and 59% (56/95) were female (Table 1).

**Table 1.** Demographic characteristics of participants.

| Characteristics | Adults (n=97) | Children (n=97) |
|---|---|---|
| Age (years), mean (SD) | 41.6 (6.3) | 12.7 (2.7) |
| **Sex, n (%)** | | |
| Male | 22 (23) | 39 (41) |
| Female | 73 (77) | 56 (59) |
| **Race, n (%)** | | |
| Non-Hispanic white | 83 (87) | 81 (85) |
| Hispanic | 5 (5) | 8 (8) |
| Asian or Asian American | 5 (5) | 4 (4) |
| Other | 2 (2) | 2 (2) |
| **Marital status, n (%)** | | |
| Married or marriage-like relationship | 84 (88) | N/A[a] |
| Divorced/separated | 9 (10) | N/A |
| Widowed | 2 (2) | N/A |
| **Level of education, n (%)** | | |
| High school graduate or General Educational Development | 8 (8) | N/A |
| Vocational or technical school | 8 (8) | N/A |
| Some college, including 2-year degree | 33 (35) | N/A |
| Bachelor's degree | 25 (26) | N/A |
| Master's/doctoral degree | 21 (22) | N/A |
| **Family income, n (%)** | | |
| <US $50,000 | 23 (24) | N/A |
| >US $50,000 | 64 (67) | N/A |
| I would rather not report this | 8 (8) | N/A |
| **Occupation location, n (%)** | | |
| Mainly indoors | 82 (86) | N/A |
| Mainly outdoors | 9 (9) | N/A |
| In a motor vehicle | 3 (3) | N/A |

[a] N/A: not applicable (not asked of children).

## Acceptability of a UV Radiation Monitoring Device

Parents and children reported moderate levels of satisfaction with wearing Shade. Parents *agreed or strongly agreed* that the device was well suited for their bodies (45/89, 51%), was comfortable (57/89, 64%), and was easy to wear (68/69, 76%; Table 2). More than 80% (73/88) of children *agreed or strongly agreed* that the device was easy to wear. Only 8% (7/89) of parents and 17% (15/88) of children *agreed or strongly agreed* that the device felt weird physically (Table 2). Only 25% (22/89) of parents *agreed or strongly agreed* that they would purchase the device.

**Table 2.** Parent's and children's reported acceptability of Shade, a UV radiation monitoring device.

| Acceptability of Shade | Agree or strongly agree | |
|---|---|---|
| | Parents, n (%) | Children, n (%) |
| Wearing the device feels weird physically | 7 (8) | 15 (17) |
| I think the device is well suited for my body | 45 (51) | 39 (44) |
| I think the device was comfortable | 57 (64) | 33 (38) |
| I think the device was boring | 57 (64) | 13 (15) |
| I think the device was annoying | 19 (21) | 28 (32) |
| I think the device was pleasant | 20 (23) | 23 (26) |
| I think the device may threaten my privacy | 3 (3) | 4 (5) |
| If most people in my environment used the device, I would be more inclined to use it as well | 24 (27) | 32 (36) |
| If people who are influential in my life recommended that I use the device for a period of time, I would do so | 43 (48) | 46 (52) |
| I think I would wear the device only if I were forced to | 14 (16) | 12 (14) |
| If it were launched on the market at an affordable price, I would likely purchase it | 22 (25) | 15 (17) |
| If the device were available to me, I would use it | 44 (49) | 37 (42) |
| It seems tiresome to use the device | 26 (29) | 16 (18) |
| It seems easy to wear the device | 68 (76) | 73 (83) |
| The device would be incompatible with most aspects of my activities | 11 (12) | 26 (30) |
| The device limits the way in which I like to perform my activities | 8 (9) | 14 (16) |
| The device could improve the quality of my activities | 15 (17) | 12 (14) |

### Feasibility of Using a UV Radiation Monitoring Device

Parents reported wearing Shade for an average of 12.7 days (SD 2.54) out of 14 days, and children reported wearing their device for an average of 12.2 days (SD 3.03) out of 14 days. The majority of parents (65/89, 73%) and children (54/89, 61%) reported that they wore Shade whenever they were outdoors between 7 AM and 7 PM during the study period. When asked when they wore the device during the day, 80% (72/90) of parents and 71% (63/89) of children reported that they put it on in the morning and wore it all day and took it off again in the evening, indicating full compliance with the study protocol.

The vast majority of the sample (89/97, 92% of parents and 87/97, 90% of children) provided responses to the open-ended feasibility questions. The most commonly endorsed themes regarding what participants liked about Shade were the device's ease of use and compact size and that the device increased the participant's awareness about their sun-safe habits and tracked their UVR exposure (see Multimedia Appendix 1). The most commonly endorsed challenges of wearing Shade among parents and children included keeping the device's UVR sensor uncovered (eg, by clothing), remembering to wear the device, its impact on daily activities (eg, the device was in the way while doing chores and the device tugged on shirts), and the device inadvertently falling off. Another challenge mentioned by parents and children included receiving questions from others about the device (Multimedia Appendix 1).

## Discussion

### Principal Findings

The results of this study indicate that the use of a wearable UVR monitoring device was moderately acceptable and feasible for adults and children. In addition, parents and children were, on average, adherent to wearing the device for the vast majority of the desired monitoring days, indicating the potential for the use of monitoring devices in future melanoma prevention interventions. Although participants found wearing the device to be both acceptable and feasible, the results of this study indicated some ways in which the device could be improved for future use.

Exposure to UVR is the primary modifiable risk factor for melanoma [39]. Despite public health efforts to increase the awareness of the harmful effects of UVR through educational interventions, adults and adolescents continue to report experiencing sunburns and do not adhere to sun protection use (eg, wearing sunscreen or long sleeve shirts) [40]. The use of a feasible and acceptable UVR monitoring device, similar to the one assessed in this study, may be beneficial when combined with other health behavior strategies (eg, wearing a sunscreen and avoiding peak UVR hours) in driving behavior changes to reduce UVR exposure. Although participants were not provided with UVR exposure data in real time, 33% of adults and 9% of children reported that wearing the device made them more aware of their outdoor sun habits. More than 11% of children reported that they liked the device because it was able to track their UVR exposure. These findings suggest participants may be interested in using a wearable UVR exposure device to monitor their UVR exposure. Future studies may be helpful in identifying whether users find it feasible and acceptable to monitor their UVR exposure through the use of a monitoring device and an integrated UVR feedback system (eg, a phone app or device display). Our team is currently designing interventions to provide participants with personalized UVR feedback in real time in an effort to promote sun protection use.

Self-monitoring of health behaviors via objective measures has been shown in other populations to increase the awareness of behaviors such as sedentary periods and to counter self-reported underestimates of the time spent in sedentary activities [41]. Similarly, in the context of UVR exposure, it is likely that individuals do not have an accurate perception of the amount of UVR exposure they receive. For example, one study reported that a participant's perception of being in the sun for a *long period* could range from 30 min to an entire day, which would vastly affect the amount of UVR exposure they received [42]. Providing feedback on the amount of UVR exposure received during daily activities and accumulated throughout the course of an entire day could help individuals have a more accurate sense of their UVR exposure. Knowledge of one's daily UVR exposure will likely become increasingly important, as the overall amount of UVR that is reaching the earth's surface is increasing because of ozone depletion, which puts people at a greater risk for the harmful effects of UVR [43,44].

The results of this study also raised considerations for modifications of UVR monitoring wearable device technology for use in future research. Participants indicated wearing the device was challenging because they forgot to wear the device, it impacted their daily activities, the device fell off at times, and the device solicited questions from others. These results are similar to findings from other studies evaluating the acceptability of wearing health tracking devices. Previous studies have found that barriers to using a wearable device include remembering to put it on, the inability to wear it during certain activities, the device not being waterproof, fear of losing the device, and increased social tension resulting from peers asking about the device [45-47]. Future devices may be more acceptable to users if the device or associated mobile phone app had an alert function to remind them to put on their device each morning, if the device could be worn as a wristwatch or in another inconspicuous location so as to not attract attention from peers, and if the device was waterproof. Parents and children had differing views about some of the aspects of Shade. For example, 64% of parents and 37.5% of children *agreed or strongly agreed* that the device was comfortable, and 21.0% of parents and 31.8% of children found the device to be annoying. These differences may emphasize the need to tailor UVR wearable devices for use in adult and child populations separately.

## Strengths and Limitations

The strengths of this study include the use of a commercially available wearable device, which can be accessed by the general population to track their UVR exposure. This study also included both adults and children, who could both potentially benefit from future interventions using UVR monitoring devices, given that both populations are at risk for UVR exposure. Exploring adults' and children's perceptions of the device's acceptability and feasibility separately helped to elucidate the most important challenges that parents and children may experience when using such devices. For example, remembering to wear the device was a bigger issue for children than for adults, and this information can be considered when developing future interventions and app functions. A limitation of this study was the focus on self-reported compliance to wearing the device. In the future, devices that employ accelerometers or other technologies that objectively assess whether participants were wearing the device would be better monitors for study protocol adherence. Future studies could minimize the potential for missing data due to these factors by emphasizing to participants the importance of regular synchronizing and charging their devices to prevent loss of data. Finally, this study was conducted within a single geographic area, which may reduce the generalizability of the results. Two limitations of Shade include its lack of an accelerometer (used to track compliance) and the lack of a waterproof sensor, making it impossible to be worn while swimming.

## Conclusions

The findings of this study indicate that a wearable UVR monitoring device can be feasibly and acceptably used by both adults and children. The use of a wearable device to monitor UVR exposure is a unique and objective method for quantifying the amount of UVR exposure and could be used to support adults and children in reducing their UVR exposure. Ultimately, decreases in personal UVR exposure could contribute to the prevention of melanoma and other types of skin cancer in the future.

## Acknowledgments

The authors appreciate Peter Kaplan and Emmanuel Dumont for their guidance on and assistance with using Shade in this study. The authors also appreciate Jared Luther for his assistance with data management. This work was supported in part by the National Cancer Institute of the National Institutes of Health (NIH; K07CA196985; YW). This work was also supported in part by a Pilot Project Award from the American Cancer Society Huntsman Cancer Institute Institutional Research Grant (129785-IRG-16-190-01-IRG) and an NIH New Innovator Award (1DP2EB022360-01; JJ). Data for this project were collected using REDCap, which is supported by the National Cancer Institute of the NIH (8UL1TR000105, formerly UL1RR025764). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Themes endorsed by parents and children on the open-ended feasibility questions.
[DOCX File , 17 KB - derma_v3i1e15711_app1.docx ]

## References

1. Surveillance, Epidemiology, and End Results Program. Cancer Stat Facts: Melanoma of the Skin URL: http://seer.cancer.gov/statfacts/html/melan.html [accessed 2019-01-01]
2. Armstrong BK, Kricker A. How much melanoma is caused by sun exposure? Melanoma Res 1993 Dec;3(6):395-401. [doi: 10.1097/00008390-199311000-00002] [Medline: 8161879]
3. Whiteman DC, Whiteman CA, Green AC. Childhood sun exposure as a risk factor for melanoma: a systematic review of epidemiologic studies. Cancer Causes Control 2001 Jan;12(1):69-82. [doi: 10.1023/a:1008980919928] [Medline: 11227927]
4. Godar DE, Wengraitis SP, Shreffler J, Sliney DH. UV doses of Americans. Photochem Photobiol 2001 Jun;73(6):621-629. [doi: 10.1562/0031-8655(2001)073<0621:udoa>2.0.co;2] [Medline: 11421067]
5. Dobbinson S, Wakefield M, Hill D, Girgis A, Aitken JF, Beckmann K, et al. Children's sun exposure and sun protection: prevalence in Australia and related parental factors. J Am Acad Dermatol 2012 Jun;66(6):938-947. [doi: 10.1016/j.jaad.2011.06.015] [Medline: 21890234]
6. Hamilton K, Cleary C, White KM, Hawkes AL. Keeping kids sun safe: exploring parents' beliefs about their young child's sun-protective behaviours. Psychooncology 2016 Feb;25(2):158-163. [doi: 10.1002/pon.3888] [Medline: 26101815]
7. Johnson K, Davy L, Boyett T, Weathers L, Roetzheim RG. Sun protection practices for children: knowledge, attitudes, and parent behaviors. Arch Pediatr Adolesc Med 2001 Aug;155(8):891-896. [doi: 10.1001/archpedi.155.8.891] [Medline: 11483115]
8. Balk SJ, Council on Environmental Health, Section on Dermatology. Ultraviolet radiation: a hazard to children and adolescents. Pediatrics 2011 Mar;127(3):e791-e817. [doi: 10.1542/peds.2010-3502] [Medline: 21357345]
9. Geller AC, Colditz G, Oliveria S, Emmons K, Jorgensen C, Aweh GN, et al. Use of sunscreen, sunburning rates, and tanning bed use among more than 10 000 US children and adolescents. Pediatrics 2002 Jun;109(6):1009-1014. [doi: 10.1542/peds.109.6.1009] [Medline: 12042536]
10. Glanz K, Steffen AD, Schoenfeld E, Tappe KA. Randomized trial of tailored skin cancer prevention for children: the Project SCAPE family study. J Health Commun 2013;18(11):1368-1383 [FREE Full text] [doi: 10.1080/10810730.2013.778361] [Medline: 23806094]
11. Lucas RM, Valery P, van der Mei I, Dwyer T, Pender MP, Taylor B, Ausimmune Investigator Group. Sun exposure over a lifetime in Australian adults from latitudinally diverse regions. Photochem Photobiol 2013;89(3):737-744. [doi: 10.1111/php.12044] [Medline: 23311500]
12. Glenn BA, Lin T, Chang LC, Okada A, Wong WK, Glanz K, et al. Sun protection practices and sun exposure among children with a parental history of melanoma. Cancer Epidemiol Biomarkers Prev 2015 Jan;24(1):169-177 [FREE Full text] [doi: 10.1158/1055-9965.EPI-14-0650] [Medline: 25587110]
13. Stump TK, Aspinwall LG, Taber JM, Edwards R, Leachman SA. Protection-Adjusted Length of Exposure (PALE): A New Self-Report Measure of UVR Exposure That Captures Multiple Modalities of Photoprotection. In: Paper Presented at the Annual Meeting of the Melanoma Genetics Consortium. 2013 Presented at: GenoMEL'13; June 2013; Philadelphia, PA.
14. McCarty CA. Sunlight exposure assessment: can we accurately assess vitamin D exposure from sunlight questionnaires? Am J Clin Nutr 2008 Apr;87(4):1097S-1101S. [doi: 10.1093/ajcn/87.4.1097S] [Medline: 18400741]
15. O'Riordan DL, Stanton WR, Eyeson-Annan M, Gies P, Roy C. Correlations between reported and measured ultraviolet radiation exposure of mothers and young children. Photochem Photobiol 2000 Jan;71(1):60-64. [doi: 10.1562/0031-8655(2000)071<0060:cbramu>2.0.co;2] [Medline: 10649890]
16. O'Riordan DL, Lunde KB, Steffen AD, Maddock JE. Validity of beachgoers' self-report of their sun habits. Arch Dermatol 2006 Oct;142(10):1304-1311. [doi: 10.1001/archderm.142.10.1304] [Medline: 17043185]
17. Rampen FH, Fleuren BA, de Boo TM, Lemmens WA. Unreliability of self-reported burning tendency and tanning ability. Arch Dermatol 1988 Jun;124(6):885-888. [Medline: 3377517]
18. Grandahl K, Mortensen OS, Sherman DZ, Køster B, Lund P, Ibler KS, et al. Solar UV exposure among outdoor workers in Denmark measured with personal UV-B dosimeters: technical and practical feasibility. Biomed Eng Online 2017 Oct 10;16(1):119 [FREE Full text] [doi: 10.1186/s12938-017-0410-3] [Medline: 29017484]
19. Bodekær M, Petersen B, Philipsen PA, Heydenreich J, Thieden E, Wulf HC. Sun exposure patterns of urban, suburban, and rural children: a dosimetry and diary study of 150 children. Photochem Photobiol Sci 2015 Jul;14(7):1282-1289. [doi: 10.1039/c5pp00052a] [Medline: 26066309]

20. Thieden E, Agren MS, Wulf HC. Solar UVR exposures of indoor workers in a Working and a Holiday Period assessed by personal dosimeters and sun exposure diaries. Photodermatol Photoimmunol Photomed 2001 Dec;17(6):249-255. [doi: 10.1034/j.1600-0781.2001.170601.x] [Medline: 11722749]
21. Thieden E, Philipsen PA, Heydenreich J, Wulf HC. UV radiation exposure related to age, sex, occupation, and sun behavior based on time-stamped personal dosimeter readings. Arch Dermatol 2004 Feb;140(2):197-203. [doi: 10.1001/archderm.140.2.197] [Medline: 14967793]
22. Thieden E, Philipsen PA, Wulf HC. Ultraviolet radiation exposure pattern in winter compared with summer based on time-stamped personal dosimeter readings. Br J Dermatol 2006 Jan;154(1):133-138. [doi: 10.1111/j.1365-2133.2005.06961.x] [Medline: 16403106]
23. Hacker E, Horsham C, Vagenas D, Jones L, Lowe J, Janda M. A mobile technology intervention with ultraviolet radiation dosimeters and smartphone apps for skin cancer prevention in young adults: randomized controlled trial. JMIR Mhealth Uhealth 2018 Nov 28;6(11):e199 [FREE Full text] [doi: 10.2196/mhealth.9854] [Medline: 30487115]
24. King L, Xiang F, Swaminathan A, Lucas RM. Measuring sun exposure in epidemiological studies: Matching the method to the research question. J Photochem Photobiol B 2015 Dec;153:373-379. [doi: 10.1016/j.jphotobiol.2015.10.024] [Medline: 26555640]
25. Allen MW, McKenzie RL. National Institute of Water and Atmospheric Research. 2010. Electronic UV Dosimeters for Research and Education URL: https://niwa.co.nz/sites/niwa.co.nz/files/electronic_uv_dosimeters.pdf [accessed 2019-01-01]
26. Banerjee S, Hoch EG, Kaplan PD, Dumont EL. A Comparative Study of Wearable Ultraviolet Radiometers. In: Proceedings of the 2017 IEEE Life Sciences Conference. 2017 Presented at: LSC'17; December 13-15, 2017; Sydney, Australia. [doi: 10.1109/LSC.2017.8268131]
27. Webb AR. Measuring UV radiation: a discussion of dosimeter properties, uses and limitations. J Photochem Photobiol B 1995 Nov;31(1-2):9-13. [doi: 10.1016/1011-1344(95)07162-4]
28. Webb AR, Slaper H, Koepke P, Schmalwieser AW. Know your standard: clarifying the CIE erythema action spectrum. Photochem Photobiol 2011;87(2):483-486. [doi: 10.1111/j.1751-1097.2010.00871.x] [Medline: 21366601]
29. World Health Organization. Ultraviolet Radiation and Health URL: http://www.who.int/uv/uv_and_health/en/ [accessed 2019-01-01]
30. Lewis ZH, Lyons EJ, Jarvis JM, Baillargeon J. Using an electronic activity monitor system as an intervention modality: A systematic review. BMC Public Health 2015 Jun 24;15:585 [FREE Full text] [doi: 10.1186/s12889-015-1947-3] [Medline: 26104189]
31. Ridgers ND, McNarry MA, Mackintosh KA. Feasibility and effectiveness of using wearable activity trackers in youth: a systematic review. JMIR Mhealth Uhealth 2016 Nov 23;4(4):e129 [FREE Full text] [doi: 10.2196/mhealth.6540] [Medline: 27881359]
32. Køster B, Søndergaard J, Nielsen JB, Allen M, Bjerregaard M, Olsen A, et al. Feasibility of smartphone diaries and personal dosimeters to quantitatively study exposure to ultraviolet radiation in a small national sample. Photodermatol Photoimmunol Photomed 2015 Sep;31(5):252-260. [doi: 10.1111/phpp.12179] [Medline: 25850335]
33. Centers for Disease Control and Prevention. US Cancer Statistics Data Visualizations Tool URL: https://www.cdc.gov/cancer/uscs/dataviz/index.htm [accessed 2019-01-01]
34. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
35. Spagnolli A, Guardigli E, Orso V, Varotto A, Gamberini L. Measuring User Acceptance of Wearable Symbiotic Devices: Validation Study Across Application Scenarios. In: Proceedings of the International Workshop on Symbiotic Interaction. 2015 Presented at: Symbiotic'15; October 7-8, 2015; Berlin, Germany p. 87-98. [doi: 10.1007/978-3-319-13500-7_7]
36. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
37. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
38. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
39. Cancer Treatment Centers of America. Risk Factors for Skin Cancer URL: https://www.cancercenter.com/cancer-types/skin-cancer/risk-factors [accessed 2019-01-01]
40. Centers for Disease Control and Prevention. 2019. Skin Cancer: Behavior Rates: Sunburns URL: https://www.cdc.gov/cancer/skin/statistics/behavior/sunburns.htm [accessed 2019-01-01]
41. Stubbs B, Williams J, Gaughran F, Craig T. How sedentary are people with psychosis? A systematic review and meta-analysis. Schizophr Res 2016 Mar;171(1-3):103-109. [doi: 10.1016/j.schres.2016.01.034] [Medline: 26805414]
42. Shuk E, Burkhalter JE, Baguer CF, Holland SM, Pinkhasik A, Brady MS, et al. Factors associated with inconsistent sun protection in first-degree relatives of melanoma survivors. Qual Health Res 2012 Jul;22(7):934-945 [FREE Full text] [doi: 10.1177/1049732312443426] [Medline: 22645220]

43. Diffey B. Climate change, ozone depletion and the impact on ultraviolet exposure of human skin. Phys Med Biol 2004 Jan 7;49(1):R1-11. [doi: 10.1088/0031-9155/49/1/r01] [Medline: 14971768]
44. Linos E, Swetter SM, Cockburn MG, Colditz GA, Clarke CA. Increasing burden of melanoma in the United States. J Invest Dermatol 2009 Jul;129(7):1666-1674 [FREE Full text] [doi: 10.1038/jid.2008.423] [Medline: 19131946]
45. Ridgers ND, Timperio A, Brown H, Ball K, Macfarlane S, Lai SK, et al. Wearable activity tracker use among Australian adolescents: usability and acceptability study. JMIR Mhealth Uhealth 2018 Apr 11;6(4):e86 [FREE Full text] [doi: 10.2196/mhealth.9199] [Medline: 29643054]
46. Schaefer SE, Ching CC, Breen H, German JB. Wearing, thinking, and moving: testing the feasibility of fitness tracking with urban youth. Am J Health Educ 2016;47(1):8-16. [doi: 10.1080/19325037.2015.1111174]
47. Schaefer SE, van Loan M, German JB. A feasibility study of wearable activity monitors for pre-adolescent school-age children. Prev Chronic Dis 2014 May 22;11:E85 [FREE Full text] [doi: 10.5888/pcd11.130262] [Medline: 24854236]

## Abbreviations

**NIH:** National Institutes of Health
**REDCap:** Research Electronic Data Capture
**SED:** standard erythemal dose
**UVR:** UV radiation

---

*Edited by G Eysenbach; submitted 31.07.19; peer-reviewed by N Ridgers, A McGar; comments to author 01.10.19; revised version received 25.11.19; accepted 30.01.20; published 29.04.20.*

*Please cite as:*
*Nagelhout ES, Lensink R, Zhu A, Parsons BG, Jensen JD, Wu YP*
*The Feasibility and Acceptability of Using a Wearable UV Radiation Exposure Monitoring Device in Adults and Children: Cross-Sectional Questionnaire Study*
*JMIR Dermatol 2020;3(1):e15711*
*URL: http://derma.jmir.org/2020/1/e15711/*
*doi:10.2196/15711*
*PMID:33043275*

Original Paper

# Skin Lesion Classification With Deep Convolutional Neural Network: Process Development and Validation

Arnab Ray[1], BTech; Aman Gupta[1], BTech; Amutha Al[1], MEng

SRM Institute of Science and Technology, Chennai, India

**Corresponding Author:**
Arnab Ray, BTech
SRM Institute of Science and Technology
SRM Nagar, Kattankulathur
Chennai, 603203
India
Phone: 91 8939336693
Email: ad733943@gmail.com

## Abstract

**Background:** Skin cancer is the most common cancer and is often ignored by people at an early stage. There are 5.4 million new cases of skin cancer worldwide every year. Deaths due to skin cancer could be prevented by early detection of the mole.

**Objective:** We propose a skin lesion classification system that has the ability to detect such moles at an early stage and is able to easily differentiate between a cancerous and noncancerous mole. Using this system, we would be able to save time and resources for both patients and practitioners.

**Methods:** We created a deep convolutional neural network using an Inceptionv3 and DenseNet-201 pretrained model.

**Results:** We found that using the concepts of fine-tuning and the ensemble learning model yielded superior results. Furthermore, fine-tuning the whole model helped models converge faster compared to fine-tuning only the top layers, giving better accuracy overall.

**Conclusions:** Based on our research, we conclude that deep learning algorithms are highly suitable for classifying skin cancer images.

***(JMIR Dermatol 2020;3(1):e18438)*** doi:10.2196/18438

**KEYWORDS**

deep convolutional neural network; VGG16, Inceptionv3; Inception ResNet V2; DenseNet; skin cancer; cancer; neural network; machine learning; melanoma

## Introduction

### Skin Cancer

One in every three cancers diagnosed is skin cancer. Although melanomas represent fewer than 5% of all skin cancers, they account for approximately 75% of all skin cancer–related deaths and are responsible for over 10,000 deaths annually. Early detection of the mole would decrease the number of skin cancer deaths.

Skin cancer is significantly lower in India due to the presence of eumelanin in India’s dark-skinned population, which provides some protection against the development of skin cancer. Still, skin cancer constituted 3.18% of all patients with cancer in India. Of this, 54.76% were basal cell carcinomas, while 36.91% were squamous cell carcinoma and malignant melanoma was only 8.33%. The majority of patients were from rural areas (88%) and many were involved in agriculture (92%) [1].

### Neural Networks in the Context of Skin Cancer

We searched for research papers that used neural networks in the context of skin cancer from Google Scholar, PubMed, Research Gate, and the ISIC (International Skin Imaging Collaboration) archive. We included the results in the literature survey. Deep learning has solved many complex modern problems. The increasing amount of data on the internet helps in this process. There is a huge improvement in image classification using convolutional neural networks (CNN). The first few layers of deep CNN (DCNN) can learn the general features of an image, which can be used for different models. Using fine-tuning, DCNN models trained on one data set can be reused for image classification of other data sets. By fine-tuning Inceptionv3, Esteva et al [2] proposed that, “CNN achieves performance on par with all tested experts, demonstrating an artificial intelligence capable of classifying skin cancer with a level of competence comparable to dermatologists". Esteva and colleagues used their own obtained dermatologist-labelled data set consisting of 129,450 clinical images, including 3374 dermoscopy images. This data set includes 2032 skin diseases, belonging to 9 skin disease partitions. By fine-tuning Inceptionv3 on this data set, Esteva and colleagues achieved up to 66% accuracy classification on these 9 classes.

Another previously published study that used DCNN used AlexNet [3]. The data set consisted of 200 pictures. However, by image augmentation (ie, rotating all the pictures), 4400 images were made. This study used the transfer learning model, in which the AlexNet model was trained on ImageNet data, and the last layer was replaced with the softmax layer that is classified into melanoma, seborrheic keratosis, and nevus. For the change of weights, they used the stochastic gradient descent (SGD) algorithmic program. They were able to achieve an accuracy of 98%.

In another study, the authors planned a mechanized strategy for malignant melanoma determination connected to an arrangement of dermoscopy photos [4]. Highlights removed relied upon using a multilayer perceptron (MLP) classifier and coevent network to distinguish between melanocytic nevi and melanoma. The authors proposed two different procedures for MLP: programmed MLP and conventional MLP. Both techniques were useful for the separation of melanocytic carcinoma with a high accuracy. Following this, the arrangement procedure was executed with an MLP classifier that involved two strategies: automatic MLP and traditional MLP. The MLP classifier displayed distinctive grouping accuracy. The programmed MLP planned 93.4% and 76% training and testing accuracy, respectively.

A different study used a model that uses support vector machine (SVM) learning algorithms [5]. Their model did not use annotated information. The feature transfer that they used allowed the system to draw similarities between observations of dermoscopic pictures and that of the natural world. It mimics the method specialists use to explain patterns in skin lesions. Two-fold cross-validation was performed 20 times for analysis (40 experiments in total), and two discrimination tasks were examined: malignant melanoma versus atypical lesions, and malignant melanoma versus all nonmelanoma lesions. This approach achieved an accuracy of 93.1% for the primary task and 73.9% accuracy for the second task.

In another study, authors designed and modelled a system that can collect and combine past pigmented skin lesion (PSL) image results, their analysis, and corresponding observations and conclusions by medical experts, using a prototyping methodology [6]. One area of the system used computational intelligence techniques to research, process, and classify the images and their probable morphology. Trained medical personnel in remote locations can use mobile knowledge acquisition devices to take pictures of PSL and input the pictures into the planned system, which would classify the imaged PSL as malignant or benign.

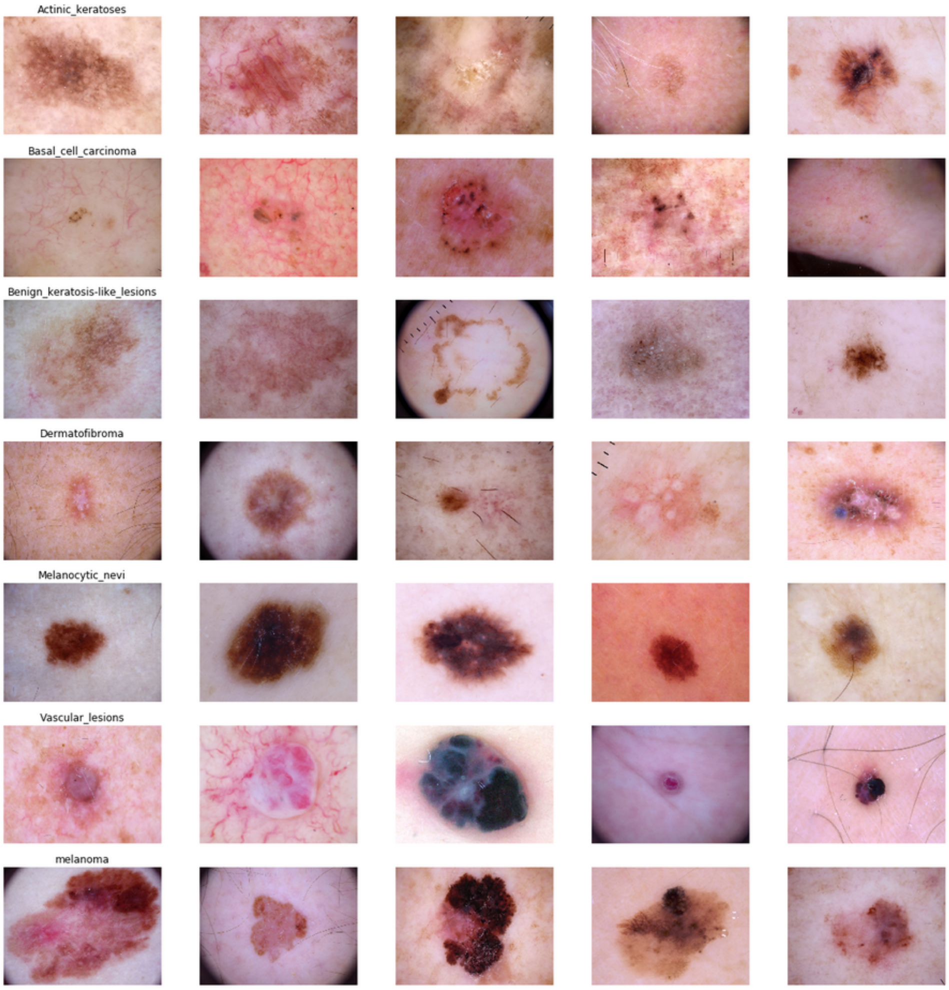
Another group used a similar concept using DCNN. They trained their model on a data set of 129,450 images. They used the Inceptionv3 architecture model and classified images among 757 different melanoma classes. The accuracy achieved was 72%; this value was relatively low due to the high number of classes in this data set [2].

Another study used lesion segmentation as the first step of processing [7]. They identified morphological features specific to certain lesions. Preprocessing steps included changing the color channel, smoothing the image, removing hairs, etc. They modelled the algorithm as a binary classification model (ie, benign or malignant). Lesion-related morphological features (including diameter, color, and magnification) were used as the input to a number of classifiers. The best accuracy (79%) was found with the k-nearest neighbors (KNN) algorithm.

In this project, we used the HAM10000 data set obtained by ViDIR Group, Department of Dermatology, Medical University of Vienna. Figure 1 shows example images from the data set that was used for this study.

In this study, we fine-tuned DCNNs and compared the performance of 4 DCNNs: VGG16, Inception-ResNet V2, Inceptionv3, and DenseNet-201. Each DCNN was fine-tuned from the top layers. Fine-tuning of all layers was performed with Inceptionv3 and DenseNet-201. Finally, we created an ensemble of Inceptionv3 and DenseNet-201 with all layers fine-tuned.

**Figure 1.** Example lesion photos from the HAM10000 data set (ViDIR Group, Department of Dermatology, Medical University of Vienna).



# Methods

## Exploratory Data Analysis

This step was performed to better understand the data and prepare the data for neural networks. In this project, we used the HAM10000 data set obtained by ViDIR Group, Department of Dermatology, Medical University of Vienna. The diagnostic accuracy for melanoma was significantly higher with dermoscopy compared to unaided eye diagnosis (respectively, log OR 4.0 [95% CI 3.0-5.1] versus log OR 2.7 [95% CI 1.9-3.4], an improvement of 49%, *P*<.001) [8]. The diagnostic accuracy solely depended on the experience and knowledge of the examiner.

We observed that this data set is biased toward melanocytic nevi, as seen in Table 1. Hence, in the worst-case scenario, our neural network model will have an accuracy higher than 60%.

All the original images (450×600 pixels) were resized to 64×4-pixel RGB images for the baseline model and 192×256 pixels for fine-tuning models. The data set was split into 7210 training examples, 1803 validation examples, and 1002 test examples.

**Table 1.** Counts for each type of lesion in the data set

| Type of lesion | Number of images |
|---|---|
| Melanocytic nevi | 6705 |
| Melanoma | 1113 |
| Benign keratosis | 1099 |
| Basal cell carcinoma | 514 |
| Actinic keratoses | 325 |
| Vascular lesions | 142 |
| Dermatofibroma | 115 |

## Baseline Model

We built a baseline CNN to estimate the difficulty of the problem. Our architecture consisted of 6 layers: (1) a convolutional layer with 16 kernels each of size 3 and padding such that the size of the image is maintained, (2) a max-pooling layer with 2×2 window, (3) a convolutional layer with 32 kernels each of size 3 and padding to maintain size, (4) a max-pooling layer with 2×2 window, (5) a convolutional layer with 64 kernels each of size 3 and padding to maintain size, and (6) a max-pooling layer with 2×2 window.

To train the model, data augmentation was required. The learning rate was initialized at 0.01 and Adam Optimizer was used. The baseline model was trained for a total of 35 epochs.

## VGG16 Model

VGG16 is a convolutional neural net architecture (Figure 2 [9]) that won the ImageNet competition in 2014 and is generally regarded as one of the best current vision models architecture. Even though it is an old model, we chose VGG16 because of its simplicity.

**Figure 2.** VGG16 architecture.

On the ImageNet data set, VGG16 achieved an accuracy of 90.1% for top-5 and 71.3% for top-1.

Data augmentation was performed to increase the data set image count. Fine-tuning was performed on the model by removing the top, fully-connected layers that were then replaced with following: (1) a max-pooling layer, (2) a fully connected layer with 512 units, (3) a dropout layer with 0.5 rate, and (4) a softmax activation layer for 7 types of skin lesions.

The first step included freezing all layers in VGG16 and performing feature extraction for newly added layers. After 3 epochs, we unfroze the final convolutional block of VGG16 and started fine-tuning a model for 20 epochs. The learning rate was set to 0.001 and Adam Optimizer was used. VGG16 was fine-tuned for a total of 30 epochs.

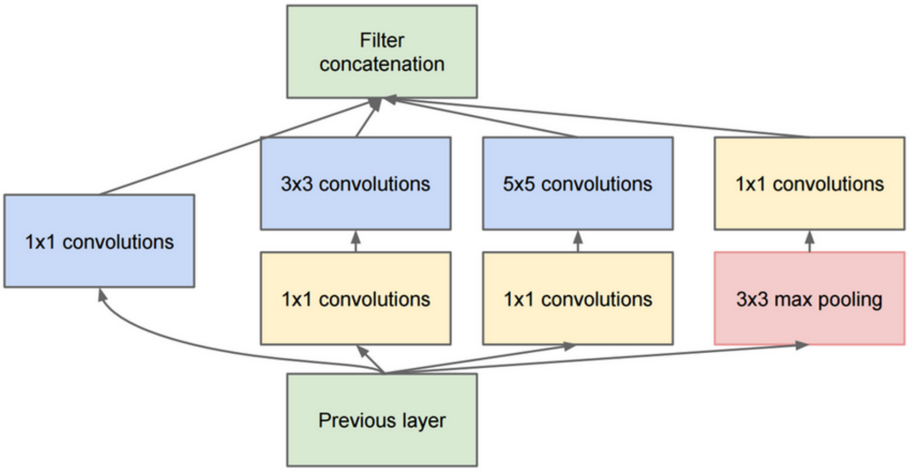
## Inception Model

Inceptionv3 produced an accuracy of 93.7% for top-5 and 77.9% for top-1 on the ImageNet data set. The Inception module has 1×1, 3×3, and 5×5 convolutions, all in parallel (Figure 3 [10]). The intention was to let the network decide, through training, what information would be learned and used. It also allows for multi-scale processing; the model can recover low-level features via small convolutional layers and high-level features with large convolutional layers.

We fine-tuned all layers of Inceptionv3 and the top two inception blocks with batch normalization layers. Inceptionv3 was fine-tuned for 20 epochs.

Additionally, we tried Inception-ResNet, a variant of Inception. It uses a residual connection, which has become necessary for training very deep convolutional models. The same training strategy used for Inceptionv3 was used for Inception-ResNet.

**Figure 3.** Inceptionv3 architecture. Published with permission.

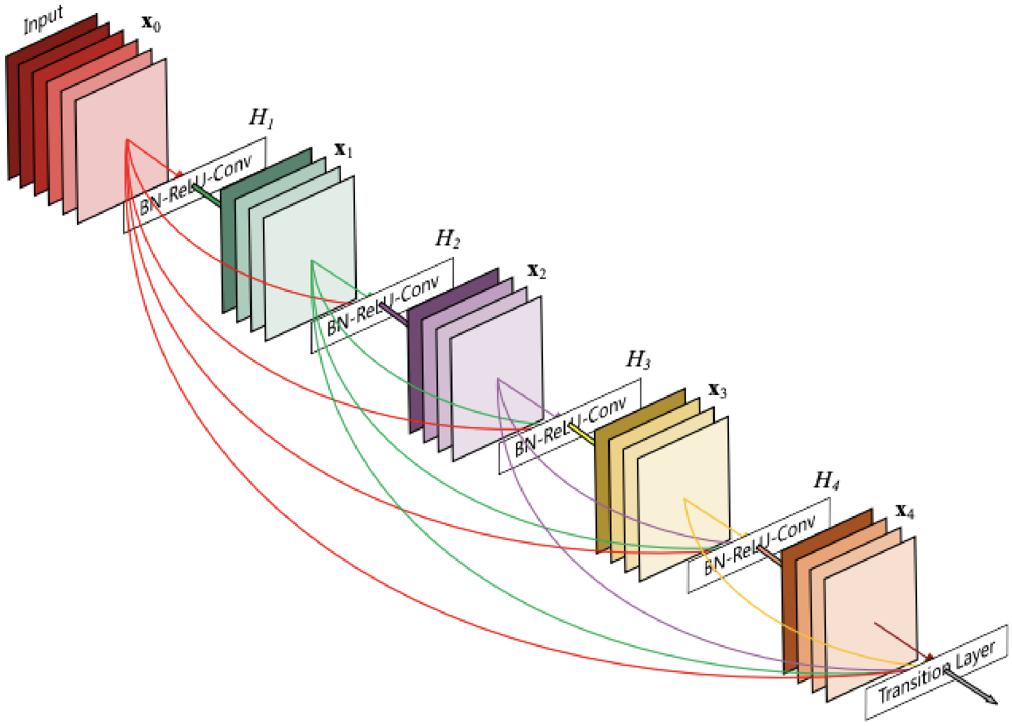


### DenseNet Model

This is a new architecture that performed exceptionally well in the ImageNet data set competition, giving an accuracy of 93.6% in top-5 and 77.3% on top-1. DenseNet has 4 dense blocks and uses approximately 20 million parameters (Figure 4 [11]).

In a dense block, one layer generates feature maps through a composite function, consisting of three consecutive operations: batch normalization, ReLU (rectified linear activation unit), and a 3×3 convolution. We used DenseNet-201, which uses 4 dense blocks, and we performed two types of fine-tuning on it: (1) fine-tuning on the last dense block (32 layers; Part A), and (2) fine-tuning on the whole network (Part B). Part A was trained for 27 epochs and Part B was trained for 20 epochs.

**Figure 4.** DenseNet architecture. Published with permission.



## Results

Table 2 shows the classification results from each model when the top layers were fine-tuned (Part A). Table 3 displays the classification results for each model when all layers were fine-tuned. All experiments were performed on a laptop with GPU NVIDIA 1050Ti. To speed up processing times, Google Colab (P100 GPU) was used.

From training a custom model, it was clear that the problem cannot be solved by a simple CNN model with a few layers. Therefore, we incorporated fine-tuning of the pretrained model. By hypertuning the pretrained model that had over 100 layers, we achieved better results. Fine-tuning all layers (Part B) gave us better results than fine-tuning only the top layers (Part A). Crucially, Part B was trained for fewer epochs, which helped the model converge faster. However, in both cases, DenseNet gave us better results than Inceptionv3. Using the concepts of ensemble learning, we created an ensemble of Inceptionv3 and DenseNet-201. This combination achieved a further improved accuracy of 88.8% on the validation set and 88.5% on the test set.

**Table 2.** Fine-tuning the top layers.

| Model | Validation (%) | Test (%) | Test loss | Depth (layers) |
|---|---|---|---|---|
| Custom model | 77.48 | 76.54 | 0.646671 | 11 |
| VGG16 | 79.82 | 79.64 | 0.708 | 23 |
| Inceptionv3 | 79.935 | 79.94 | 0.7482 | 315 |
| Inception-ResNet V2 | 80.82 | 82.53 | 0.6691 | 784 |
| DenseNet-201 | 85.8 | 83.9 | 0.691 | 711 |

**Table 3.** Fine-tuning all layers.

| Model | Validation (%) | Test (%) | Test loss |
|---|---|---|---|
| Inceptionv3 | 86.92 | 86.826 | 0.6241 |
| DenseNet-201 | 86.696 | 87.725 | 0.5587 |
| Ensemble (Inceptionv3 and DenseNet-201) | 88.8 | 88.52 | 0.41156 |

## Discussion

Our results indicate that deep learning algorithms are highly suitable for classifying skin cancer images. Additionally, by using the concepts of fine-tuning and the ensemble learning model, improved results were achieved. Finally, we found that fine-tuning the whole model helped the model converge faster compared with fine-tuning only the top layers, giving an overall better accuracy.

### Conflicts of Interest

None declared.

### References

1. Lal ST, Banipal RP, Bhatti DJ, Yadav HP. Changing Trends of Skin Cancer: A Tertiary Care Hospital Study in Malwa Region of Punjab. J Clin Diagn Res 2016 Jun;10(6):PC12-PC15 [FREE Full text] [doi: 10.7860/JCDR/2016/18487.8051] [Medline: 27504344]
2. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017 Jan 25;542(7639):115-118. [doi: 10.1038/nature21056]
3. Hosny KM, Kassem MA, Foaud MM. Skin Cancer Classification using Deep Learning and Transfer Learning. 2018 Presented at: 9th Cairo International Biomedical Engineering Conference (CIBEC); December 20-22; Cairo, Egypt p. 90-93. [doi: 10.1109/CIBEC.2018.8641762]
4. S.Mabrouk M, A.Sheha M, Sharawy A. Automatic Detection of Melanoma Skin Cancer using Texture Analysis. IJCA 2012 Mar 31;42(20):22-26. [doi: 10.5120/5817-8129]
5. Codella N, Cai J, Abedini M, Garnavi R, Halpern A, Smith JR. Deep Learning, Sparse Coding, and SVM for Melanoma Recognition in Dermoscopy Images. In: Machine Learning In Medical Imaging. Cham: Springer; 2015:118-126.
6. Okuboyejo DA, Olugbara OO, Odunaike SA. Automating Skin Disease Diagnosis Using Image Classification. 2013 Presented at: Proceedings of the World Congress on Engineering and Computer Science; October 23-25; San Francisco.
7. Codella N, Rotemberg V, Tschandl P, Celebi ME, Dusza S, Gutman D, et al. Skin lesion analysis toward melanoma detection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC). 2018 Presented at: IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018); April 4; Washington, DC, USA.
8. Kittler H, Pehamberger H, Wolff K, Binder M. Diagnostic accuracy of dermoscopy. Lancet Oncol 2002 Mar;3(3):159-165. [doi: 10.1016/s1470-2045(02)00679-4] [Medline: 11902502]
9. VGG16 architecture. 2020 Jun 02. Classez et segmentez des données visuelles URL: https://tinyurl.com/ya49hbpk
10. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2015 Jun [FREE Full text]
11. Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely connected convolutional networks. arXiv 2016 Aug 25 [FREE Full text]

### Abbreviations

**CNN:** convolutional neural network
**DCNN:** deep convolutional neural network
**ISIC:** International Skin Imaging Collaboration
**KNN:** k-nearest neighbor
**MLP:** multilayer perceptron
**PSL:** pigmented skin lesion

*Edited by G Eysenbach; submitted 26.02.20; peer-reviewed by A Chitranshi, K Ray; comments to author 03.03.20; revised version received 03.03.20; accepted 21.03.20; published 07.05.20.*

*Please cite as:*
*Ray A, Gupta A, Al A*
*Skin Lesion Classification With Deep Convolutional Neural Network: Process Development and Validation*
*JMIR Dermatol 2020;3(1):e18438*
*URL: http://derma.jmir.org/2020/1/e18438/*
*doi:10.2196/18438*
*PMID:*

Original Paper

# Information About Sunscreen on YouTube and Considerations for Sun Safety Promotion: Content Analysis

Anne K Julian[1], MA, PhD; Jessica Welch[1], BS; Maddison M Bean[2], MS; Sarah Shahid[1], BS; Frank M Perna[1], EdD, PhD

[1]Behavioral Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Rockville, MD, United States
[2]College of Public Health and Human Sciences, Oregon State University, Corvallis, OR, United States

**Corresponding Author:**
Anne K Julian, MA, PhD
Behavioral Research Program
Division of Cancer Control and Population Sciences
National Cancer Institute
9609 Medical Center Dr
Rockville, MD, 20877
United States
Phone: 1 240 276 6175
Email: anne.julian@nih.gov

## Abstract

**Background:** Sunscreen use is a popular sun protection method; however, application of sunscreen rarely meets the standards recommended for effectiveness. Access to information about how to effectively use sunscreen may play a role in proper sunscreen application. The internet is a common health information source; however, the quality of sunscreen-related content varies.

**Objective:** The objective of this study was to examine information about sunscreen in YouTube videos by video source.

**Methods:** In November 2017, the authors identified the 20 most popular YouTube videos (sorted by view count and relevance) for each of these 5 search terms: sunscreen cancer, sunscreen health, sunscreen information, sunscreen ingredients, and sunscreen natural. The inclusion criteria were English language and view count >1000 (N=111 unique videos). We double-coded videos for standard recommendations for sunscreen use (eg, apply 20 minutes before sun exposure), use of outdated terminology, and recommendation of complementary sun safety strategies.

**Results:** The view counts ranged from 1100 to 671,142 (median 17,774, SD 109,651) and the average daily views ranged from 1 to 1448 (median 23, SD 234). End users (46/111, 41.4%) and health care providers (24/111, 21.6%) were the most common sources, and none of the most popular videos were produced by federal agencies or cancer-related nongovernmental organizations. Health care provider videos included marginally more recommendations than end user videos (mean 1.46, SD 1.96 vs mean 1.05, SD 1.20), but few (19/111, 17.1%) mentioned reapplication. The videos were generally positive toward sunscreen (82/111, 73.9%); however, some videos were negative (29/111, 26.1%), with warnings about the health risks of chemical sunscreens and their ingredients. Do-it-yourself sunscreen tutorials represented 19/111 (17.1%) of the sample.

**Conclusions:** YouTube is a potential source for disseminating sun safety messages; however, the quality of its sunscreen content varies. Most of the videos in our study failed to include important sunscreen use recommendations. Clinicians should be prepared to address the information needs of patients by discussing effective, evidence-based sunscreen application and recommending a combined sun safety approach.

***(JMIR Dermatol 2020;3(1):e14411)*** doi:10.2196/14411

**KEYWORDS**

sunscreen; social media; sun safety; cancer prevention

## Introduction

Skin cancer affects an estimated 5 million people annually in the United States, with a treatment cost of over $8.1 billion [1]. Ultraviolet (UV) radiation is a known carcinogen that causes direct and indirect DNA damage; intentional or unintentional UV exposure is responsible for the majority of skin cancer incidence [2]. Sunscreen is a primary sun protection strategy that is used by approximately one third of US adults [3]; however, sunscreen application rarely reflects the standards necessary to meet the advertised protection levels [4,5]. Furthermore, sunscreen is intended for use in combination with other sun protection strategies [6]. In 2012, the US Food and Drug Administration (FDA) issued new regulations for sunscreen labeling [7]; however, research has demonstrated that sunscreen application rarely reflects the thickness, evenness, and timing (of both initial application and reapplication) necessary to meet the advertised protection level of the sunscreen [8]. In 2014, the Surgeon General's Call to Action to Prevent Skin Cancer included a renewed call to promote a combination sun protection strategy (eg, wearing sunscreen, a hat, and protective clothing) [5].

The Centers for Disease Control and Prevention (CDC) and the American Academy of Dermatology support the daily use of sunscreen with a sun protection factor (SPF) of at least 15 and 30, respectively, for UV protection, and dermatologists report recommending the use of sunscreen to their patients [9]. Although health care providers are trusted sources of health information, individuals increasingly supplement these sources with information garnered from the internet and social media [10]. In 2012, 72% of US adults had sought health information on the internet within the past 12 months [10]. The growing influence of social media in health decision making is not equally distributed across demographic and age groups; young age is a consistent predictor of social media use [11]; however, social media usage among adults is steadily increasing [12,13]. YouTube, the most prominent user-generated video-sharing website, allows any person to upload content on any topic [10]; based on traffic, it was the top-ranked website on the internet as of 2019 [14]. Furthermore, YouTube is increasingly being used as a platform to disseminate health information [15]. The American Cancer Society, National Institutes of Health, and National Cancer Institute have adopted this outlet to communicate health messages to broad audiences.

Despite these efforts, misleading or incorrect health information is common on YouTube [15]. For example, over one half of videos on the topic of immunization feature negative views or contradict the reference standard, which may lead viewers to make unsafe health decisions [16]. With respect to skin cancer, Basch et al [17] found YouTube content promoting untested home remedies for skin cancer, such as black salve; this finding highlights the challenge of health communication on social media. Moreover, a study of sun safety topics on YouTube [18] found that 17% of videos retrieved using the search term "sunscreen" included false or misleading information; however, less than a quarter of the videos in their sample were English language videos. Sunscreen use in practice is less thorough than recommended, and the question remains whether this is due in part to incomplete or misleading information available on social media platforms such as YouTube; this is especially likely to be the case if the content is not derived from governmental or professional medical organizations who create sun safety recommendations. Here, we analyzed a large sample of English language YouTube videos about sunscreen to address the following research questions: What primary sources, content types, positive/negative product attributions (valence), and adherence to standard recommendations [7] are included in YouTube videos about sunscreen? Do these videos contain the terms "sunblock" and "waterproof," which the FDA prohibits on sunscreen packaging? What complementary sun protection strategies do these videos include?

These data may inform public health video messaging to improve sun safety practices for effective sunburn prevention.

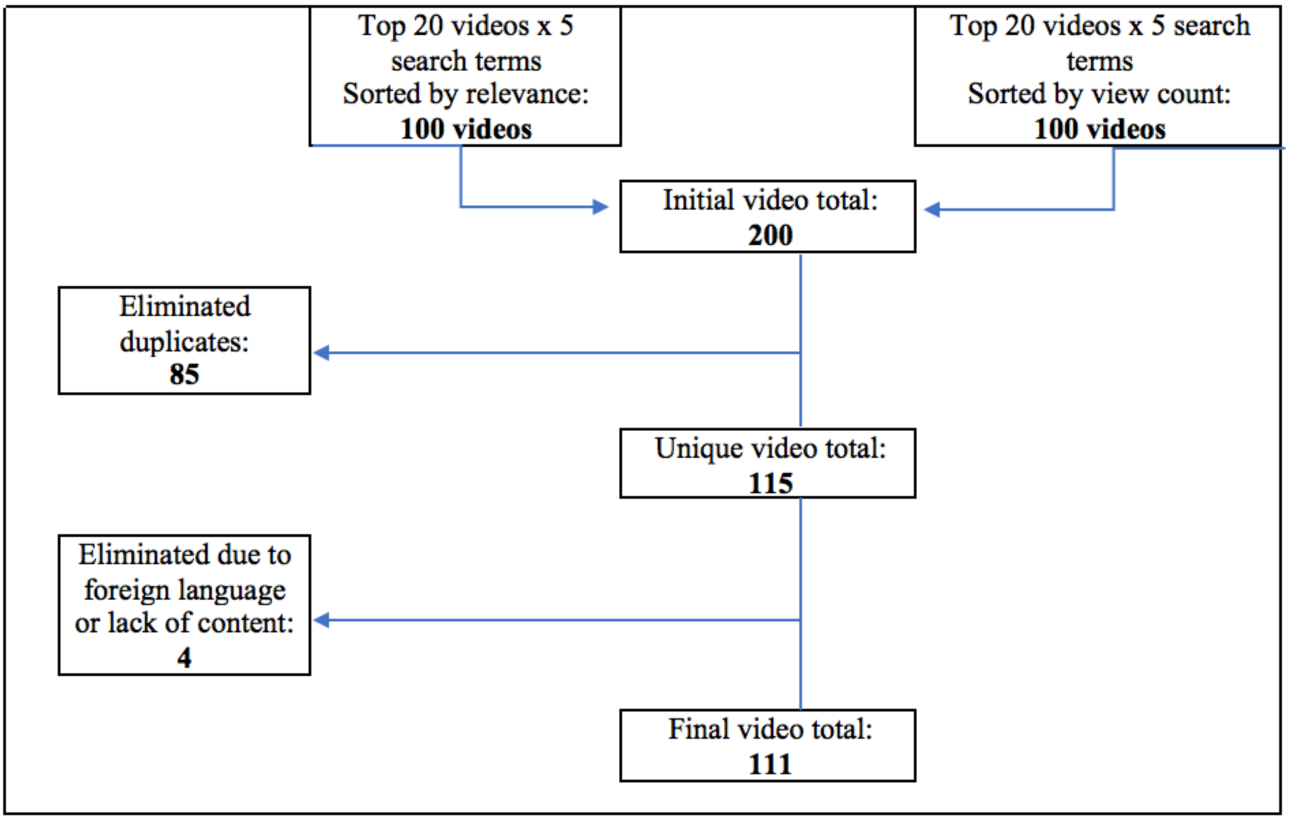
## Methods

We conducted direct YouTube searches between November 13 and November 15, 2017 (search terms: sunscreen cancer, sunscreen health, sunscreen information, sunscreen ingredients, and sunscreen natural), deleting cookies after each search to avoid influencing subsequent results [19]. The YouTube region was set to United States. We identified the 20 most frequently appearing videos for each search term and removed duplicates to obtain the analytic sample of videos (see Figure 1).

We developed an early draft codebook in consultation with sun safety and communication experts to reflect common metrics in YouTube-based health information studies and sunscreen recommendation reference standards. This draft was then expanded and refined to include the recommendation of nonsunscreen sun safety methods, use of the outdated terms "sunblock" and "waterproof," and possible themes related to sunscreen safety. The codebook was pilot-tested using a test set of 5 videos by a team that included 3 experts in sun safety and 1 postbaccalaureate fellow. These videos were chosen using an extreme cases approach and reflected a diverse range of content and positions toward sunscreen. The codebook development team met to discuss and resolve coding discrepancies, generate additional constructs and codes, and refine coding response options.

After the initial codebook refinement, the coding team (MG, SS, JW, and AJ) double-coded all videos for the presence of 6 standard recommendations for effective sunscreen use based on FDA and CDC recommendations: SPF level, broad spectrum classification, application timing (20 minutes before exposure), liberal/generous application, and reapplication (every 2 hours and separately after toweling). We coded the overall valence toward sunscreen as negative or positive. We coded the use of the terms "sunblock" and "waterproof"; since 2011, the FDA has prohibited these terms on sunscreen labels, as they may be misleading [7]. We further coded the video source based on the posting source of the video and the posting channel's thumbnail and description. Lastly, we coded recommendation of the following complementary sun protection strategies (operationalized as instruction to adopt): seeking shade, wearing a hat, wearing protective clothing, wearing sunglasses, and staying indoors during peak UV hours (Cohen kappa range 0.80-0.96). Viewer attention to social media content is short [20]; therefore, we stopped coding each video after 5 minutes. Any final disagreement was resolved through discussion. We used Stata/SE 14 [21] to calculate descriptive statistics (means, medians and standard deviations, ranges and percentages) overall and stratified by video source.

**Figure 1.** YouTube video search flow chart depicting the number of unique videos retrieved from YouTube in November 2017 using the search terms sunscreen information, sunscreen health, sunscreen natural, sunscreen ingredients, and sunscreen cancer and the final number of videos meeting the English language requirement. Relevance is a proprietary search algorithm owned by YouTube.



## Results

The videos had view counts of 1100 to 671,142 (mean 58,642, median 17,774, SD 109,651), were between 40 seconds and 41 minutes long (mean 5:54 minutes, median 4:19 minutes), and were uploaded between 2009 and 2017 (average post duration 1127 days, median 1101, range 64-3135, SD 784). Average daily views ranged from 1 to 1448 (mean 96.8, median 23, SD 234). Primary video sources were health care providers or their associations (24/111, 21.6%), end users (51/111, 45.9%), and corporate entities (15/111, 13.5%) (see Table 1 for the descriptive characteristics of the videos). No government health agencies were among the sources of videos retrieved. Valence toward sunscreen use was generally positive (82/111, 73.9%); however, positive content included promotion of mineral sunscreens that warned about health risks of chemical sunscreens. Videos that discouraged sunscreen use (29/111, 26.2%) included misleading content, such as overstatement of the benefits of vitamin D, and outright falsehoods, such as the claim that sunscreen causes cancer.

Of the 6 recommendations for effective sunscreen use, the most frequently mentioned were SPF (35/111, 31.5%) and broad-spectrum classification (30/111, 27.0%); meanwhile, application timing, generous application, reapplication every 2 hours, and reapplication after toweling were each mentioned in fewer than 20% of videos (see Table 2). Most videos included 0, 1, or 2 recommendations. SPF was mentioned in 5/9 (55%) of news and television coverage videos. Corporate entities, such as sunscreen or skin care companies, most frequently recommended generous application (4/14 videos, 29%); however, 6/14 (43%) of these videos paired this message with a poor demonstration of sunscreen application.

Few videos advocated the use of complementary sun protection strategies in conjunction with sunscreen (19/111, 17.1%). Of videos with positive valence toward sunscreen, 15/82 (18.3%) suggested complementary sun protection strategies, compared to 3/29 (10.3%) of videos with negative valence (see Table 2).

Few videos cited scientific sources for the presented information (30/111, 27.0%). Sources of scientific information were the Environmental Working Group (EWG), an environmental advocacy group (13/111, 11.7%), the FDA (7/111, 6.3%), and peer-reviewed articles (5/111 videos, 4.5%). Lastly, an unexpected finding was that 19/111 (17.1%) of videos demonstrated do-it-yourself recipes for how to make sunscreen at home.

**Table 1.** Video content types, sources, and characteristics (N=111).

| Video characteristic | Frequency, n (%) |
|---|---|
| **Content type** | |
| Public service announcement | 4 (3.6) |
| Medical advice | 24 (21.5) |
| Scientific explanation | 12 (10.8) |
| Product demonstration/promotion | 14 (12.6) |
| Do-it-yourself sunscreen tutorial | 19 (17.1) |
| Opinion/commentary | 23 (20.7) |
| Televised news clip | 15 (13.5) |
| **Source** | |
| Nongovernmental organization | 5 (4.5) |
| News organization | 9 (8.1) |
| Health care provider[a] | 24 (21.6) |
| End user | 51 (45.9) |
| Corporation | 15 (13.5) |
| Topic-based group channel | 7 (6.3) |
| **Use of outdated terms[b]** | |
| Sunblock | 13 (11.7) |
| Waterproof | 5 (4.5) |
| **Valence[c]** | |
| Positive | 82 (73.9) |
| Negative | 29 (26.1) |
| **Total number of recommendations** | |
| 0 | 54 (48.6) |
| 1 | 20 (18.0) |
| 2 | 19 (17.1) |
| 3 | 9 (8.1) |
| 4 | 3 (2.7) |
| 5 | 4 (3.6) |
| 6 | 2 (1.8) |

[a] "Health care provider" was broadly defined to include physicians, nurses, and their organizations as well as people who self-identified as doctors but whose credentials did not include a medical degree. Videos featuring a doctor as the primary messenger were coded for the credentials of the provider; these included allopathic doctors (MD, DO), nonmedical doctors (PhD), and complementary or alternative medicine practitioners (DC, ND, MD(H)).

[b] In videos using the terms "sunblock" and "waterproof," 5 and 2 respectively used it in the context of an explanation that the term is considered misleading and should no longer be used.

[c] Videos coded as positive toward sunscreen include those that promoted the use or home manufacture of mineral or "natural" sunscreens but that also included warnings about potential negative effects of chemical or commercial sunscreens on humans or the environment.

**Table 2.** Standard recommendations for sunscreen use, overall and by video source[a].

| Recommendation | Overall (N=111) | Health care provider (n=24) | Corporation (n=14) | News (n=9) | End user (n=47) |
|---|---|---|---|---|---|
| **Sunscreen application, n (%)** | | | | | |
| Choose a broad spectrum | 30 (27) | 5 (22) | 2 (14) | 4 (44) | 14 (30) |
| SPF level 15+ | 35 (32) | 8 (36) | 3 (21) | 5 (55) | 16 (37) |
| Apply 20 minutes before exposure | 14 (13) | 6 (26) | 2 (14) | 1 (11) | 2 (3) |
| Reapply every 2 hours | 16 (14) | 5 (22) | 2 (14) | 1 (11) | 2 (3) |
| Apply a generous amount | 19 (17) | 5 (22) | 4 (29) | 1 (11) | 5 (10) |
| Reapply after toweling | 12 (11) | 4 (17) | 2 (14) | 1 (11) | 2 (3) |
| **Complementary sun protection, n (%)** | | | | | |
| Wear a wide-brimmed hat | 1 (4) | 1 (4) | 1 (0) | 2 (22) | 2 (4) |
| Wear protective clothing | 12 (11) | 1 (4) | 0 (0) | 3 (33) | 3 (6) |
| Wear sunglasses | 8 (7) | 1 (4) | 0 (0) | 2 (22) | 1 (2) |
| Seek shade/use an umbrella | 11 (10) | 0 (0) | 0 (0) | 3 (33) | 3 (6) |
| Stay indoors during peak UV periods | 6 (5) | 0 (0) | 0 (0) | 1 (11) | 1 (2) |

[a] Due to small cell sizes, videos from nongovernmental organizations and sources coded as “Other,” which included topic-based group channels and sources that were not classifiable, were omitted from these group analyses. Therefore, the rows do not total 100%.

## Discussion

### Principal Findings

Incomplete or misleading information about sunscreen is common on YouTube and may play a role in sunscreen misuse, as reported in the literature [4,5]. In light of research on the quality of videos covering other areas of health information, the variable accuracy we found with respect to sunscreen use is not surprising [9]. Consistent with previous literature, no videos produced by government health agencies or cancer-related NGOs were among the most popular or relevant [13]. The EWG was the most-cited source of scientific information, but these video references were typically decontextualized statements about possible harms of common sunscreen chemicals (eg, oxybenzone and octinoxate) used in pitches to promote and sometimes sell other products. The outdated terms “sunblock” and “waterproof” were rarely used, but the videos generally failed to include all standard recommendations for effective sunscreen use or to recommend a combined sun protection approach. In the absence of instruction on application thickness, reapplication, and complementary sun protection measures, the emphasis we observed on SPF and broad-spectrum classification may perpetuate ineffective sunscreen use.

### Unexpected Findings

Some videos contradicted sun safety recommendations, warned about health risks or dangerous ingredients of sunscreen, or included do-it-yourself recipes. Although these videos were not the majority, this finding is surprising and somewhat alarming in that some of these videos featured health care–affiliated spokespersons. These providers represented a subset of health care providers on YouTube and were primarily promoters of alternative or naturopathic medicine. The do-it-yourself sunscreen tutorial was a concerning and unexpected content category, as our search strategy did not target this content. Do-it-yourself sunscreen recipes have not undergone the testing that is required by the FDA and cannot be assumed to be effective. Some “recipes” included no approved sunscreen ingredients and instead promoted ingredients such as green tea and coconut oil. Clinicians should be aware of such videos and should be prepared to address the topic with patients who are considering making their own sunscreen. While recent research on sunscreen absorption [22] may raise public concern, especially about the use of sunscreen by children, the extant recommendations are unchanged; we can conclude from available evidence that sunscreen is safe and that its use is superior to unprotected exposure for children and adults (for infants under 6 months of age, sunscreen is not recommended; rather, sun exposure should be avoided completely) [23]. Guidance such as use of a mineral sunscreen, sun avoidance strategies, and use of protective clothing and hats [23] may be particularly useful to parents who are concerned about using sunscreen products on children.

### Limitations

The YouTube videos analyzed here represent one point in time. Videos that are highly popular at one point may fade in popularity over time and may therefore not be prominent in subsequent searches. In our approach, we intentionally used neutral search terms associated with sunscreen and then gauged the valence of the most popular and relevant videos. Positive or negative search terms may yield different types of content; also, the search algorithms such as “relevance” used by YouTube are proprietary and may change over time, thus yielding different search results. Sunscreen videos are a narrow topic, and the most popular videos still do not qualify as generally popular, as viral videos have millions of views. Furthermore, not all adults use sunscreen, and individuals seeking information about sunscreen specifically from government health agencies may not use the YouTube application programming interface and may instead use Google or search the agency's website. The number of videos contradicting standard recommendations for sun safety or demonstrating do-it-yourself sunscreen was unanticipated; however, the behavioral significance of this finding is unclear.

## Conclusion

Sunscreen use is a common and important form of sun protection, and YouTube is a common source of health information; it is especially popular among teens and young adults, which are the developmental groups with the highest rates of sunburn [3]. This study fills an important gap in the literature by offering a comprehensive examination of sunscreen information on YouTube, a common source of information for young people. The public currently bears the burden of judging the credibility of sunscreen messages; however, a challenge inherent in our findings is that health care providers depicted in videos (who are generally trusted sources of information) were not, as a rule, arbiters of high-quality sunscreen information. While the popular videos sampled were generally positive toward sunscreen use, the information content appears to be problematic in 2 ways. First, sunscreen-related content on YouTube infrequently recommends complementary sun safety strategies. This is particularly problematic when videos explicitly discourage sunscreen use [24]. Second, a small but not unimportant percentage of popular videos contained misleading information.

YouTube is a promising tool to share information on effective sunscreen use; however, incomplete or negative content about sunscreen is common, and in some cases, this content is communicated by health care–affiliated messengers. Further, health agencies that produce YouTube videos related to sun protection were not among the more popular or relevant video sources. These data highlight a challenge of using YouTube for public health communication: how to create eye-catching content that will reliably be identified by social media search algorithms. To better harness YouTube for disseminating sun safety messages, health agencies may look to increase their content visibility by better incorporating features such as memorable titles and keyword tags to increase their popularity. Considering the increasing popularity of YouTube as a source of health information, further research is warranted to examine viewer response to these videos and to determine whether exposure to the videos impacts subsequent sun safety behavior.

## Acknowledgments

The authors would like to acknowledge the contributions of Anne Hartman, who assisted in codebook development.

## Conflicts of Interest

None declared.

## References

1. Guy GP, Machlin SR, Ekwueme DU, Yabroff KR. Prevalence and costs of skin cancer treatment in the U.S., 2002-2006 and 2007-2011. Am J Prev Med 2015 Feb;48(2):183-187 [FREE Full text] [doi: 10.1016/j.amepre.2014.08.036] [Medline: 25442229]
2. Armstrong BK, Kricker A. The epidemiology of UV induced skin cancer. J Photoch Photobio B 2001 Oct;63(1-3):8-18. [doi: 10.1016/s1011-1344(01)00198-1]
3. Holman DM, Ding H, Guy GP, Watson M, Hartman AM, Perna FM. Prevalence of Sun Protection Use and Sunburn and Association of Demographic and Behaviorial Characteristics With Sunburn Among US Adults. JAMA Dermatol 2018 May 01;154(5):561-568 [FREE Full text] [doi: 10.1001/jamadermatol.2018.0028] [Medline: 29541756]
4. Pratt H, Hassanin K, Troughton LD, Czanner G, Zheng Y, McCormick AG, et al. UV imaging reveals facial areas that are prone to skin cancer are disproportionately missed during sunscreen application. PLoS One 2017;12(10):e0185297 [FREE Full text] [doi: 10.1371/journal.pone.0185297] [Medline: 28968413]
5. US Dept of Health and Human Services, Office of the Surgeon General. 2014. The Surgeon General's Call to Action to Prevent Skin Cancer URL: https://www.hhs.gov/sites/default/files/call-to-action-prevent-skin-cancer.pdf [accessed 2020-04-16]
6. Morris KL, Perna FM. Decision Tree Model vs Traditional Measures to Identify Patterns of Sun-Protective Behaviors and Sun Sensitivity Associated With Sunburn. JAMA Dermatol 2018 Aug 01;154(8):897. [doi: 10.1001/jamadermatol.2018.1646]
7. 21CFR201.: U.S. Food and Drug Administration Code of Federal Regulations Title 21 URL: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?fr=201.327 [accessed 2019-10-25]
8. Petersen B, Wulf HC. Application of sunscreen--theory and reality. Photodermatol Photoimmunol Photomed 2014;30(2-3):96-101. [doi: 10.1111/phpp.12099] [Medline: 24313722]
9. Farberg AS, Glazer AM, Rigel AC, White R, Rigel DS. Dermatologists' Perceptions, Recommendations, and Use of Sunscreen. JAMA Dermatol 2017 Jan 01;153(1):99. [doi: 10.1001/jamadermatol.2016.3698]
10. Koch-Weser S, Bradshaw YS, Gualtieri L, Gallagher SS. The Internet as a health information source: findings from the 2007 Health Information National Trends Survey and implications for health communication. J Health Commun 2010 Dec;15 Suppl 3:279-293. [doi: 10.1080/10810730.2010.522700] [Medline: 21154099]

11. Chou WS, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009 Nov;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
12. Perrin A. Social Media Usage: 2005-2015. 2005. Pew Research Center 2015 Oct 8 URL: https://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/ [accessed 2019-10-25]
13. Mitchell SJ, Godoy L, Shabazz K, Horn IB. Internet and Mobile Technology Use Among Urban African American Parents: Survey Study of a Clinical Population. J Med Internet Res 2014 Jan 13;16(1):e9. [doi: 10.2196/jmir.2673]
14. Alexa. The top 500 sites on the web URL: https://www.alexa.com/topsites [accessed 2019-10-25]
15. Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: A systematic review. Health Informatics J 2015 Sep;21(3):173-194. [doi: 10.1177/1460458213512220] [Medline: 24670899]
16. Keelan J, Pavri-Garcia V, Tomlinson G, Wilson K. YouTube as a source of information on immunization: a content analysis. JAMA 2007 Dec 5;298(21):2482-2484. [doi: 10.1001/jama.298.21.2482] [Medline: 18056901]
17. Basch CH, Basch CE, Hillyer GC, Reeves R. YouTube Videos Related to Skin Cancer: A Missed Opportunity for Cancer Prevention and Control. JMIR Cancer 2015 Mar 02;1(1):e1. [doi: 10.2196/cancer.4204]
18. Ruppert L, Køster B, Siegert A, Cop C, Boyers L, Karimkhani C, et al. YouTube as a source of health information: Analysis of sun protection and skin cancer prevention related issues. Dermatol Online J 2017 Jan 15;23(1) [FREE Full text] [Medline: 28329465]
19. Sampson M, Cumber J, Li C, Pound CM, Fuller A, Harrison D. A systematic review of methods for studying consumer health YouTube videos, with implications for systematic reviews. PeerJ 2013 Sep;1:e147 [FREE Full text] [doi: 10.7717/peerj.147] [Medline: 24058879]
20. Does Video Length Matter?. URL: https://wistia.com/learn/marketing/does-length-matter-it-does-for-video-2k12-edition [accessed 2020-04-16]
21. StataCorp. Stata Statistical Software: Release 14. College Station, TX: Statacorp LP; 2015. URL: https://www.stata.com/stata14/ [accessed 2020-04-16]
22. Matta MK, Zusterzeel R, Pilli NR, Patel V, Volpe DA, Florian J, et al. Effect of Sunscreen Application Under Maximal Use Conditions on Plasma Concentration of Sunscreen Active Ingredients. JAMA 2019 Jun 04;321(21):2082. [doi: 10.1001/jama.2019.5586]
23. Holman DM, Qin J, Gottschlich EA, Balk SJ. Clinical counseling on sun protection and indoor tanning avoidance: A survey of current practices among U.S. health care providers. Prev Med 2019 Sep;126:105783. [doi: 10.1016/j.ypmed.2019.105783]
24. Bleakley A, Lazovich D, B Jordan A, Glanz K. Compensation Behaviors and Skin Cancer Prevention. Am J Prev Med 2018 Dec;55(6):848-855. [doi: 10.1016/j.amepre.2018.06.019] [Medline: 30344038]

---

## Abbreviations

**CDC:** Centers for Disease Control and Prevention
**EWG:** Environmental Working Group
**FDA:** Food and Drug Administration
**SPF:** sun protection factor
**UV:** ultraviolet

---

*Edited by G Eysenbach; submitted 18.04.19; peer-reviewed by C Basch, E Warner, K Glanz; comments to author 01.10.19; revised version received 20.11.19; accepted 22.03.20; published 02.06.20.*

*Please cite as:*
*Julian AK, Welch J, Bean MM, Shahid S, Perna FM*
*Information About Sunscreen on YouTube and Considerations for Sun Safety Promotion: Content Analysis*
*JMIR Dermatol 2020;3(1):e14411*
*URL: http://derma.jmir.org/2020/1/e14411/*
*doi:10.2196/14411*
*PMID:*

Original Paper

# Associations Between Emotions, Social Media Use, and Sun Exposure Among Young Women: Ecological Momentary Assessment Study

Jessica Fitts Willoughby[1], PhD; Jessica Gall Myrick[2], PhD; Stephanie Gibbons[1], MA; Clark Kogan[3], PhD

[1]The Edward R Murrow College of Communication, Washington State University, Pullman, WA, United States
[2]Donald P Bellisario College of Communications, Pennsylvania State University, University Park, PA, United States
[3]Center for Interdisciplinary Statistical Education and Research, Washington State University, Pullman, WA, United States

**Corresponding Author:**
Jessica Fitts Willoughby, PhD
The Edward R Murrow College of Communication
Washington State University
208 Goertzen Hall
PO Box 642520
Pullman, WA, 99164-2520
United States
Phone: 1 5093357926
Email: jessica.willoughby@wsu.edu

## Abstract

**Background:** Research has pointed to a connection between social media use, emotions, and tanning behaviors. However, less is known about the role specific emotions may play in influencing social media use and how emotions and social media use may each be associated with outdoor tanning.

**Objective:** This paper aims to examine the connection between emotions, social media use, and outdoor tanning behaviors among young women, a group particularly important for skin cancer prevention efforts.

**Methods:** We used ecological momentary assessment to collect data from 197 women aged 18 to 25 years 3 times a day for 7 days in July 2018. We collected data from women in 2 states.

**Results:** We found that boredom was associated with increased time spent on social media and that increased time spent on social media was associated with increased time spent outdoors without sun protection.

**Conclusions:** Our results highlight that social media may be a particularly important channel for skin cancer prevention efforts targeting young women, as more social media use was associated with increased time spent outdoors with skin exposed. Researchers should consider the role of emotions in motivating social media use and subsequent tanning behaviors. Additionally, as boredom was associated with social media use, intervention developers would benefit from developing digital and social media interventions that entertain as well as educate.

***(JMIR Dermatol 2020;3(1):e18371)*** doi:10.2196/18371

**KEYWORDS**

social media; skin cancer; emotions; cancer prevention; health communication; ecological momentary assessment

## Introduction

### Background

It is estimated that more than 100,000 new cases of melanoma, the deadliest type of skin cancer, will be diagnosed in the United States in 2020 [1], and melanoma is one of the most prevalent cancers in young women [2]. Younger skin is particularly vulnerable to the effects of ultraviolet light and the skin damage it may cause [3]. Despite the health implications of exposure to ultraviolet light, young women often express a desire to have tan skin and believe having a tan makes them look healthier [4]. Taken together, these statistics and preferences point to an increased risk of skin cancer among young women.

As such, researchers are eager to better understand the psychosocial and behavioral factors that predict risky skin behaviors like tanning. Previous research has found that emotions are associated with indoor tanning behaviors in young women [5], as is young women's use of social media [6]. Additionally, research has found that intentional outdoor tanning is significantly greater among girls who spend more time on their cell phones [7].

One psychosocial predictor that may help predict tanning is emotions, which may be particularly important for understanding both media use and tanning behaviors for a number of reasons. Appraisal theory suggests that specific emotions may be linked with action tendencies based on automatic subjective evaluations [8]. For example, a person experiences happiness when they have gained something they desire. The action tendencies associated with happiness include moving toward the happiness-inducing event, sharing positive outcomes, and celebration [8,9]. However, there is a lack of understanding about the role different emotions might play in motivating social media use and outdoor tanning, a risk behavior related to skin cancer.

In this manuscript, we examine the associations among emotions and social media use, emotions and risky outdoor skin behaviors, and social media use and outdoor tanning behavior using ecological momentary assessment. This method allows us to see how emotions, media use, and tanning behavior interrelate in real time throughout the course of a week. This may be important, as previous research has found that emotions vary throughout the day in response to stressors (eg, minor hassles that happen throughout the day in people's lives) [10]. Therefore, capturing emotion data at a single point in time may substantially increase the noise, thereby reducing the power. Additionally, digital media use is often frequent and dynamic, leading to challenges in determining how to best measure media exposure [11]. Previous research in other topic areas has found that retrospective self-report in communication may be difficult [12]. As the use of ecological momentary assessment is beneficial for collecting data and may better assess complex or dynamic data [13], we used 3 daily surveys over the course of a week to assess social media use, emotions, and related outcomes among young women.

## Emotions and Social Media

Research has unearthed both positive and negative consequences of using social media, warranting a perpetual need to understand the evolving media platform [14]. A great deal of social media research has focused on the linear effect that social media may have on one's emotions [15,16]. This includes outcomes such as increased depression, anxiety, need of belonging, empathy, social desirability, social support, and others. However, the relationship between social media use and one's emotions is often cyclical, with certain emotions and personality types perpetuating social media consumption [14].

For example, rather than observing how social media use leads to loneliness, it is important to also observe that users turn to social media as a substitute for social interaction [16]. Certain personality traits and emotions, such as extroversion and levels of instability, are associated with social media use and willingness to interact [17]. Forgas [18] argues that emotions influence every aspect of an individual's life and stimulate each behavior in some aspect. As such, emotions influence social media use, too. Motives for social media use are unique, since they encompass "a diverse and complex set of (often unconscious) decision-making activities, influenced by a large number of individual, collective, and environmental factors" [14]. Positive outcomes of social media use include entertainment or connectivity, while negative emotional outcomes may include users' feelings that they are wasting time or distracting themselves from more meaningful social interaction [14]. These emotional experiences have been shown to predict social media use, especially with regard to information sharing, entertainment, and convenience [19-21]. Research has found that young adults use social media for a variety of reasons, including information [22], entertainment and leisure [23], socialization [24], and keeping in contact with others [23]. The previously discussed research leads to our first research question (research question 1): Which emotions are associated with the increased use of social media among young women?

## Emotions and Tanning

Many individuals have positive emotional associations with the sun [25]. In fact, researchers have identified a physiological link between exposure to UV light and the feel-good chemicals released by our bodies, called endorphins [26]. As such, it may not be surprising that ultraviolet exposure can make people feel good. Expectations or beliefs that tanning will result in a positive emotional state have been consistently found to predict both indoor and outdoor tanning [27-29]. Conversely, individuals who have negative emotional associations with tanning (eg, they think they will be uncomfortable while doing it or are anxious about the health risks) are less likely to perform the behavior [30].

According to the broaden-and-build theory, positive emotions exist to motivate individuals to broaden the scopes of their attention and skills while also building resources and relationships to help them cope with future negative situations [31]. Tanning is often a social activity done with friends and reported to be an enjoyable, positive experience [25]. However, positive emotions can also reinforce behaviors that elicit them, even if individuals know the behavior is risky [32]. As such, the positive emotional association that individuals build with tanning behaviors may continue to motivate them to pursue this cancer-causing behavior to maintain a positive emotional state or alleviate a negative one.

This supposition is also in line with mood management theory, which argues that people often seek stimuli (be it a behavior or a media message) that help them achieve a more positive mood [33,34]. For example, recent work using the National Cancer Institute's Family Life, Activity, Sun, Health, and Eating (FLASHE) cross-sectional survey found that adolescents who reported being frequent outdoor tanners were more likely to experience loneliness than adolescents who reported they did not tan outdoors [7]. Although more research is needed to understand this relationship, it could be that adolescents in the survey tanned outdoors in order to reap the positive emotional benefits that could help them cope with feelings of loneliness. This lack of clear consensus around the role of positive emotions and tanning behaviors led to the following research question (research question 2): Are more positive emotions associated with increased time spent outdoors without sun protection among young women?

### Social Media and Tanning

Although there is some understanding of how social media may influence well-being, researchers have called for more research to be conducted related to digital natives and social media's impact on behavior, health, and society [23]. Research has found that media use is connected to a variety of tanning behaviors [4,6,35,36]. Of particular note, young women may desire to look like people in the media [37]. In US culture, tan skin is a commonly held beauty standard, and women often engage in outdoor or indoor tanning to achieve this standard. Although previous research has observed how media sources such as television [38] and magazines [35,36] influence tanning behaviors, a recent focus has been on social media [6,7]. Social media expands on traditional media platforms, as there are greater opportunities to interact within the media and exchange images. Additionally, social media is a channel frequently used by young women [39]. Social media has been found to have an indirect but positive relationship with indoor tanning use among young women, while traditional mass media has not [5]. Such research purports that social media may perpetuate the desire to engage in tanning behaviors [5,40].

A theoretical rationale for the effects of social media on tanning behaviors comes from social cognitive theory [41,42], which posits that people can learn from behaviors modeled in media. Although media may not always directly model tanning behavior, people who are rewarded or viewed positively and are tan could influence young women to aim to achieve a similar look. Social media has been found to be related to the belief that indoor tanning leads to positive outcomes, such as improved appearance, greater mood, and improved health, while simultaneously being negatively associated with health and social consequences [5].

In part, this relationship may also have to do with the connections and relationships that can occur around tanning. Researchers have found tanning to be a social experience, as indoor tanners are likely to discuss tanning with friends [43]. In a study that looked at different platforms and their connections to tanning behaviors, Willoughby and Myrick [6] found that image-based social media platforms, such as Snapchat, Instagram, and Pinterest, were all positively associated with sunbathing, suggesting that the inclusion of images of tanned or tanning people may encourage young women to spend time sunbathing. Based on the previously discussed literature, we propose the following hypothesis: Increased social media use will be associated with increased time spent outdoors without sun protection among young women.

## *Methods*

### Study Design

We conducted an ecological momentary assessment (EMA) of young women's emotions, social media use, and outdoor tanning behaviors over a 7-day period in July 2018. EMA is a data collection method that allows for real-time data collection. It uses mobile devices to ask participants to report on moods or behaviors at specific points in time close to when the events occurred. Consequently, EMA has potential advantages over traditional data collection methods, including its usefulness for measuring mood and behavior [44].

### Recruitment

Participants were recruited from 2 universities in the United States, one in the Pacific Northwest and one in the Northeast. A mix of in-person and online methods were used to recruit young adult college women at the participating universities. If they were interested and met inclusion criteria (ie, identified as a woman and were aged 18-29 years), participants signed up via a text message program to receive study updates. All participants received the EMA over the same 7-day period in July 2018.

### Procedures

Participants received 3 messages a day. The first assessment was sent at 10 AM and asked participants about their current emotions. The second assessment was sent at 4 PM and asked about social media use over the past 6 hours, time spent outdoors, and current emotions. The third assessment was sent at 8 PM and asked about social media use over the past 4 hours, current emotions, time spent outdoors, and the weather. We selected the timing of the surveys to fall around peak UV exposure, which occurs between 10 AM and 4 PM [45], with an end time of 8 PM to avoid having the surveys sent too late in the evening for some participants. At the end of the week, participants were asked to complete a longer survey that asked about demographics, attitudes, and behaviors. For completing at least half of the EMA surveys and the posttest, participants received US $50 in cash. Prior to conducting the research, the first author's institutional review board approved all procedures.

### Measures

Individual emotions were measured at 3 time points by asking participants to "rate how much of each emotion describes how you are feeling" on a 5-point Likert-type scale from "none of this emotion at all" to "a great deal of this emotion." The emotions assessed included the following: content, excited, sad, envious, anxious, and bored. We selected these specific emotions based on previous research that examined emotions and tanning [27,46,47]. We examined both the discrete emotions and a composite measure of positivity. Composite measures of positivity have been used in previous research [48]. We created the composite measure to indicate whether a person was experiencing a greater level of positive emotions by adding together contentment and excitement and then subtracting sadness, envy, anxiety, and boredom. The emotions data were summarized to obtain the average of each emotional state over all 7 days and over all 3 time points. The creation of the composite measure is similar to composite creation measures in other work on emotion regulation [49].

Social media was measured at 2 time points daily (second and third survey) with the question "How often have you used the following social media since the last survey?" on a slider scale with half-hour increments marked. The options were Instagram, Facebook, Twitter, Snapchat, and text messaging. As we were focused on the role of social media collectively for this study, we added together the amount of time participants reported spending on Instagram, Facebook, Twitter, and Snapchat. We calculated the sum of the hours spent on social media over the day divided by the proportion of the daily hours that were observed via survey response. For instance, if only morning social media usage was recorded, then the total amount of time on social media was divided by 0.6, as the time period represents 60% of the time participants reported social media use. If participants responded to both surveys, then it would be the total number of hours observed that day.

Sun exposure was measured with the question “How much time did you spend outdoors with your skin exposed?” on a slider scale with half-hour increments marked at both 4 PM and 8 PM. If there was an observation at 8 PM but no observation at 4 PM (missing data), we used the amount of time outdoors divided by 0.4 (as the 4 PM to 8 PM time period represents 40% of the overall time they were questioned about outdoor tanning). If data were available at both times, we summed the amount of time to determine total time outdoors. The time periods used were selected because they encompass the peak sun hours in which participants could have spent time outdoors. The item is similar to items used in previous research that considered time spent outside without sun protection as a measure of sun exposure [50,51], but it included specific time frames to take advantage of the multiple measurements in EMA.

Perceived weather was measured with 3 items that asked participants to “please indicate how strongly you disagree or agree with the following statements regarding the weather today” on a 5-point Likert-type scale: “It was sunny outside,” “It was warm outside,” and “The weather was nice.” We averaged the 3 weather variables to obtain a new weather variable that indicated whether the weather was nice (Cronbach α=.853).

### Analysis Strategy

We collected data from 197 individuals, but due to missing data, we had 192 participants in the data frame that was used to address research question 1 and 179 participants in the data frame that was used to address research question 2.

To assess research question 1, which asked which specific emotions were associated with increased use of social media, we fit a linear model relating the logarithm of the average daily amount of time (out of the 10 hours assessed) that participants used social media to the average emotional state of an individual. Emotions were included in the regression using natural cubic splines in order to allow for a nonlinear relationship between the emotional scale and the log of the social media usage. We computed an adjusted generalization of the Spearman rank correlation to motivate the number of knots to use for each cubic spline. Correlations between each emotion and overall social media use were small for all variables except boredom. We therefore considered a restricted cubic spline for boredom (with 3 knots) and included all other terms as linear. Between 3 and 5 knots is typically found to be sufficient for a spline, and as our correlation between boredom and social media use was only moderate, 3 knots was selected [52]. We constructed residual plots to assess the plausibility of the normality and variance homogeneity assumption and found little evidence to suggest major deviation. Wald *F* statistics were computed to summarize the effect of each emotion on social media use.

To assess research question 2, which asked if increased positive emotions were associated with increased time spent outdoors without sun protection, and hypothesis 1, which posited that increased use of social media would be associated with increased time spent outdoors without sun protection, we performed a generalized least squares regression to relate daily time outdoors to positivity and social media use after adjusting for weather and accounting for correlation within individuals via the use of a compound symmetric correlation structure. As all correlations between variables were very small, we considered only linear terms in our regression analysis. We prefer generalized least squares to mixed-effects modeling, as it has a certain elegance for continuous responses [52]. We used the square root transformation for the time spent outdoors to help improve the plausibility of residual normality. Residuals were assessed for the plausibility of normality and variance homogeneity. Mild deviations from normality were found, but they were not substantial enough to warrant further transformations or a change in the distributional assumptions. Furthermore, there was little evidence to suggest more than minor deviations from variance homogeneity. We conducted Wald tests to assess the overall effect of positivity, weather, and social media on time spent outdoors.

## Results

### Sample

We collected demographic information at the end with a posttest survey (N=149). Some people did not respond to all surveys; as such, we examined the sample descriptive statistics by each analysis frame for comparison, finding minimal differences. We report on the descriptive data for individuals who completed the posttest survey.

On average, participants were aged 20.5 (SD 6.81) years. A total of 69.8% (104/149) of participants were from Washington state, and 26.2% (39/149) of participants were from Pennsylvania; 4.0% (6/149) reported being from another state. Nearly two-thirds (110/149, 73.8%) of participants reported being White or Caucasian. The majority of participants reported spending time outdoors in the sun a few times per week (65/149, 43.6%) or once per week (31/149, 20.8%). In addition, 36.2% (54/149) of participants had indoor tanned at some time in their life. Table 1 presents additional demographic information.

**Table 1.** Demographic information of study participants.

| Variable | Participants, n (%)[a] |
|---|---|
| **Race** | |
| White or Caucasian | 110 (74) |
| Asian or Pacific Islander | 19 (13) |
| Hispanic or Latina | 16 (11) |
| Black or African American | 9 (6) |
| American Indian or Native American | 4 (3) |
| Other | 1 (1) |
| **Grade** | |
| Freshman | 37 (25) |
| Sophomore | 19 (13) |
| Junior | 42 (28) |
| Senior | 39 (27) |
| Other or not applicable | 12 (7) |
| Not diagnosed with skin cancer | 149 (100) |
| Someone close ever diagnosed with skin cancer (yes) | 64 (43) |
| **Area in which they live** | |
| Rural (eg, in the country) | 44 (30) |
| Urban (eg, in the city) | 20 (13) |
| Suburban (eg, near a city, outside a city) | 79 (53) |
| Indoor tan ever (yes) | 54 (36) |

[a] Descriptive data included 149 individuals, as some participants did not complete the posttest questionnaire.

## Research Questions and Hypothesis Testing

For research question 1, we found evidence suggesting an association between one or more emotions and social media use ($F_{7,184}$=5.94; $P$<.001). Specifically, boredom was associated with increased social media use ($F_{2,7}$=14.70; $P$<.001). To interpret the effect size of boredom, we fit a simplified model in which boredom was reduced to a linear effect. We found an estimated increase of 67% (95% CI 37%-103%) in social media usage for every 1-unit increase in boredom (Figure 1). We assessed the evidence that the effect of boredom on social media use was nonlinear and found no evidence against linearity ($P$=.10). We did not find evidence that levels of contentment, excitement, sadness, envy, or anxiousness were associated with social media use with the current sample (Table 2).

**Figure 1.** Average daily social media use (in hours) plotted against average boredom. The blue line shows the model predictions after back transformation across different boredom ratings, computed at the mean value for all other emotions.

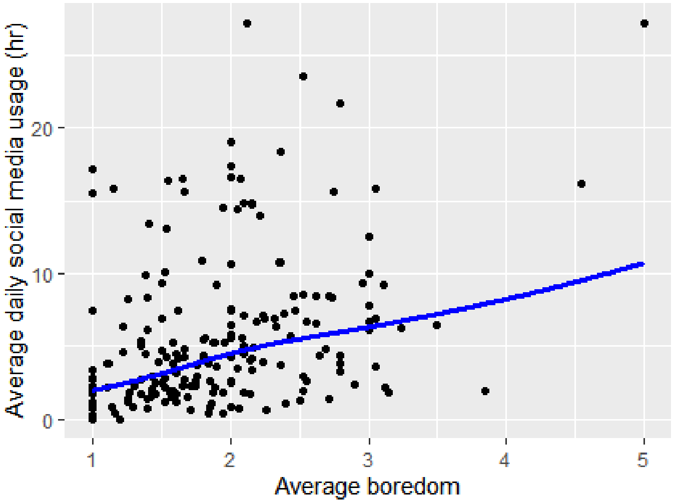


**Table 2.** Social media use analysis of variance.

| Emotion | Partial sum of squares | Mean squares | *F* value (*df*) | *P* value |
|---|---|---|---|---|
| Bored | 23.80 | 11.90 | 14.70[a] (2) | <.001 |
| Content | 0.53 | 0.53 | 0.66 (1) | .42 |
| Excited | 0.94 | 0.94 | 1.16 (1) | .28 |
| Sad | 0.38 | 0.38 | 0.46 (1) | .50 |
| Envious | 1.42 | 1.42 | 1.76 (1) | .19 |
| Anxious | 0.34 | 0.34 | 0.41 (1) | .52 |

[a] Significant at the *P*<.001 level.

For research question 2, there was evidence suggesting some association between time spent outdoors and the explanatory variables of positivity, weather, and social media (N=184; $\chi^2_3$=17.9; *P*<.001). Further Wald tests found slight evidence of an effect of positivity (*P*=.08) and strong evidence for a modest effect of social media (*P*<.001) on time spent outdoors without sun protection, providing support for hypothesis 1 (Figure 2 and Table 3).

**Figure 2.** Average daily time outdoors (in hours) plotted against average daily social media use. The blue line shows the model predictions after back transformation across different social media usage levels, computed at the mean value for positivity and weather.

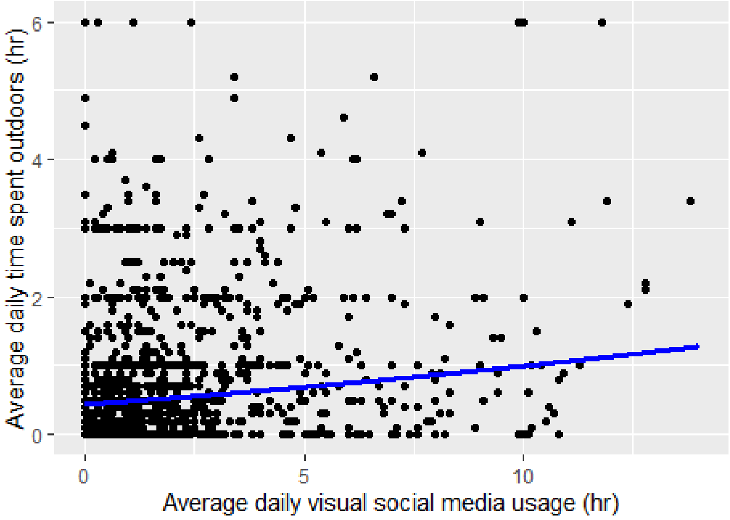


**Table 3.** Results of Wald analysis of variance on increased amount of time spent outdoors.

| Factor | Chi-square (*df*) | *P* value |
|---|---|---|
| Positivity | 3.1 (1) | .08 |
| Weather | 3.7 (1) | .06 |
| Social media | 13.9 (1) | <.001 |

## Discussion

### Principal Findings

This research used an EMA study conducted with young women over the course of a week in July to assess the relationships between emotions, social media use, and time spent outdoors. Of the emotions we assessed, we found only boredom associated with increased use of social media, with increased reports of boredom associated with increased time spent on social media. The magnitude of the effect was substantial, with boredom at greater levels nearly doubling social media use. Additionally, increased use of social media was significantly associated with increased time spent outdoors without sun protection. Below, we discuss these findings in greater depth.

Boredom, a low-arousal negative emotional state, occurs when an individual is unable or unwilling to dedicate their attention to whatever is currently in front of them [53]. Another definition of boredom is monotony combined with frustration [54]. Individuals experiencing boredom have been found to visit more websites during an internet search task than overstimulated, stressed individuals [55]. In our study, increased levels of boredom were associated with increased social media use, highlighting how individuals may use social media as a distraction or way to entertain themselves [23]. If bored individuals turn to social media that portray tan celebrities and peers, it could then motivate them to tan as another means of shifting their emotional state to something more pleasant.

Of the other emotions we assessed, including contentment, excitement, envy, anxiety, and sadness, none were significantly associated with social media use in our analyses. There may be a few different reasons for this. First, according to the appraisal theory of emotions, different emotions predispose people to take certain actions (called action tendencies) [8]. However, those action tendencies are typically directed at the underlying cause of that emotion. For example, if an individual is afraid of skin cancer, the action tendency is to avoid the threat somehow (possibly by not tanning). Since we measured emotions more generally (ie, not directly related to tanning), they may have predicted a number of potential actions more generally too. Future work could collect emotional responses with a specific target.

Second, previous research has found that the use of specific social media sites (ie, Facebook) for surveillance was associated with depression, with envy mediating the relationship. However, no direct effects between social media use and depression were found [56]. Additional research has found similar links, with envy and social comparison mediating the association between Facebook use and depression [57]. As these studies highlight, there are potential mediators and moderators that may be associated with social media use and emotions. Perhaps different, unaccounted for variables could help connect emotions to social media.

Lastly, it is possible that, as we used 3 daily surveys 4 or more hours apart, our time measurement wasn’t specific enough to see the resulting impact of emotions on social media use except with boredom, which provided a strong enough result to lead to increased social media use at the time of measurement. Future research could benefit from varying the timing to see if emotions then impact resulting social media use.

In addition to finding an association between boredom and subsequent social media use, we found that increased use of social media was associated with increased time spent outdoors without sun protection. This supports previous research, which has found a connection between social media and tanning behaviors [5-7]. Americans frequently use social media, with young adults (aged 18-24 years) using a variety of platforms and most of them (71%) visiting the platforms multiple times each day [39]. Nearly 3 out of 4 young people use Instagram, a highly visual form of social media, with young women often reporting frequent use. The content to which young women are exposed could be showcasing the benefits of tanning behaviors. In line with social cognitive theory [42], when individuals see behaviors that are modeled and rewarded, they may be more likely to engage in such behaviors. Additionally, as peers and socialization may impact tanning behaviors [43,58], social media may play an especially important role. This supports other research that has highlighted the possibility of using social media for skin cancer prevention efforts [59].

Taken together, these findings have implications for skin cancer prevention strategies among young women. In particular, social media is a well-suited platform for prevention efforts, as individuals who spend time in the sun without sun protection are also likely to use social media. Additionally, as young women turn to social media when bored, strategies that employ some form of entertainment may be particularly well suited to engaging the audience and garnering attention toward prevention efforts. Strategies such as entertainment education, which is the inclusion of educational content in entertaining formats [60], have been found to be effective for a variety of health topics related to risky behaviors [61,62]. Entertainment education efforts targeting young women in digital media have been found to be effective at influencing elements of narrative that may lead to behavior change, such as transportation and engagement [63,64]. As young women may be turning to social media when bored, including prevention efforts that entertain while also educating may help by first garnering attention, which can be difficult in the crowded social media environment. A recent meta-analysis of narrative game-based health behavior interventions found that such interventions had a large effect on improving knowledge, a medium effect on self-efficacy, and small effects on increasing enjoyment and encouraging health behaviors [65]. Narrative game-based interventions provided through or promoted on social media could be a useful option to explore for tanning prevention efforts.

## Limitations and Future Directions

As with any research, there are limitations to keep in mind with the current study. As we waited until the end of the ecological momentary assessment to elicit demographic information, demographic information was not available for a small proportion of participants. Additionally, while the sample came from 2 locations, it was a convenience sample that cannot be generalized more broadly. Additionally, we assessed only 6 emotions that participants may have felt, which were selected based on previous research related to tanning and social media use. Other emotions that were not captured in this study may have motivated participant behaviors. We also selected our time based on peak hours of UV exposure, so it is possible that there was social media use that occurred outside of our window of assessment, which could further impact findings associated with the role of emotions and social media use in particular. Perhaps emotions are stronger drivers of social media use only at certain times of day. We tried to mitigate the effects of a lack of data at certain time points in our creation of the composite variable, but we recognize that some instances of social media use were likely not captured, such as use during times that were outside our data collection efforts (eg, 11 PM) or on platforms that were not included (eg, Pinterest). However, the study does help advance the research in this area by using a different form of measurement through the use of ecological momentary assessment, allowing participants to report on media use, emotions, and behaviors closer in time to when they occurred, which should help reduce some of the potential difficulties of assessing dynamic constructs and issues with retrospective recall. However, although the analyses we conducted allowed us to assess the proposed research questions and hypotheses in a manner comparable to other research that has examined emotions, social media, and tanning in between-subject analyses [5,7,47], this analytical decision did not allow us to look at within-person differences. Future work could benefit from the addition of within-person analyses to further address the connections between emotions, social media use, and tanning.

Future research would also benefit from looking at younger audiences, as research suggests adolescents may also be impacted by social media in regard to indoor tanning behaviors [5]. Examining more specific social media channels and their associations with tanning behaviors, as well as the content presented in social media, would also be beneficial. Previous research has found different associations with tanning behaviors based on specific platforms [6]. A greater understanding of the specific content to which young people are exposed could help further explain effects beyond mere exposure, as previous research in other topic areas has found that perceptions of media messages may be influential among young people [66,67]. Future research also should continue to explore the relationship between social media and tanning behaviors among young women to further parse out the mechanisms through which effects occur and to continue to bolster prevention efforts. Additionally, more work in measurement, specifically for use in EMA studies, would be beneficial.

## Conclusion

Our week-long ecological momentary assessment study of young women found that boredom, more than other emotional predictors, was associated with increased use of social media and that increased use of social media was associated with time spent outdoors without sun protection. This means that social media such as Snapchat, Instagram, Twitter, and Facebook may be prime channels for intervention and present an opportunity for reaching young women who have an increased risk of engaging in risky sun behaviors. Additionally, interventions developed for these channels should aim to be entertaining and engaging, as boredom was associated with social media use in our sample. By creating interventions that can entertain and educate, attention may be increased, potentially leading to greater intervention effects.

## Conflicts of Interest

None declared.

## References

1. Key statistics for melanoma skin cancer. American Cancer Society. URL: https://www.cancer.org/cancer/melanoma-skin-cancer/about/key-statistics.html [accessed 2020-11-20]
2. Bleyer A, O'Leary M, Barr M, Ries L. Cancer epidemiology in older adolescents and young adults 15 to 29 years of age, including SEER incidence and survival: 1975-2000. National Cancer Institute Surveillance, Epidemiology, and End Results Program. 2006. URL: https://seer.cancer.gov/archive/publications/aya/ [accessed 2020-11-20]
3. Volkmer B, Greinert R. UV and children's skin. Prog Biophys Mol Biol 2011 Dec;107(3):386-388. [doi: 10.1016/j.pbiomolbio.2011.08.011] [Medline: 21907231]
4. Dennis LK, Lowe JB, Snetselaar LG. Tanning behavior among young frequent tanners is related to attitudes and not lack of knowledge about the dangers. Health Educ J 2009 Sep;68(3):232-243 [FREE Full text] [doi: 10.1177/0017896909345195] [Medline: 22707763]
5. Myrick JG, Noar SM, Kelley D, Zeitany AE. The relationships between female adolescents' media use, indoor tanning outcome expectations, and behavioral intentions. Health Educ Behav 2016 Sep 02;44(3):403-410. [doi: 10.1177/1090198116667251] [Medline: 27590838]
6. Willoughby JF, Myrick JG. Entertainment, social media use and young women's tanning behaviours. Health Educ J 2018 Dec 23;78(3):352-365. [doi: 10.1177/0017896918819643]
7. Niu Z, Parmar V, Xu B, Coups EJ, Stapleton JL. Prevalence and correlates of intentional outdoor and indoor tanning among adolescents in the United States: Findings from the FLASHE survey. Prev Med Rep 2018 Sep;11:187-190. [doi: 10.1016/j.pmedr.2018.06.014] [Medline: 29992085]
8. Lazarus RS. Progress on a cognitive-motivational-relational theory of emotion. Am Psychol 1991 Aug;46(8):819-834. [doi: 10.1037//0003-066x.46.8.819] [Medline: 1928936]
9. Roseman IJ. Emotional behaviors, emotivational goals, emotion strategies: multiple levels of organization integrate variable and consistent responses. Emotion Rev 2011 Sep 20;3(4):434-443. [doi: 10.1177/1754073911410744]
10. Stawski RS, Sliwinski MJ, Almeida DM, Smyth JM. Reported exposure and emotional reactivity to daily stressors: the roles of adult age and global perceived stress. Psychol Aging 2008;23(1):52-61. [doi: 10.1037/0882-7974.23.1.52] [Medline: 18361654]
11. Niederdeppe J. Meeting the challenge of measuring communication exposure in the digital age. Commun Methods Meas 2016 Apr 20;10(2-3):170-172. [doi: 10.1080/19312458.2016.1150970]
12. Eveland WP, Hutchens MJ, Shen F. Exposure, attention, or "use" of news? Assessing aspects of the reliability and validity of a central concept in political communication research. Commun Methods Meas 2009 Dec 09;3(4):223-244. [doi: 10.1080/19312450903378925]
13. Smyth JM, Stone AA. Ecological momentary assessment research in behavioral medicine. J Happiness Stud 2003;4(1):35-52. [doi: 10.1023/A:1023657221954]
14. Widén G, Lindström J, Brännback M, Huvila I, Nyström AG. Mixed emotions in active social media use - fun and convenient or shameful and embarrassing? 2015 Presented at: iConference 2015; March 24-27, 2015; Newport Beach, CA.

15. Escobar-Viera CG, Shensa A, Bowman ND, Sidani JE, Knight J, James AE, et al. Passive and active social media use and depressive symptoms among United States adults. Cyberpsychol Behav Soc Netw 2018 Jul;21(7):437-443. [doi: 10.1089/cyber.2017.0668] [Medline: 29995530]
16. Berryman C, Ferguson CJ, Negy C. Social media use and mental health among young adults. Psychiatr Q 2018 Jun;89(2):307-314. [doi: 10.1007/s11126-017-9535-6] [Medline: 29090428]
17. Correa T, Hinsley AW, de Zúñiga HG. Who interacts on the web?: the intersection of users' personality and social media use. Comput Hum Behav 2010 Mar;26(2):247-253. [doi: 10.1016/j.chb.2009.09.003]
18. Forgas JP. Affect in social thinking and behavior. New York, NY: Psychology Press; 2006.
19. Choi M, Toma CL. Social sharing through interpersonal media: patterns and effects on emotional well-being. Comput Hum Behav 2014 Jul;36:530-541. [doi: 10.1016/j.chb.2014.04.026]
20. Stieglitz S, Dang-Xuan L. Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. J Manage Inf Syst 2013 Jun 7;29(4):217-248. [doi: 10.2753/MIS0742-1222290408]
21. Williams DL, Crittenden VL, Keo T, McCarty P. The use of social media: an exploratory study of usage among digital natives. J Public Aff 2012 Mar 06;12(2):127-136. [doi: 10.1002/pa.1414]
22. Kim K, Yoo-Lee E, Joanna Sin S. Social media as information source: undergraduates' use and evaluation behavior. 2011 Presented at: ASIST 2011: Annual Meeting of The American Society for Information Science and Technology; Oct 7-12, 2011; New Orleans, LA. [doi: 10.1002/meet.2011.14504801283]
23. Bolton RN, Parasuraman A, Hoefnagels A, Migchels N, Kabadayi S, Gruber T, et al. Understanding Generation Y and their use of social media: a review and research agenda. J Serv Manage 2013 Jun 14;24(3):245-267. [doi: 10.1108/09564231311326987]
24. Valkenburg PM, Peter J, Schouten AP. Friend networking sites and their relationship to adolescents' well-being and social self-esteem. Cyberpsychol Behav 2006 Oct;9(5):584-590. [doi: 10.1089/cpb.2006.9.584] [Medline: 17034326]
25. Potente S, Coppa K, Williams A, Engels R. Legally brown: using ethnographic methods to understand sun protection attitudes and behaviours among young Australians 'I didn't mean to get burnt--it just happened!'. Health Educ Res 2010 Nov 08;26(1):39-52. [doi: 10.1093/her/cyq066] [Medline: 21059798]
26. Feldman SR, Liguori A, Kucenic M, Rapp SR, Fleischer AB, Lang W, et al. Ultraviolet exposure is a reinforcing stimulus in frequent indoor tanners. J Am Acad Dermatol 2004 Jul;51(1):45-51. [doi: 10.1016/j.jaad.2004.01.053] [Medline: 15243523]
27. Myrick JG, Noar SM, Kelley D, Zeitany AE, Morales-Pico BM, Thomas NE. A longitudinal test of the Comprehensive Indoor Tanning Expectations Scale: the importance of affective beliefs in predicting indoor tanning behavior. J Health Psychol 2016 Jul 10;22(1):3-15. [doi: 10.1177/1359105315595116] [Medline: 26231616]
28. Yoo J, Hur W. Body-tanning attitudes among female college students. Psychol Rep 2014 Apr;114(2):585-596. [doi: 10.2466/06.07.pr0.114k21w5] [Medline: 24897909]
29. Carcioppolo N, Orrego Dunleavy V, Myrick JG. A closer look at descriptive norms and indoor tanning: investigating the intermediary role of positive and negative outcome expectations. Health Commun 2018 Sep 10;34(13):1619-1627. [doi: 10.1080/10410236.2018.1517632] [Medline: 30198759]
30. Noar SM, Myrick JG, Morales-Pico B, Thomas NE. Development and validation of the Comprehensive Indoor Tanning Expectations Scale. JAMA Dermatol 2014 May 01;150(5):512-521. [doi: 10.1001/jamadermatol.2013.9086] [Medline: 24500373]
31. Fredrickson BL. The role of positive emotions in positive psychology. The broaden-and-build theory of positive emotions. Am Psychol 2001 Mar;56(3):218-226 [FREE Full text] [Medline: 11315248]
32. Kourosh AS, Harrington CR, Adinoff B. Tanning as a behavioral addiction. Am J Drug Alcohol Abuse 2010 Jun 15;36(5):284-290. [doi: 10.3109/00952990.2010.491883] [Medline: 20545604]
33. Zillmann D. Mood management through communication choices. Am Behav Sci 2016 Jul 27;31(3):327-340. [doi: 10.1177/000276488031003005]
34. Knobloch-Westerwick S. Mood management: theory, evidence, and advancements. In: Bryant J, Vorderer P, editors. Psychology of Entertainment. Mahwah, NJ: Lawrence Erlbaum Associates; 2006:239-254.
35. Dixon HG, Warne CD, Scully ML, Wakefield MA, Dobbinson SJ. Does the portrayal of tanning in Australian women's magazines relate to real women's tanning beliefs and behavior? Health Educ Behav 2011 Jan 24;38(2):132-142. [doi: 10.1177/1090198110369057] [Medline: 21263062]
36. Cho H, Lee S, Wilson K. Magazine exposure, tanned women stereotypes, and tanning attitudes. Body Image 2010 Sep;7(4):364-367. [doi: 10.1016/j.bodyim.2010.04.002] [Medline: 20573553]
37. Cafri G, Thompson JK, Roehrig M, Rojas A, Sperry S, Jacobsen PB, et al. Appearance motives to tan and not tan: evidence for validity and reliability of a new scale. Ann Behav Med 2008 Mar 21;35(2):209-220. [doi: 10.1007/s12160-008-9022-2] [Medline: 18365298]
38. Fogel J, Krausz F. Watching reality television beauty shows is associated with tanning lamp use and outdoor tanning among college students. J Am Acad Dermatol 2013 May;68(5):784-789. [doi: 10.1016/j.jaad.2012.09.055] [Medline: 23261546]
39. Social Media Use in 2018. Pew Research Center. URL: https://www.pewresearch.org/internet/2018/03/01/social-media-use-in-2018/ [accessed 2020-11-20]

40. Noar SM, Myrick JG, Zeitany A, Kelley D, Morales-Pico B, Thomas NE. Testing a social cognitive theory-based model of indoor tanning: implications for skin cancer prevention messages. Health Commun 2015;30(2):164-174. [doi: 10.1080/10410236.2014.974125] [Medline: 25470441]
41. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [Medline: 847061]
42. Bandura A. Social Foundations of Thought and Action. Upper Saddle River, NJ: Prentice-Hall; 1986.
43. Rodríguez VM, Daniel CL, Welles BF, Geller AC, Hay JL. Friendly tanning: young adults’ engagement with friends around indoor tanning. J Behav Med 2017 Feb 8;40(4):631-640. [doi: 10.1007/s10865-017-9832-4] [Medline: 28181003]
44. Moskowitz DS, Young SN. Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. J Psychiatry Neurosci 2006 Jan;31(1):13-20 [FREE Full text] [Medline: 16496031]
45. Sun safety. American Skin Association. URL: https://www.americanskin.org/resource/safety.php [accessed 2020-11-20]
46. Myrick JG, Holton AE, Himelboim I, Love B. #Stupidcancer: exploring a typology of social support and the role of emotional expression in a social media community. Health Commun 2015 Oct 09;31(5):596-605. [doi: 10.1080/10410236.2014.981664]
47. Myrick J, Willoughby J. Beyond demographics: The relationship of young women's emotional states with skin damaging and skin safety behaviors. 2017 Presented at: Annual Meeting of the National Communication Association; Nov 16-19, 2017; Dallas, TX.
48. Kraus A, Myrick JG. Feeling bad about feel-good ads: the emotional and body-image ramifications of body-positive media. Commun Res Rep 2017 Oct 17;35(2):101-111. [doi: 10.1080/08824096.2017.1383233]
49. Miyamoto Y, Ma X. Dampening or savoring positive emotions: a dialectical cultural script guides emotion regulation. Emotion 2011 Dec;11(6):1346-1357. [doi: 10.1037/a0025135] [Medline: 21910543]
50. Hanwell H, Vieth R, Cole D, Scillitani A, Modoni S, Frusciante V, et al. Sun exposure questionnaire predicts circulating 25-hydroxyvitamin D concentrations in Caucasian hospital workers in southern Italy. J Steroid Biochem Mol Biol 2010 Jul;121(1-2):334-337. [doi: 10.1016/j.jsbmb.2010.03.023] [Medline: 20298782]
51. Diffey B, Norridge Z. Reported sun exposure, attitudes to sun protection and perceptions of skin cancer risk: a survey of visitors to Cancer Research UK's SunSmart campaign website. Br J Dermatol 2009 Jun;160(6):1292-1298. [doi: 10.1111/j.1365-2133.2009.09149.x] [Medline: 19416235]
52. Harrell F. Regression modeling strategies: With application to linear models, logistic and ordinal regression, and survival analysis. New York, NY: Springer International Publishing; 2015.
53. Eastwood JD, Frischen A, Fenske MJ, Smilek D. The unengaged mind: defining boredom in terms of attention. Perspect Psychol Sci 2012 Sep;7(5):482-495. [doi: 10.1177/1745691612456044] [Medline: 26168505]
54. Hill A, Perkins R. Towards a model of boredom. Br J Psychol 1985 May;76(2):235-240. [doi: 10.1111/j.2044-8295.1985.tb01947.x] [Medline: 4027489]
55. Mastro DE, Eastin MS, Tamborini R. Internet search behaviors and mood alterations: a selective exposure approach. Media Psychol 2002 May;4(2):157-172. [doi: 10.1207/s1532785xmep0402_03]
56. Tandoc EC, Ferrucci P, Duffy M. Facebook use, envy, and depression among college students: is facebooking depressing? Comput Hum Behav 2015 Feb;43:139-146. [doi: 10.1016/j.chb.2014.10.053]
57. Appel H, Gerlach A, Crusius J. The interplay between Facebook use, social comparison, envy, and depression. Curr Opinion Psychol 2016 Jun;9:44-49 [FREE Full text] [doi: 10.1016/j.copsyc.2015.10.006]
58. Holman DM, Watson M. Correlates of intentional tanning among adolescents in the United States: a systematic review of the literature. J Adolesc Health 2013 May;52(5 Suppl):S52-S59 [FREE Full text] [doi: 10.1016/j.jadohealth.2012.09.021] [Medline: 23601612]
59. Falzone AE, Brindis CD, Chren M, Junn A, Pagoto S, Wehner M, et al. Teens, tweets, and tanning beds: rethinking the use of social media for skin cancer prevention. Am J Prev Med 2017 Sep;53(3S1):S86-S94 [FREE Full text] [doi: 10.1016/j.amepre.2017.04.027] [Medline: 28818251]
60. Singhal A, Rogers E. Entertainment-education: A communication strategy for social change. New York, NY: Routledge; 2012.
61. Kim K, Lee M, Macias W. An alcohol message beneath the surface of ER: how implicit memory influences viewers' health attitudes and intentions using entertainment-education. J Health Commun 2014;19(8):876-892. [doi: 10.1080/10810730.2013.837556] [Medline: 24479746]
62. van Leeuwen L, Renes RJ, Leeuwis C. Televised entertainment-education to prevent adolescent alcohol use: perceived realism, enjoyment, and impact. Health Educ Behav 2013 Apr;40(2):193-205. [doi: 10.1177/1090198112445906] [Medline: 22773596]
63. Willoughby JF, Liu S. Do pictures help tell the story? An experimental test of narrative and emojis in a health text message intervention. Comput Hum Behav 2018 Feb;79:75-82. [doi: 10.1016/j.chb.2017.10.031]
64. Willoughby JF, Niu Z, Liu S. Assessing the potential use of narrative and the entertainment education strategy in an mHealth text-message intervention. J Health Commun 2018;23(1):20-27. [doi: 10.1080/10810730.2017.1396628] [Medline: 29252131]
65. Zhou C, Occa A, Kim S, Morgan S. A meta-analysis of narrative game-based interventions for promoting healthy behaviors. J Health Commun 2020;25(1):54-65. [doi: 10.1080/10810730.2019.1701586] [Medline: 31829829]

66. Rodgers KB, Hust SJT, Willoughby JF, Wheeler J, Li J. Adolescents' sex-related alcohol expectancies and alcohol advertisements in magazines: the role of wishful identification, realism, and beliefs about women's enjoyment of sexualization. J Health Commun 2019;24(4):395-404. [doi: 10.1080/10810730.2019.1630523] [Medline: 31215356]
67. Willoughby JF, Hust SJT, Li J, Couto L. Social media, marijuana and sex: an exploratory study of adolescents' intentions to use and college students' use of marijuana. J Sex Res 2020 Oct 09:1-13. [doi: 10.1080/00224499.2020.1827217] [Medline: 33034543]

## Abbreviations

**EMA:** ecological momentary assessment
**FLASHE:** Family Life, Activity, Sun, Health, and Eating

---

*Edited by G Eysenbach; submitted 22.02.20; peer-reviewed by Anonymous, N Nguyen, J Makin; comments to author 01.05.20; revised version received 14.07.20; accepted 11.11.20; published 10.12.20.*

*Please cite as:*
*Willoughby JF, Myrick JG, Gibbons S, Kogan C*
*Associations Between Emotions, Social Media Use, and Sun Exposure Among Young Women: Ecological Momentary Assessment Study*
*JMIR Dermatol 2020;3(1):e18371*
*URL: http://derma.jmir.org/2020/1/e18371/*
*doi:10.2196/18371*
*PMID:*

Original Paper

# A Psychosocial Support Website From the British Association of Dermatologists for People Living With a Skin Condition: Mixed Methods Evaluation

Connor Heapy[1*], BSc, MSc, PhD; Kerry Montgomery[1*], BSc, MSc, PhD; Steven Ersser[2], BSc (Hons), RGN, PhD, PFHEA; Matt Gass[3], BA (Hons); Nina Goad[3], BA (Hons); Andrew R Thompson[4], BA (Hons), DClinPsy

[1]Department of Psychology, University of Sheffield, Sheffield, United Kingdom
[2]Department of Nursing, University of Bournemouth, Bournemouth, United Kingdom
[3]British Association of Dermatologists, London, United Kingdom
[4]Cardiff and Vale University Health Board / Cardiff University, South Wales Clinical Psychology Training, Cardiff, United Kingdom
[*]these authors contributed equally

**Corresponding Author:**
Andrew R Thompson, BA (Hons), DClinPsy
Cardiff and Vale University Health Board / Cardiff University
South Wales Clinical Psychology Training
11th Floor, Tower Building
70 Park Place
Cardiff, CF10 3AT
United Kingdom
Phone: 44 292 087 0582
Email: thompsona18@cardiff.ac.uk

## Abstract

**Background:** There is a lack of psychological support for individuals with skin conditions, and few low-intensity self-help interventions are available.

**Objective:** This study aimed to test the acceptability and usability of a support website and its embedded self-help resources.

**Methods:** A mixed methods approach was utilized. A total of 583 participants (426 with a skin condition, 97 relatives and friends, and 60 dermatology professionals) viewed the British Association of Dermatologist’s *SkinSupport* website and then completed a survey about their well-being and the usability of the website. A comparison group comprising 816 participants also completed the well-being measures. In total, 37 participants (19 living with a skin condition, and the relatives and friends of individuals with a skin condition, and 18 dermatology professionals) viewed the *SkinSupport* website and then took part in focus groups. Participants were recruited via social media, professional networks, and volunteer lists. Data from the survey were analyzed using descriptive and inferential statistics and qualitative content analysis.

**Results:** Both quantitative and qualitative responses suggest that the *SkinSupport* website was viewed positively by both patients and health professionals. Overall, 79.8% (417/523) of individuals with a skin condition, and the relatives and friends of individuals with a skin condition, said that they would use the website again; and 86.7% (52/60) of dermatology professionals said that they would recommend the site to somebody with a skin condition. Qualitative responses related to the website fell into 4 key themes: (1) appearance, (2) use and navigation, (3) information, and (4) areas for development.

**Conclusions:** The *SkinSupport* website was considered acceptable and usable. A range of areas requiring modification were identified. The website provides a useful resource that patients can access freely. Given the lack of services available to patients with skin conditions, health care professionals could routinely inform patients of this resource at assessment.

***(JMIR Dermatol 2020;3(1):e17052)*** doi:10.2196/17052

**KEYWORDS**

dermatology; psychology; surveys and questionnaires; interview; mobile phone

# Introduction

## Background

Skin conditions are common, affecting 54% of the UK population in any 12-month period [1]. Approximately 13 million people in England and Wales visit their general practitioner (GP) with a skin complaint each year, making it the most frequent reason for patients to visit their GP with a new problem [1,2]. Skin conditions may lead to difficulties in emotional, psychological, and social functioning [1,3]. Despite the potential negative impact that skin conditions can have on psychological well-being, access to psychological support for dermatology patients is acknowledged to be limited [1,4]. Although support websites and interventions exist for the group more broadly described as having a *visible difference*, many individuals with a skin condition will not identify themselves as having a visible difference and may, therefore, not find, or feel targeted by, such resources. In addition, there are many nonappearance-based issues associated with having a skin condition that do not apply to many other conditions that lead to a *visible difference*, such as pain and itch. As a result, the All Parliamentary Group on Skin has identified the provision of psychological services for people with skin conditions as a priority area for attention in 2019 [5].

Web-based self-help support has the potential to offer a cost-efficient method of addressing some of the gaps in the provision of psychosocial support for dermatology patients [5]. Such support can be considered as a *low-intensity* psychological intervention and is well-suited to individuals experiencing low-to-moderate levels of psychological distress [6]. Individuals could access such support in place of, or before, receiving more *intensive* forms of psychological therapy or support. Indeed, web-based self-help support has been found to improve individual well-being with a range of chronic health conditions [7]. However, there are few web-based self-help options available for people with skin conditions, and those that are available are condition-specific and not widely accessible (eg, cognitive-behavioral therapy–based web support for psoriasis [8]). Therefore, there is a need for a widely available self-help support tool for individuals with a range of skin conditions.

To address this need for psychological resources for individuals with a skin condition, the British Association of Dermatologists (BAD) launched a website, called *SkinSupport* [9], in July 2015. *Skin Support* was designed to support people living with a range of skin conditions and their relatives and friends, and as a signposting tool for professionals. The website was developed by a range of health care professionals and patient representatives. It includes access to downloadable self-help resources that have been subject to some evaluation. *SkinSupport* hosts a range of skin-specific psychosocial self-help information aimed at helping people to address common problems that are sometimes associated with skin conditions, such as how to regain confidence and improve low mood [10,11]. It also contains information and resources on different skin conditions, including treatment and causes.

## Objectives

This study aimed to examine the acceptability and usability of the BAD *SkinSupport* website for people living with skin conditions and their relatives and friends. In addition, the acceptability and usability of the website as a signposting resource for professionals working with people living with skin conditions was examined. Evaluating this web-based resource is vital in ensuring that it serves its intended purpose of providing low-intensity psychosocial support for dermatology patients.

# Methods

## Design

A mixed methods approach was used to collect both qualitative and quantitative data using a web-based survey and focus group discussions. Qualitative methods using focus groups were adopted to capture novel, unanticipated responses that may be missed by a predetermined questionnaire [12]. Quantitative questionnaire survey methods were adopted to help summarize key acceptability and usability information across a large sample of individuals, allowing the generalizability of the findings to be examined. Using a single method approach to this question, therefore, would mean that crucial information on the acceptability and usability of the website would be missed. This pragmatic, mixed methods approach, therefore, allowed for a comprehensive assessment of the acceptability and usability of the *SkinSupport* website [13,14].

## Participants

Both the questionnaire survey and the focus group were open to (1) individuals with a skin condition, (2) the relatives and friends of someone with a skin condition, and (3) dermatology professionals. To be involved in the study, participants were required to be able to speak fluent English and be aged ≥16 years. Any professional working with people with skin conditions was eligible to participate (eg, dermatologists, nurses, GPs, mental health workers and charities, etc).

Participants for the questionnaire survey (evaluation group) were recruited using a convenience sampling method. Participants for the focus groups were recruited via a purposive sampling method. For both parts of the study, participants living with a skin condition and their relatives and friends were recruited via a number of sources, including social media and volunteer lists (university and charity lists). Dermatology professionals were recruited via professional dermatology bodies, dermatology special interest groups, and staff working at a local National Health Service well-being service.

A comparison group of participants was also recruited for this study. These were individuals who had chosen to visit the *SkinSupport* website for purposes unrelated to the research project and then followed a link to a questionnaire on the website, inviting them to take part. This sample was recruited to check whether the participants who were involved in the website evaluation were typical of those who visited the *SkinSupport* website. Variables relevant to individuals with skin conditions were recorded and compared between the 2 samples.

## Procedure

For the survey, participants in the evaluation group followed a web-based link to the questionnaire on the web-based survey platform Qualtrics. The survey first instructed participants to follow a link to the *SkinSupport* website before answering related questions. Participants in the comparison group were individuals who had visited the *SkinSupport* website before the current evaluation started. These participants were also asked to complete a series of established, validated questionnaires (patient health questionnaire 2 [PHQ-2], generalized anxiety disorder questionnaire 2 [GAD-2], and dermatology quality of life index [DLQI]).

In total, 3 focus groups were conducted for dermatology professionals and 3 focus groups for people living with skin conditions, and the relatives and friends of individuals with a skin condition. Participants were first sent a link to the website and asked to look at it. In addition, the website was displayed during the focus group discussion and participants were asked to comment on particular aspects. Focus groups were run by 2 authors of this paper, separately: CH and KM. Some focus group discussions took place in university settings and others in hotel conference room settings. A semistructured interview schedule was used to guide the focus groups, and the discussions were recorded using an encrypted digital recorder. The focus groups lasted approximately 60 min. Data were transcribed and uploaded into NVivo version 11 (QSR International) for thematic analysis. The qualitative data analysis was conducted by KM with supervision provided by AT.

Ethical approval was obtained from the Department of Psychology Ethics Committee, University of Sheffield.

## Materials

The web-based survey assessed the acceptability and usability of the website as well as relevant demographic and clinical information. The acceptability and usability of the website were assessed by asking participants in the evaluation group for their opinions on a range of aspects of the website, including: (1) the appeal, (2) ease of use, (3) quality of information provided, (4) quality of self-help information, and (5) whether they would recommend the site.

The appeal of the website was assessed by asking participants to rate the home page and the rest of the website on a scale from *0* (not appealing) to *100* (very appealing). Participants were also asked to assess the balance of pictures and text on a 4-point Likert scale from *poor* to *excellent*. Ease of use was assessed by asking participants to rate on a 4-point Likert scale how easy they found the website to use, from *very easy to use* to *very difficult to use*. Quality of information provided was assessed by asking participants how helpful they found the information on the website from *0* (very unhelpful) to *100* (very helpful). Quality of self-help information provided was assessed by asking participants to rate on a 4-point Likert scale how easy the information was to understand, from *clear and easy to understand* to *very difficult to understand*. Participants were also asked to rate how helpful the self-help resource was from *0* (very unhelpful) to *100* (very helpful) and if they would use the techniques described in the self-help information from *0* (very unlikely) to *100* (very likely). Participants were also asked to rate on a 5-point Likert scale from *definitely* to *definitely not* whether they would use the website again, whether they would recommend the website to people living with a skin condition, and whether they would recommend the website to people supporting someone with a skin condition. Participants were also asked to provide any further comments on each of these areas in free-text response boxes.

The following demographic information was collected: age, gender, country of residence, ethnic group, and diagnosed skin condition (if relevant).

The DLQI is a 10-item measure of the impact of the skin condition on the patient's life in the last week. The DLQI has demonstrated good validity, reliability, and responsiveness to changes in a range of skin conditions [15,16].

PHQ-2 [17] is a 2-item measure of depression. PHQ-2 demonstrated good criterion and construct validity. PHQ-2 has a maximum score of 6, and a score of >3 is indicative of symptoms of major depression, with a sensitivity of 83% and specificity of 90% [17].

GAD-2 is a 2-item measure of the symptoms of GAD. GAD-2 has a maximum score of 6, and a score of >3 is indicative of clinically significant symptoms of GAD [18]. GAD-2 showed good sensitivity for GAD (88%), panic disorder (76%), and social anxiety disorder (70%) and good specificity (81%-83%) for all 3 disorders [17].

## Website Development

The website development was guided from the outset by an expert advisory panel, comprising patients, psychiatrists, psychologists, dermatologists, and communications professionals. This panel was able to help guide the BAD through the many complexities involved in providing advice to potentially vulnerable service users or those in acute distress, and, crucially, to collate, evaluate, and commission the profusion of materials required. A large proportion of time and resources were allocated to the development and testing of the website, before a national rollout. This included beta testing of the site with medical professionals, the public, and patients. Patient information materials on the website were reviewed by a medical committee and seconded by experts every 3 years. This review date can be brought forward in response to known changes, such as withdrawal of a named drug or updates to guidelines.

## Analysis

For the survey data, descriptive statistics were used to provide information on the demographics of participants in the evaluation group completing the survey. The current survey data for people living with skin conditions were compared with data collected by the BAD, before the launch of the survey (n=816), to examine any differences in anxiety, depression, and quality of life between website visitors and survey participants.

Qualitative free-text responses from the survey were analyzed using qualitative content analysis (QCA) [19]. QCA is a systematic text analysis technique that preserves the advantages of quantitative analysis, allowing frequencies of data to be reported while also providing a method to examine the experiences of participants using the *SkinSupport* website.

Qualitative data collected from the focus groups were analyzed using thematic analysis [20]. Thematic analysis is a method used to “identify, analyse and report patterns within data” to explain a particular phenomenon of interest [20]. Analysis of the focus group transcripts began by using line-by-line analysis, from which a list of key themes was generated from each group. These initial themes were then compared and contrasted to identify how the themes from each group fitted together. The aim was to preserve the integrity of the feedback gained from the individual groups within the final set of themes and to achieve internal saturation [21]. Detailed records of the analysis containing excerpts from the transcripts were maintained to enable checking.

Demographic data (age, gender, and ethnicity) of individuals in the evaluation group who would recommend the website to others were compared with those who would not recommend the website to others using a *t* test and chi-square test of homogeneity. Ethnicity was recoded into a dichotomous variable, to *white* and *nonwhite*, due to the relatively small number of participants with many of the *nonwhite* ethnic categories.

The qualitative results from the survey and focus groups were analyzed separately. These findings were then combined and summarized for the purpose of this study.

## Results

### Participant Characteristics

In total, 816 participants completed the web-based survey as part of the comparison group (no demographic information was collected) and 583 participants completed the web-based survey as part of the evaluation group (426 participants living with a skin condition, 97 relatives and friends of individuals living with skin conditions, and 60 dermatology professionals). The participants living with a skin condition, and the relatives and friends of individuals with a skin condition, were predominately white (394/463, 85.1%), female (407/477, 85.3%), living in the United Kingdom (465/473, 98.3%), and aged between 16 and 74 years (mean age 38.6, SD 12.6 years). The most common skin conditions reported were eczema (132/396, 33.3%), psoriasis (61/396, 15.4%), and acne (43/396, 10.9%), with some participants reporting more than one skin condition (Table 1). Health professionals included dermatologists (26/60, 43.3%), dermatology nurses (16/60, 26.7%), psychological practitioners (6/60, 10.0%), GPs and general practice nurses (3/60, 5.0%), and charity workers (2/60, 3.4%).

In addition, 37 participants took part in the focus groups (19 participants living with a skin condition, and the relatives and friends of people living with a skin condition, across the 3 focus groups, and 18 dermatology professionals across the 3 focus groups).

### Psychological Well-Being

Of those participants in the evaluation group, 33.9% (132/389) reported clinically significant symptoms of depression. In addition, 37.0% (144/389) of the participants in the evaluation group reported clinically significant symptoms of anxiety. Quality of life related to the skin condition was examined in those with a skin condition in the evaluation group (those without a skin condition were not assessed on quality of life), and 88.7% (345/389) of participants indicated that their skin condition had an effect on their quality of life, ranging from a small to an extremely large effect (Table 2).

A two-tailed independent *t* test found that participants in the comparison group scored significantly higher in depression (PHQ-2: mean 2.77, SD 1.91) than participants in the evaluation group (mean 2.23, SD 1.89; $t_{1203}$=−4.67; *P*<.001). In addition, participants in the comparison group scored significantly higher in anxiety (GAD-2: mean 2.75, SD 1.97) than those in the evaluation group (mean 2.38, SD 1.99; $t_{1203}$=−3.05; *P*=.002). Finally, participants in the comparison group reported poorer quality of life related to their skin condition (DLQI: mean 13.16, SD 7.84) than participants in the evaluation group (mean 9.53, SD 7.48; $t_{1207}$=−7.63; *P*<.001).

**Table 1.** Sample characteristics from individuals with a skin condition, and the relatives and friends of individuals with a skin condition, completing the web-based survey (N=523).

| Sample characteristic | Values, n (%) |
|---|---|
| **Skin condition**[a] | |
| Eczema, dermatitis, and prurigo | 132 (25.2) |
| Acne | 43 (8.2) |
| Psoriasis | 61 (11.6) |
| Undiagnosed skin problems (dry, oily, itchy, or flaky skin) | 31 (5.9) |
| Rosacea | 24 (4.5) |
| Alopecia | 13 (2.4) |
| Vitiligo | 12 (2.3) |
| Scarring | 12 (2.3) |
| Other | 68 (13.0) |
| Total that provided skin condition information | 396 (75.7) |
| **Gender** | |
| Male | 65 (12.4) |
| Female | 407 (77.8) |
| Other | 5 (1.0) |
| Total that provided gender information | 477 (91.2) |
| **Ethnicity** | |
| White | 394 (75.3) |
| Asian or Asian British | 39 (7.4) |
| Black or African or Caribbean or Black British Caribbean | 9 (1.7) |
| Other | 8 (1.5) |
| Preferred not to say | 13 (2.4) |
| Total that provided ethnicity information | 463 (88.5) |

[a] These data are from participants with a skin condition and not relatives and friends of individuals with a skin condition.

**Table 2.** Dermatology quality of life scores from participants with a skin condition who completed the web-based survey (N=389).

| Impact of skin condition on quality of life | Values, n (%) |
|---|---|
| No effect | 44 (11.3) |
| Small effect | 91 (23.4) |
| Moderate effect | 110 (28.3) |
| Very large effect | 104 (26.7) |
| Extremely large effect | 40 (10.3) |

The following information presented refers only to participants from the evaluation group, as only data on depression, anxiety, and quality of life were collected from the comparison group.

## Appearance of the Website

### *Descriptive Statistics*

Participants were asked to rate the appearance of the *SkinSupport* homepage and other pages on a scale of 0 (*not appealing*) to 100 (*very appealing*). Participants with a skin condition, and the relatives and friends of individuals with a skin condition, positively rated the appearance of the homepage (mean 74.92, SD 17.86) and the rest of the website (mean 71.49, SD 20.30). Similarly, professionals positively rated the appearance of the homepage (mean 77.32, SD 17.11) and the rest of the website (mean 76.48, SD 17.05).

The balance of pictures and text was rated on a 4-point Likert scale, ranging from poor to excellent. The majority of participants with a skin condition, and the relatives and friends of individuals with a skin condition, rated the balance of pictures and text as excellent (183/569, 32.2%) or good (309/569, 54.3%). Similarly, the majority of professionals rated the balance of text and images as excellent (18/58, 31%) or good (29/58, 50%).

### *Views of Individuals With Skin Conditions and the Relatives and Friends of Individuals With a Skin Condition*

The appearance, layout, and images used on the website were generally viewed positively by those with a skin condition, and the relatives and friends of someone with a skin condition (Multimedia Appendix 1). The website was described as looking professional and visually appealing. The images of patients being used on the website were viewed positively, although some commented that the images were not diverse enough and omitted younger age groups and ethnic minorities. In addition, some people suggested that more pictures of the different skin conditions should be added. There were conflicting views of the balance of text and image, with some saying the balance was good and others commenting that the website was too *text heavy*. Although some participants described the colors positively (eg, as *calming*), others described the colors as bland and the words as not being distinct enough, potentially affecting readability.

### *Views of Dermatology Professionals*

In general, professionals viewed the website positively, describing it as professional and visually appealing (Multimedia Appendix 1). As with the patient group, there were conflicting views regarding the images used on the website and the balance of images and text. Some participants commented that there was a lack of images in general, and more specifically, a lack of images representing ethnic minorities. Others commented that they liked the use of images of *real patients*. Participants suggested that adding videos to describe people's experiences could be helpful. One professional suggested that the medical information on the website was not accurate, and links to other websites needed updating. Finally, 1 dermatologist suggested that a separate area on the website for children and young people would be helpful, as the current website may not be appealing to a younger age group.

## Use and Navigation

### *Descriptive Statistics*

Participants were asked to rate the difficulty in understanding the information displayed on the website, and the language and terminology used on the website, on a scale from 0 (*very difficult to understand*) to 100 (*very easy to understand*). Those with a skin condition, and the relatives and friends of individuals with a skin condition, rated the information displayed on the website (mean 78.04, SD 20.15) and the language and terminology used (mean 77.49, SD 20.94) as easy to understand. Dermatology professionals also rated the information displayed on the website (mean 79.15, SD 20.23) and the language and terminology used (mean 77.90, SD 22.25) as easy to understand. In addition, 96.6% (504/522) of participants with a skin condition, and the relatives and friends of individuals with a skin condition, and 96.6% (56/58) of dermatology professionals, reported that the website was *easy*, or *very easy* to use.

### *Views of Individuals With Skin Conditions and the Relatives and Friends of Individuals With a Skin Condition*

Participants reported finding the website easy to use and navigate through, even on a smartphone (Multimedia Appendix 1). However, some participants said that it was difficult to locate materials on the website and that the structure could be improved, as it was difficult to navigate to previous pages. Similarly, 1 participant commented that the navigation bar was inconsistent throughout the website and should be made consistent. Some participants also suggested that a search function should be added to each page. In addition, it was suggested that the search function should be made more intelligent so that it could predict what condition you were looking for, even if there was imprecision in the typed search term. Finally, some participants commented on the difficulty of navigating through the *A-Z* list of conditions.

### *Views of Dermatology Professionals*

Professionals generally thought that the website was user friendly and easy to navigate (Multimedia Appendix 1). Although 1 participant felt that the use of medical jargon was avoided on the website, other participants reported that the patient information leaflets were written in unfriendly language, and some of the leaflets were too complex. It was suggested that a button that quickly navigates participants back to the home page would be useful. Finally, 1 participant was concerned about links on the website that took participants to information that did not inform or reassure sufficiently.

## Information on the Website

### *Descriptive Statistics*

Participants rated how helpful the information on the website was on a scale from 0 (*not helpful*) to 100 (*very helpful*). Participants with a skin condition had a mean score of 77.27 (SD 18.50), and the relatives and friends of someone with a skin condition had a mean score of 78.53 (SD 19.64), suggesting that they found the information helpful. In addition, 97.3% (468/481) of participants with a skin condition and the relatives and friends of someone with a skin condition and 100% (37/37) of dermatology professionals rated the self-help information on the website as clear and easy to understand, or somewhat easy to understand.

### *Views of Individuals With Skin Conditions and the Relatives and Friends of Individuals With a Skin Condition*

Participants generally reported that the website was comprehensive and helpful (Multimedia Appendix 1). However, some commented on the fact that some information was too long and wordy. Some participants noticed missing information, for example, on different subtypes of certain skin conditions. Others commented on the fact that they learned no new information from the website. It was also suggested that the purpose of the website should be clarified. Some commented that the information should be made less medical; in particular, the use of the term *psychodermatology* should be removed. Finally, some participants said that the information was too heavily focused on adults, with too much focus on psoriasis specifically.

### Views of Dermatology Professionals

The information on the website received mixed comments from dermatology professionals. Some thought the website needed more information on eczema (Multimedia Appendix 1). Others said that there was too much information on certain topics and that this may be off putting for some patients. One professional suggested that the use of videos to help with mindfulness exercises would be useful. In addition, some professionals suggested that the language of the website was overly medical in places and might benefit from revision. On certain topics, the information was said to be inaccurate. Furthermore, some professionals commented that the material was too focused on adults with a greater focus on psoriasis. Finally, 2 professionals reported being disappointed with the quality of information contained in some third-party materials on the website, including inaccurate crisis line information.

## Areas for Development

### Descriptive Statistics

Individuals with a skin condition, and the relatives and friends of individuals with a skin condition, were asked whether they would visit the website again; 42.3% (221/523) reported that they *definitely would* and 37.5% (196/523) reported that they *probably would*. In terms of recommending the website, 48.9% (256/523) said that they *definitely would* and 35.0% (183/523) said they *probably would* recommend the website to others living with a skin condition.

Dermatology professionals were asked whether they would recommend the website to people living with a skin condition; 75% (45/60) said that they *definitely would*, and 12% (7/60) said that they *probably would*. In terms of recommending the website to individuals supporting others with a skin condition, 65% (39/60) of dermatology professionals reported that they *definitely would*, and 18% (11/60) reported that they *probably would*.

Demographic data (ie, age, gender, and ethnicity) were compared between those individuals with a skin condition, and the relatives and friends of individuals with a skin condition, who would recommend (*definitely would* or *probably would*) the website to others and those who would not. There was no difference in age between those who would recommend the website to others (mean 38.7, SD 12.67) and those who would not recommend the website to others (mean 38.0, SD 12.45; $t_{473}$=4.56; *P*=.65). In addition, no differences were found in gender; 85.6% (338/395) women would recommend the website compared with 90% (69/77) women who would not recommend the website, a nonstatistically significant difference in proportions of 0.04 (*P*=.35). Similarly, no differences were found in ethnicity; 84.1% (332/395) white participants would recommend the website compared with 64.9% (61/94) white participants who would not recommend the website, a nonstatistically significant difference in proportions of 0.19 (*P*=.21).

### Views of Individuals With Skin Conditions and the Relatives and Friends of Individuals With a Skin Condition

Participants were asked if there were areas of *SkinSupport* that could be developed. A total of 2 key areas were highlighted: (1) case studies and stories of people living with skin conditions, and (2) addressing the impact of stigma on people living with skin conditions (Multimedia Appendix 1). Some in 2 of the focus groups felt that the stigma surrounding skin conditions needed to be acknowledged on *SkinSupport* and could be added as another support category. Participants also thought it beneficial to have information about how skin is portrayed in the media, the view of skin conditions as contagious, and what the BAD is doing to challenge the myths around skin conditions.

### Views of Dermatology Professionals

Professionals outlined several key areas that could be developed on the *SkinSupport* website. Regarding content, participants reported that information on habit reversal would be useful, and a question and answer section (Multimedia Appendix 1). One participant suggested that providing further information on common problems such as relationship issues and sleep would be beneficial, while also highlighting the connection between physical and mental health. Participants discussed having more condition-specific self-help available would be useful as a lot of the current information relates to psoriasis. Information on social stigma was also mentioned as being beneficial for patients. Finally, participants felt that the website needed further development to be applicable to young people.

# Discussion

## Principal Findings

This mixed methods study provides a detailed evaluation of the acceptability and usability of a psychosocial web-based support resource: *SkinSupport*. Participants included dermatology patients, and the relatives and friends of individuals with a skin condition, and professionals who work with individuals with skin conditions. Overall, *SkinSupport* was evaluated positively, and as such has the potential to be a useful resource for people living with skin conditions. For the quantitative measures of the website, the majority of participants gave positive ratings of the website's appearance, use, and navigation, and included information. In addition, participants provided positive feedback on the colors used on the website, the balance of pictures and text, and viewed the information provided as comprehensive for their condition. These quantitative results were generally supported by qualitative responses from participants in the questionnaire response boxes and in the focus groups.

Although the website was generally evaluated positively, several areas for improvement were identified. More specifically, recommendations relating to the appearance of the website include, adding more diverse images (ie, ethnic minorities, other age groups), adding more pictures of the conditions themselves, replacing text heavy sections with images, considering the use of videos, and adding further experiential accounts or stories.

Recommendations relating to the use and navigation of the website include adding a home button to each page to improve navigation, adding a search function to each page to improve navigation, improving the search function so that it is more general and intelligent (ie, specific words do not need to be typed in to find particular resources), and developing a separate area on the website for children and young people.

Recommendations relating to the information contained on the website include checking the accuracy of all the medical information on the website (including third-party links), reducing medical or unfriendly language (eg, the use of the word *psychodermatology*), making sure the information is more balanced and less focused on psoriasis, adding information that addresses the impact of stigma, relationship issues, and better signposting of the information on sleep, and adding a *question and answer* section to the website.

Some caution should be taken when generalizing the findings of the survey in this study to all users of the *SkinSupport* website. Participants who were involved in the evaluation of the website in this study demonstrated fewer symptoms of depression and anxiety and higher dermatology-related quality of life than participants in the comparison sample drawn from the existing visitors. This finding suggests that those who provided feedback on the survey may not be representative of those who typically visit the website. However, although there are some areas of the website where it could be plausible that individuals who are more distressed may evaluate the site differently (eg, self-help materials), for most areas (eg, appearance, use, and navigation), this seems unlikely. It is important to note that even those participants in the evaluation group of the website scored highly in depression (mean 2.23 SD 1.89 from a maximum of 6, the cutoff for major depression is 3; 132/389, 33.9% reached clinically significant levels of depression) and anxiety (mean 2.38 SD 1.99 from a maximum of 6, the cutoff for generalized anxiety is 3; 144/389, 37.0% reached clinically significant levels of anxiety) supporting the suggestion that such a self-help resource is needed for this population. Asking patients if they have visited this website may provide a useful rough guide for clinicians in identifying those individuals that require further psychological support.

It is important to note that 2 different constructs were being measured in this study: perceived usefulness of the website (ie, views from patients) and perceived usefulness of the website by others (ie, views from relatives, friends, and health care professionals). In general, the views of all groups aligned regarding each aspect of the website, meaning we can be confident in our conclusions on the strengths and weaknesses of the website in its current form. Indeed, many comments were repeated across groups (eg, images not diverse enough). Understandably, health care professionals were more likely to comment on the accuracy of the information and to suggest the addition of specific information (eg, on habit reversal). Overall, the website was perceived as useful (by patients) and perceived as useful by others (by relatives, friends, and health care professionals).

## Limitations

A methodological limitation of this study is that participants may not have reviewed the entire website before completing the survey or being involved in the focus groups. The website contains a lot of content, and it would take a significant amount of time to review it all. As a result, participants may have been commenting only on small sections of the website that they had reviewed. We cannot comment on the areas of the website that participants visited, or the time they spent on the website, as these data were not collected. Although all visitors to the website were invited to complete questionnaires (to form the comparison group), it is not possible to comment on the overall response rate of those who visited the website as we did not collect this data. In addition, demographic data were not collected from this comparison sample to reduce the burden and increase the response rate. As a result, the comparison group may not be representative of those who visit the website. Despite these limitations, the study involved a large sample size and captured a wide range of views from patients, relatives, and friends of individuals with a skin condition, and health care professionals. The mixed methods approach meant that comprehensive responses were collected using different types of data regarding many areas of the *SkinSupport* website.

A further limitation of the study is that some of the authors of the paper (MG, AT, and NG) were also on the advisory panel for development of the website, increasing the risk of bias. However, neither the quantitative nor qualitative data were analyzed by these authors to reduce the likelihood of bias and increase the rigor of the analysis.

Future studies are needed to investigate the effectiveness of some of the untested self-help materials contained on the website, and the tested interventions [6,7] might be moved into webpage format to encourage higher usage. In addition, a brief evaluation should be carried out after the suggested changes have been implemented on the website.

## Conclusions

Overall, the *SkinSupport* website is an acceptable and useful resource for people living with skin conditions, and the relatives and friends of individuals with a skin condition, and dermatology professionals. Some of the suggested changes around inaccurate information have already been implemented by BAD. Future decisions on which suggested changes to make will be made by BAD, with the help of this study, once further funding has been secured. After suggested improvements have been made, the website will serve as a valuable resource that can fit into the stepped care model of psychosocial care in dermatology, offering skin-specific self-help to people experiencing distress.

## Acknowledgments

This project was created by the BAD with funding received from the UK Department of Health Voluntary Sector Investment Programme Innovation, Excellence and Strategic Development Fund (award number 251784).

## Conflicts of Interest

The funders of this project were not involved in the review or approval of the manuscript. Authors NG and MG are employees of the BAD and authors AT and SE were unpaid members of the professional advisory panel for the initial construction of the website. Data were independently collected and analyzed by authors KM and CH under supervision of authors AT and SE.

Multimedia Appendix 1
Table of qualitative responses (survey and focus group) for participants in the evaluation group.
[DOCX File , 25 KB - derma_v3i1e17052_app1.docx ]

## References

1. The All Party Parliamentary Group on Skin. 2013. The Psychological and Social Impact of Skin Diseases on People's Lives URL: https://www.appgs.co.uk/publication/view/the-psychological-and-social-impact-of-skin-diseases-on-peoples-lives-final-report-2013/ [accessed 2019-02-12]
2. Schofield J, Fleming D, Grindlay D, Williams H. Skin conditions are the commonest new reason people present to general practitioners in England and Wales. Br J Dermatol 2011 Nov;165(5):1044-1050. [doi: 10.1111/j.1365-2133.2011.10464.x] [Medline: 21692764]
3. Grandfield T, Thompson A, Turpin G. An attitudinal study of responses to a range of dermatological conditions using the implicit association test. J Health Psychol 2005 Nov;10(6):821-829. [doi: 10.1177/1359105305057316] [Medline: 16176959]
4. Lavda A, Webb T, Thompson A. A meta-analysis of the effectiveness of psychological interventions for adults with skin conditions. Br J Dermatol 2012 Nov;167(5):970-979. [doi: 10.1111/j.1365-2133.2012.11183.x] [Medline: 22924999]
5. The All Party Parliamentary Group on Skin. 2019. Purpose and Aims URL: http://www.appgs.co.uk/wp-content/uploads/2019/02/APPGS_2019_Work_Priorities.pdf [accessed 2019-06-10]
6. Thompson A. Self-help for management of psychological distress associated with skin conditions. In: Bewley BA, Taylor RE, Reichenberg JS, Magid M, editors. Practical Psychodermatology. West Sussex: Wiley-Blackwell; 2014:60-65.
7. Beatty L, Lambert S. A systematic review of internet-based self-help therapeutic interventions to improve distress and disease-control among adults with chronic health conditions. Clin Psychol Rev 2013 Jun;33(4):609-622. [doi: 10.1016/j.cpr.2013.03.004] [Medline: 23603521]
8. Bundy C, Pinder B, Bucci S, Reeves D, Griffiths C, Tarrier N. A novel, web-based, psychological intervention for people with psoriasis: the electronic targeted intervention for psoriasis (eTIPs) study. Br J Dermatol 2013 Aug;169(2):329-336. [doi: 10.1111/bjd.12350] [Medline: 23551271]
9. British Association of Dermatologists. URL: http://www.skinsupport.org.uk/ [accessed 2020-04-26]
10. Pasterfield M, Clarke S, Thompson AR. The development of a self-help intervention to build social confidence in people living with visible skin conditions or scars: a think-aloud study. Scars Burn Heal 2019;5:1-8 [FREE Full text] [doi: 10.1177/2059513118822954] [Medline: 30675396]
11. Shah R, Hunt J, Webb TL, Thompson AR. Starting to develop self-help for social anxiety associated with vitiligo: using clinical significance to measure the potential effectiveness of enhanced psychological self-help. Br J Dermatol 2014 Aug;171(2):332-337. [doi: 10.1111/bjd.12990] [Medline: 24665869]
12. Bradbury K, Dennison L, Little P, Yardley L. Using mixed methods to develop and evaluate an online weight management intervention. Br J Health Psychol 2015 Feb;20(1):45-55. [doi: 10.1111/bjhp.12125] [Medline: 25406436]
13. Yardley L, Bishop F. Using mixed methods in health research: benefits and challenges. Br J Health Psychol 2015 Feb;20(1):1-4. [doi: 10.1111/bjhp.12126] [Medline: 25582981]
14. Yardley L, Bishop F. Mixing qualitative and quantitative methods. In: Willig C, Stainton-Rogers W, editors. The Sage Handbook Of Qualitative Research In Psychology. Thousand Oaks, CA: Sage Publications; 2007:352-369.
15. Finlay A, Khan G. Dermatology life quality index (DLQI)-a simple practical measure for routine clinical use. Clin Exp Dermatol 1994 May;19(3):210-216. [doi: 10.1111/j.1365-2230.1994.tb01167.x] [Medline: 8033378]
16. Basra M, Fenech R, Gatt R, Salek M, Finlay A. The dermatology life quality index 1994-2007: a comprehensive review of validation data and clinical results. Br J Dermatol 2008 Nov;159(5):997-1035. [doi: 10.1111/j.1365-2133.2008.08832.x] [Medline: 18795920]
17. Kroenke K, Spitzer R, Williams J. The patient health questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
18. Kroenke K, Spitzer R, Williams J, Monahan P, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 6;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
19. Mayring P. Qualitative content analysis. In: Steinke I, Flick U, Kardoff EV, editors. A Companion to Qualitative Research. Thousand Oaks, CA: Sage Publications; 2004:266-269.
20. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
21. Glaser B, Strauss A. The Discovery of Grounded Theory. Chicago: Aldine Pub Co; 1967.

## Abbreviations

| | |
|---|---|
| **BAD:** | British Association of Dermatologists |
| **DLQI:** | dermatology quality of life index |
| **GAD:** | generalized anxiety disorder |
| **GP:** | general practitioner |
| **PHQ:** | patient health questionnaire |
| **QCA:** | qualitative content analysis |

---

*Edited by G Eysenbach; submitted 13.11.19; peer-reviewed by B Eapen, D Harcourt; comments to author 10.01.20; revised version received 03.04.20; accepted 08.04.20; published 15.06.20.*

*Please cite as:*
*Heapy C, Montgomery K, Ersser S, Gass M, Goad N, Thompson AR*
*A Psychosocial Support Website From the British Association of Dermatologists for People Living With a Skin Condition: Mixed Methods Evaluation*
*JMIR Dermatol 2020;3(1):e17052*
*URL: http://derma.jmir.org/2020/1/e17052/*
*doi:10.2196/17052*
*PMID:*

Original Paper

# Comparison of Traditional Citation Metrics and Altmetrics Among Dermatology Journals: Content and Correlational Analysis Study

Gregg Murray[1], BSc, MSc, MD; Rebecca Hellen[1], MD; James Ralph[1], BSc, MD; Siona Ni Raghallaigh[1], MD, PhD

Beaumont Hospital, Dublin, Ireland

**Corresponding Author:**
Gregg Murray, BSc, MSc, MD
Beaumont Hospital
Beaumont Road
Dublin,
Ireland
Phone: 353 813000
Email: murrgr@gmail.com

## Abstract

**Background:** Research impact has traditionally been measured using citation count and impact factor (IF). Academics have long relied heavily on this form of metric system to measure a publication's impact. A higher number of citations is viewed as an indicator of the importance of the research and a marker for the impact of the publishing journal. Recently, social media and online news sources have become important avenues for dissemination of research, resulting in the emergence of an alternative metric system known as altmetrics.

**Objective:** We assessed the correlation between altmetric attention score (AAS) and traditional scientific impact markers, namely journal IF and article citation count, for all the dermatology journal and published articles of 2017.

**Methods:** We identified dermatology journals and their associated IFs available in 2017 using InCites Journal Citation Reports. We entered all 64 official dermatology journals into Altmetric Explorer, a Web-based platform that enables users to browse and report on all attention data for every piece of scholarly content for which Altmetric Explorer has found attention.

**Results:** For the 64 dermatology journals, there was a moderate positive correlation between journal IF and journal AAS ($r_s$=.513, $P$<.001). In 2017, 6323 articles were published in the 64 dermatology journals. Our data show that there was a weak positive correlation between the traditional article citation count and AAS ($r_s$=.257, $P$<.001).

**Conclusions:** Our data show a weak correlation between article citation count and AAS. Temporal factors may explain this weak association. Newer articles may receive increased online attention after publication, while it may take longer for scientific citation counts to accumulate. Stories that are at times deemed newsworthy and then disseminated across the media and social media platforms border on sensationalism and may not be truly academic in nature. The opposite can also be true.

***(JMIR Dermatol 2020;3(1):e15643)*** doi:10.2196/15643

**KEYWORDS**

dermatology; altmetrics; impact factor; citations; medical informatics

## Introduction

Research impact has traditionally been measured using citation count and impact factor (IF). Academics have long relied on this form of metric system to measure a publication's impact [1]. A higher number of citations is viewed as an indicator of the importance of the research and a marker for the impact of the publishing journal [2,3]. More recently, social media and online news sources have become important avenues for dissemination of research, resulting in the emergence of an alternative metric system known as altmetrics [4,5]. Altmetrics is data that can explain both the nature and volume of attention that research receives. It measures how many people have engaged with and shared research and allows the researchers or publishers to see in what manner it was shared.

Numerous altmetrics harvesting tools have recently been developed, including the Altmetric Explorer. Altmetric Explorer is a Web-based platform that enables users to browse and report on all attention data for every piece of scholarly content for which the platform has found attention [5]. The system tracks the online attention research receives by aggregating data from numerous avenues of online sharing sources, such as public policy documents, Mendeley, PubMed, mainstream media and social media outputs like Facebook and Twitter, and many more [5]. An altmetric attention score (AAS), derived by an automated algorithm, indicates the amount of attention and, in some cases, public engagement that research has received [5]. Therefore, articles that generate more attention are likely to have higher scores [6]. The AAS is based on three main principles: Volume, sources, and authors [5]. In addition, each form of mention contributes a different amount to the final score; for example, if an article is mentioned and shared via an online newspaper, it receives a higher score than if it is mentioned in a single tweet [5].

The aim of this study was to investigate the relationship between altmetrics and traditional citation metrics among dermatology journals and articles. Does sharing research online via social media and other news sources correlate with article citation counts and journal IFs? We assessed the correlation between AAS and traditional scientific impact markers, namely journal IF and article citation count, for all the dermatology journals and published articles of 2017.

## Methods

We identified dermatology journals available in 2017 and their associated IFs using InCites Journal Citation Reports [7]. We entered all 64 dermatology journals into the Altmetric Explorer search platform, and the search year was limited to 2017. This generated the AAS for each journal and its published articles. Using Dimensions, an online research database that provides the current citation count for published research, we gathered the citation counts for each published article [8]. We applied the Spearman rank correlation coefficient to assess the correlation between dermatology journal IF and AAS as well as dermatology article citation count and AAS. Statistical analysis was performed using SPSS V21 (IBM Corporation, Somers).

## Results

From the Journal Citation Report 2017, there were 64 dermatology journals with IFs ranging from 8.1 to 0.08. Of these journals, there was a moderate positive correlation between journal IF and journal AAS ($r_s$=.513, *P*<.001; Figure 1). From these journals, 6323 articles were published in 2017. Our data show a weak positive correlation between traditional citation count and AAS ($r_s$=.257, *P*<.001; Figure 2).

**Figure 1.** Spearman correlation between journal impact factor and altmetric attention score (AAS) in dermatology journals in 2017 (rs = .513, *P*&amp;lt;.001).

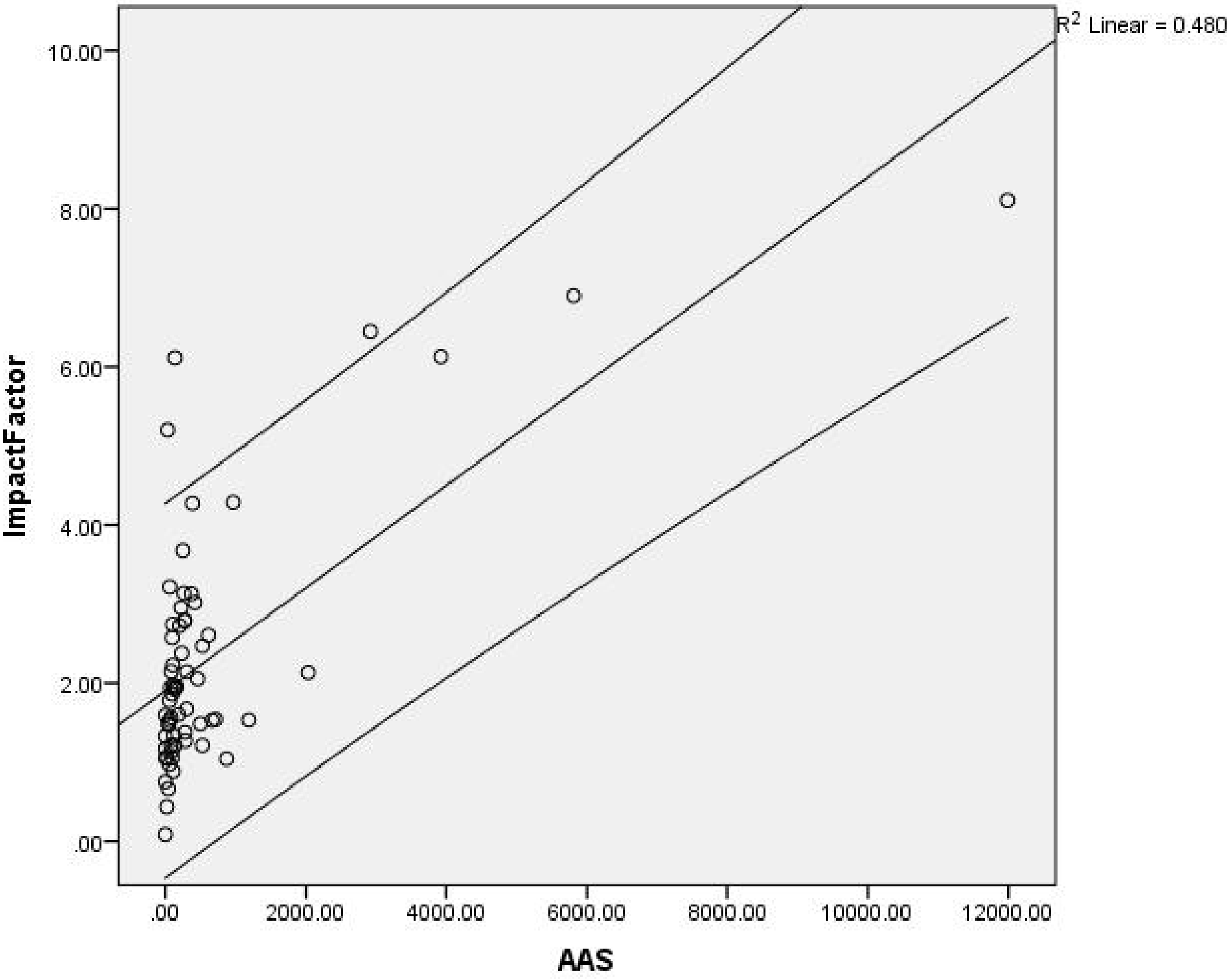
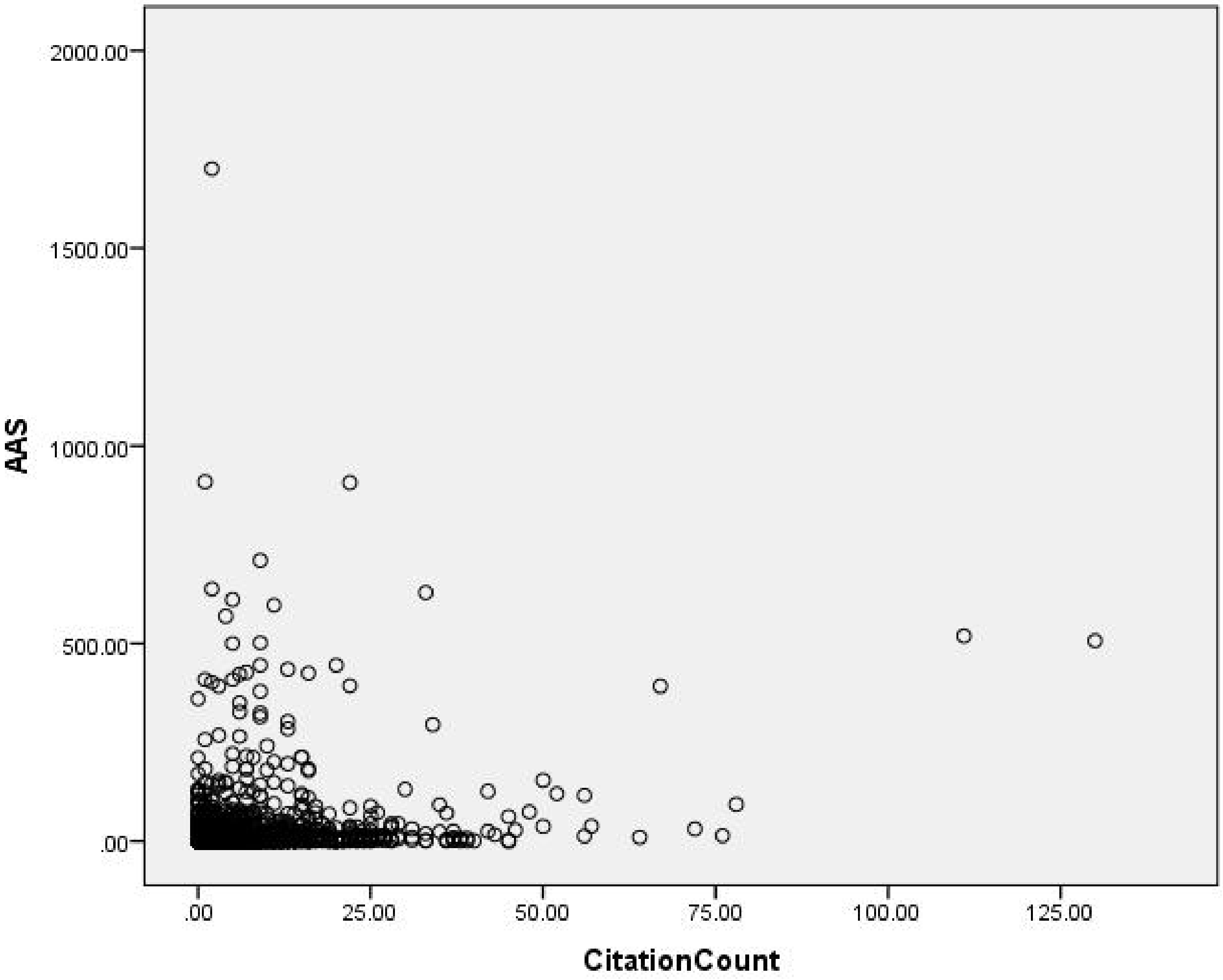


**Figure 2.** Spearman correlation between article citation count and altmetric attention score (AAS) in dermatology articles published in 2017 (rs = .257, *P*&amp;lt;.001).



## Discussion

Altmetrics complements, rather than replaces, the traditional citation metric system. Researchers and publishers can see not only how their research is distributed online but also in what form it is distributed, either through social media platforms such as Twitter and Facebook or via news articles. Our data demonstrate a moderate correlation between journal IF and AAS among dermatology journals, suggesting that the higher IF dermatology journals have a larger impact across social media and mainstream media. Our data also show a weak correlation between article citation count and AAS. Temporal factors may explain this weak association: Newer articles may receive increased online attention after publication, while it may take longer for scientific citation counts to accumulate. The weak correlation between the number of article citations and AAS in our study is consistent with recently published work of weak correlations between AAS and both cardiology and pediatric surgery citation counts [4,9]. These findings suggest that articles achieving high AAS, and therefore online attention, are not of equivalent interest to academics.

It is also important to note the occasional disconnect between what are deemed newsworthy publications and comprehensive academic research. Altmetrics reflects online crowd attention but does not reflect the quality, validity, and originality of research. The obverse could be said about traditional citation metrics, which focus on quality and validity but not dissemination of new research. For these metrics, the journal reach is limited by subscription fees or firewalls.

The near instantaneous nature of the altmetric score, with the ability for an article to go viral and a corresponding rapid boost in the AAS, enables the gaming of altmetrics. With enough time, coupled with the speed of the internet and freedom to post, research AAS could be artificially inflated by repeated tweets and widespread dissemination of research online. However, altmetrics addresses this issue by weighing author contribution with each mention to the overall attention score.

A limitation of this study was the exclusion of non-dermatological journals such as *Nature*, *New England Journal of Medicine*, and the *Lancet*, which also publish highly cited dermatological articles.

To our knowledge, this is the first study to assess the relationship between dermatology journal citation metrics and AAS.

## Conflicts of Interest

None declared.

## References

1. Sud P, Thelwall M. Evaluating altmetrics. Scientometrics 2014;98(2):1131-1143. [doi: 10.1007/s11192-013-1117-2]
2. Selvaraj S, Borkar DS, Prasad V. Media Coverage of Medical Journals: Do the Best Articles Make the News? PLoS ONE 2014 Jan 17;9(1):e85355. [doi: 10.1371/journal.pone.0085355]
3. O'Connor EM, Nason GJ, O'Kelly F, Manecksha RP, Loeb S. Newsworthiness vs scientific impact: are the most highly cited urology papers the most widely disseminated in the media? BJU Int 2017 May 18;120(3):441-454. [doi: 10.1111/bju.13881]
4. Patel RB, Vaduganathan M, Bhatt DL, Bonow RO. Characterizing High-Performing Articles by Altmetric Score in Major Cardiovascular Journals. JAMA Cardiol 2018 Dec 01;3(12):1249-1251. [doi: 10.1001/jamacardio.2018.3823]
5. Altmetric. 2018. URL: https://www.altmetric.com/about-our-data/how-it-works/ [accessed 2019-10-01]
6. Heidenreich A, Bastian PJ, Bellmunt J, Bolla M, Joniau S, van der Kwast T, et al. EAU Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent—Update 2013. Eur Urol 2014 Jan;65(1):124-137. [doi: 10.1016/j.eururo.2013.09.046]
7. InCites Journal Citation Report 2017. URL: https://incites.clarivate.com [accessed 2019-10-01]
8. Dimensions. URL: https://app.dimensions.ai/discover [accessed 2019-07-01]
9. Chang J, Desai N, Gosain A. Correlation Between Altmetric Score and Citations in Pediatric Surgery Core Journals. J Surg Res 2019 Nov;243:52-58. [doi: 10.1016/j.jss.2019.05.010] [Medline: 31154133]

## Abbreviations

**AAS:** altmetric attention score
**IF:** impact factor

---

*Edited by G Eysenbach; submitted 25.07.19; peer-reviewed by F Kaliyadan, C Liebram; comments to author 14.09.19; revised version received 11.10.19; accepted 17.12.19; published 12.02.20.*

*Please cite as:*
*Murray G, Hellen R, Ralph J, Ni Raghallaigh S*
*Comparison of Traditional Citation Metrics and Altmetrics Among Dermatology Journals: Content and Correlational Analysis Study*
*JMIR Dermatol 2020;3(1):e15643*
*URL: https://derma.jmir.org/2020/1/e15643*
*doi:10.2196/15643*
*PMID:*

Original Paper

# Evaluation of Spin in the Abstracts of Systematic Reviews and Meta-Analyses Focused on the Treatment of Acne Vulgaris: Cross-Sectional Analysis

Ryan Ottwell[1], BS; Taylor C Rogers[1], BS; J Michael Anderson[1], BS; Austin Johnson[1], BS; Matt Vassar[1], PhD

Oklahoma State University Center for Health Sciences, Tulsa, OK, United States

**Corresponding Author:**
Ryan Ottwell, BS
Oklahoma State University Center for Health Sciences
1111 W 17th St
Tulsa, OK, 74107
United States
Phone: 1 9189918718
Email: ryan.ottwell@okstate.edu

## Abstract

**Background:** Spin is the misrepresentation of study findings, which may positively or negatively influence the reader's interpretation of the results. Little is known regarding the prevalence of spin in abstracts of systematic reviews, specifically systematic reviews pertaining to the management and treatment of acne vulgaris.

**Objective:** The primary objective of this study was to characterize and determine the frequency of the most severe forms of spin in systematic review abstracts and to evaluate whether various study characteristics were associated with spin.

**Methods:** Using a cross-sectional study design, we searched PubMed and EMBASE for systematic reviews focusing on the management and treatment of acne vulgaris. Our search returned 316 studies, of which 36 were included in our final sample. To be included, each systematic review must have addressed either pharmacologic or nonpharmacologic treatment of acne vulgaris. These studies were screened, and data were extracted in duplicate by two blinded investigators. We analyzed systematic review abstracts for the nine most severe types of spin.

**Results:** Spin was present in 31% (11/36) of abstracts. A total of 12 examples of spin were identified in the 11 abstracts containing spin, with one abstract containing two instances of spin. The most common type of spin, *selective reporting of or overemphasis on efficacy outcomes or analysis favoring the beneficial effect of the experimental intervention*, was identified five times (5/12, 42%). A total of 44% (16/36) of studies did not report a risk of bias assessment. Of the 11 abstracts containing spin, six abstracts (55%) had not reported a risk of bias assessment or performed a risk of bias assessment but did not discuss it. Spin in abstracts was not significantly associated with a specific intervention type, funding source, or journal impact factor.

**Conclusions:** Spin is present in the abstracts of systematic reviews and meta-analyses covering the treatment of acne vulgaris. This paper raises awareness of spin in abstracts and emphasizes the importance of its recognition, which may lead to fewer incidences of spin in future studies.

***(JMIR Dermatol 2020;3(1):e16978)*** doi:10.2196/16978

**KEYWORDS**

acne vulgaris; systematic review; abstracts; dermatology

## Introduction

### Background

Acne vulgaris is one of the most common medical diagnoses made by dermatologists each year, with over 50 million Americans affected by this condition [1]. In 2013 alone, more than 5 million people with acne vulgaris sought medical treatment, resulting in over US $1.2 billion spent annually on the disease [1]. This condition can cause physical and emotional discomfort for those affected and can lead to long-term scarring [2]. Owing to the prevalence of acne vulgaris, it is important for the providers to be well informed about the current treatment options that are supported with the highest level of evidence.

Systematic reviews offer valuable insights to clinicians about current treatments [3]. Clinicians will often look specifically at the abstract of systematic reviews—which may be all that is accessible to them—as a succinct summary of results to guide treatment decisions. Barry et al [4] found that abstracts influenced the treatment decisions of nearly three-quarters of family physicians. Similarly, Marcelo et al [5] reported that nearly two-thirds of the residents used the abstracts of papers alone to guide them in their clinical decision making. Johnson et al [6] reported that the majority of nurse practitioner students found abstracts useful in guiding their clinical decision making. Haynes et al [7], investigating the usefulness of Medical Literature Analysis and Retrieval System Online (MEDLINE) in the clinical setting, showed that almost half of the searches performed influenced clinical decision making. They also found that, more often than not, decisions were made from limited information such as the abstract, article title, subject heading, or a combination of these. Therefore, it is imperative that the information provided in the abstracts of publications be representative of the study findings and devoid of spin. Boutron et al [8] defined spin as “a specific way of reporting, intentional or not, to highlight that the beneficial effect of the experimental treatment in terms of efficacy or safety is greater than that shown by the results.” Spin, specifically in abstracts, can influence the interpretation of data, which can lead to misinformation and, ultimately, misguided decisions about treatment.

### Objectives

Spin has been demonstrated to be prevalent in the abstracts of randomized controlled trials [9-14]. Similar lines of inquiry have noted poor reporting quality in systematic reviews [15-18]. On the basis of these two complementary lines of work, we hypothesized that spin would also be present in the abstracts of systematic reviews on the topic of treatment of acne vulgaris. This study aimed to characterize and determine the frequency of each type of spin in systematic review abstracts. More specifically, this study evaluated the top nine most severe types of spin, as previously outlined by Yavchitz et al [8]. The secondary objective was to evaluate whether various study characteristics corresponded with the presence of spin in systematic review abstracts focusing on acne vulgaris.

## *Methods*

### Publication Search

This study was performed in accordance with a previously written protocol available publicly on Open Science Framework (OSF) [19]. A search of PubMed and EMBASE, which incorporates MEDLINE, was performed on September 2, 2019. The search string for PubMed was as follows: “acne vulgaris”[MeSH Terms] OR acne vulgaris[Text Word] AND (systematic[sb] OR Meta-Analysis[ptyp]). The search string for EMBASE was as follows: ‘acne vulgaris’/exp AND (‘meta analysis’/de OR ‘systematic review’/de). These searches were modeled from the search strategies provided by a number of Cochrane systematic reviews on acne vulgaris [20,21]. Search results were then added to Rayyan [22], a screening platform.

### Training

Before screening, training was conducted, which included face-to-face training sessions for the authors RO and CR. Author MV led the training sessions as his studies regarding the presence of spin in randomized controlled trials have previously been published. During these training sessions, the authors (RO and CR) received instruction and education using example abstracts and full studies from different areas of medicine. During this training, the most severe types of spin described by Yavchitz et al [8] were discussed, and agreement was reached on the definition of each item. A Google form containing the nine most severe types of spin was developed and pilot tested for ease of use and to ensure that all the necessary data elements were included. The Google form was pilot tested by the authors (RO and CR) on numerous papers known to contain spin to ensure the form contained all the items needed for data extraction and that the Google form worked correctly. Studies were then screened independently by title and abstract by RO and CR. To be included in this study, a published article had to be a systematic review or meta-analysis designed to address either the pharmacological or nonpharmacological treatment of acne vulgaris.

### Data Extraction

Data extraction was performed in duplicate fashion using the Google form, and the investigators were blinded during the extraction process. After data extraction was completed, the investigators were unblinded and met to discuss any inconsistency until agreement was reached. On items for which agreement could not be reached, a third investigator was available for adjudication. The process of dual extraction of data was used to substantiate the work of each individual investigator. The items extracted on the Google form included the study title, journal title, number of studies included, number of subjects included, types of studies included, types of interventions, when the review was performed, databases searched, funding source, and if the review/primary outcome was significant. Furthermore, we evaluated whether the article contained any of the nine most severe examples of spin in abstracts [8].

To characterize and determine the frequency of spin in the abstracts of systematic reviews on the topic of the treatment of acne vulgaris, we employed the classification system previously developed by Yavchitz et al [8], specifically evaluating the top nine most severe examples of spin found in the abstracts of systematic reviews. These nine types of spin are outlined in Table 1.

**Table 1.** Frequency of each type of spin (N=12).

| Type of spin | Value, n (%)[a] |
|---|---|
| Conclusion contains recommendations for clinical practice not supported by the findings. | 0 (0) |
| Title claims or suggests a beneficial effect of the experimental intervention not supported by the findings. | 0 (0) |
| Selective reporting of or overemphasis on efficacy outcomes or analysis favoring the beneficial effect of the experimental intervention. | 5 (42) |
| Conclusion claims safety based on nonstatistically significant results with a wide confidence interval. | 0 (0) |
| Conclusion claims the beneficial effect of the experimental treatment despite a high risk of bias in primary studies. | 4 (33) |
| Selective reporting of or overemphasis on harm outcomes or analysis favoring the safety of the experimental intervention. | 1 (83) |
| Conclusion extrapolates the review's findings to a different intervention (ie, claiming efficacy of a specific intervention although the review covers a class of several interventions). | 2 (17) |
| Conclusion extrapolates the review's findings from a surrogate marker or a specific outcome to the global improvement of the disease. | 0 (0) |
| Conclusion claims the beneficial effect of the experimental treatment despite reporting bias. | 0 (0) |

[a] More than one type of spin may have been present in the same systematic review.

### Data Synthesis

To evaluate the most common forms of spin within systematic review abstracts, we calculated frequencies and percentages. To evaluate particular study characteristics associated with spin in systematic review abstracts, we planned in our protocol, *a priori*, to conduct a logistic regression in which the presence or absence of spin would be coded as a dichotomous criterion variable (0=presence of spin and 1=no spin), and intervention type, journal impact factor, funding source, whether there was medical writer assistance, and whether the journal requires adherence to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) or PRISMA for Abstracts (PRISMA-A) [23,24] would be included in the predictor set. However, the final sample size of 36 systematic reviews suggested that examining the associations between categorical variables using the Chi-square test and point biserial correlations between categorical and continuous variables would be more appropriate, and it was thus decided upon before data analysis was conducted. Only two systematic reviews were written by medical writers; thus, we did not include this variable. Furthermore, none of the systematic reviews reported the use of PRISMA-A, so this variable was also eliminated from the analysis. Interrater reliability was estimated using Gwet's AC1. All statistical analyses were performed using Stata 15.1. [25].

## Results

### Simple Characteristics

Our search string retrieved 316 studies, of which 58 studies were deleted as duplicates and 204 studies were excluded by the title and abstract. The studies excluded from this study are mentioned in Figure 1. We further excluded 20 studies after screening the full text. This screening process left 36 studies that met the inclusion criteria (Figure 1). Among the 36 systematic reviews included, pharmacological intervention for acne vulgaris treatment was the most common intervention type (23/36, 64%), and 31% (11/36) of studies focused on nonpharmacological interventions. The most common source of funding was through public mechanisms (12/36, 33%); however, equally as many studies made no mention of funding (12/36, 33%). The median impact factor was 3.092 for the journals with the systematic reviews included in this study's sample. Detailed characteristics of the included studies can be found in Table 2.

**Figure 1.** Flow diagram of study selection.

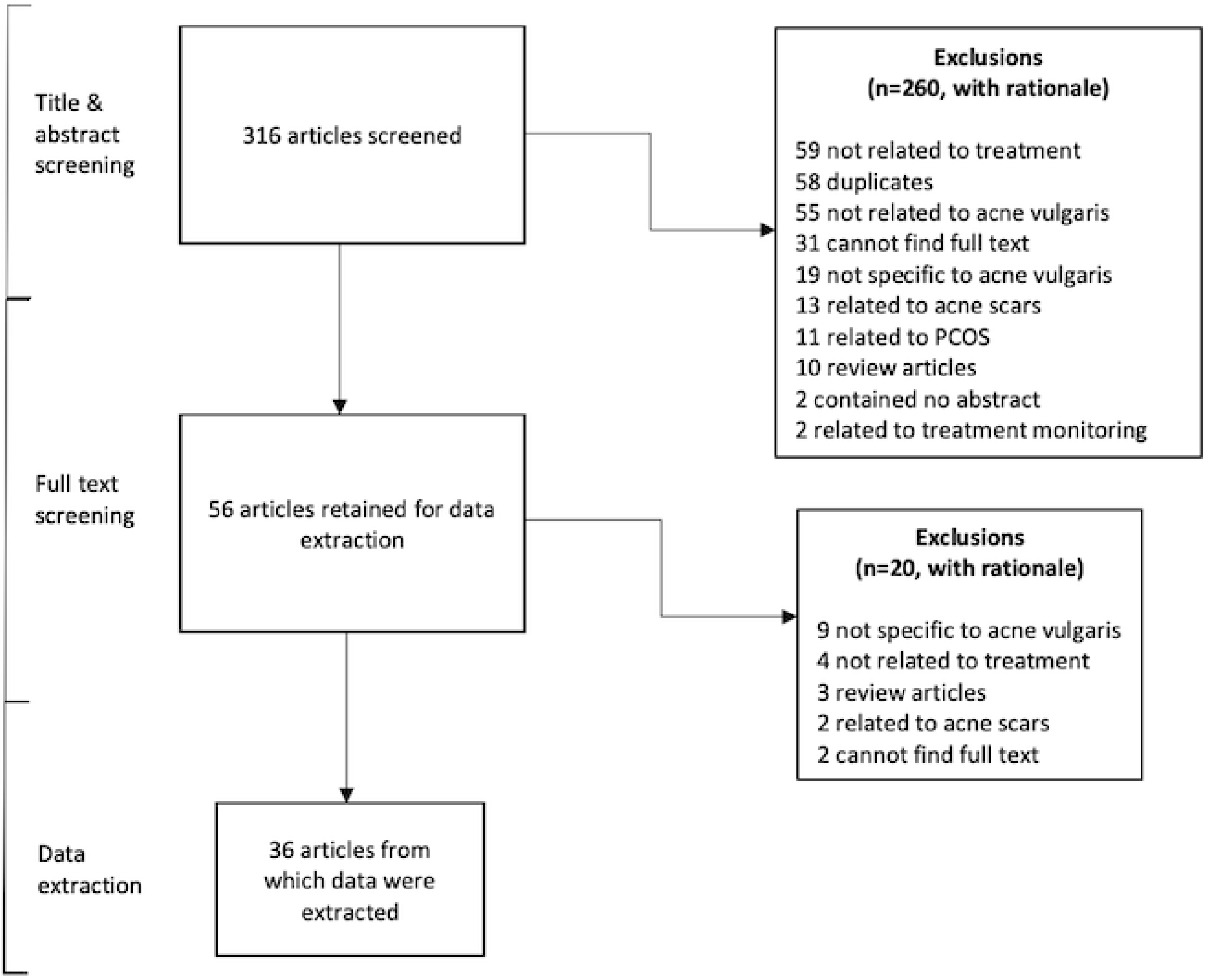


**Table 2.** Characteristics of included studies (N=36).

| Characteristics | Total (N=36), n (%) | With spin (n=11), n (%) |
|---|---|---|
| **Source of funding** | | |
| Public | 12 (33) | 3 (27) |
| Industry | 6 (17) | 2 (18) |
| Not funded | 6 (17) | 3 (27) |
| Funding not mentioned | 12 (33) | 3 (27) |
| **Intervention type** | | |
| Pharmacologic | 23 (64) | 8 (73) |
| Nonpharmacologic | 11 (31) | 3 (27) |
| Combined | 2 (6) | 0 (0) |
| **Journal requirement for Preferred Reporting Items for Systematic Reviews and Meta-Analyses** | | |
| Yes | 21 (58) | 7 (64) |
| No | 15 (42) | 4 (36) |
| **Journal requirement for Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Abstracts** | | |
| Yes | 0 (0) | 0 (0) |
| No | 0 (0) | 0 (0) |
| **Use of a medical writer** | | |
| Yes | 2 (6) | 1 (9) |
| No | 34 (94) | 0 (0) |

### Primary Outcome

For the primary outcome, interrater reliability was assessed using Gwet’s AC1, which was 0.68 (95% CI 0.44-0.92). Each discrepancy was subsequently resolved by group discussion, achieving an agreement of 100%. Spin was determined to be present in 31% (11/36) of abstracts in this study’s sample. A total of 12 separate instances of spin were identified in the 11 abstracts containing spin, with one of the abstracts containing two examples of spin. The most common type of spin identified was type 3, *selective reporting of or overemphasis on efficacy outcomes or analysis favoring the beneficial effect of the experimental intervention* (5/12, 42%); followed by type 5, *conclusion claims the beneficial effect of the experimental treatment despite a high risk of bias in primary studies* (4/12, 33%)*;* type 6, *conclusion extrapolates the review’s findings to a different intervention (ie, claiming efficacy of one specific intervention although the review covers a class of several interventions*; 2/12, 17%); and type 7, *selective reporting of or overemphasis on harm outcomes or analysis favoring the safety of the experimental intervention* (1/12, 8%). Table 1 summarizes this study’s findings for the top nine most severe types of spin in the abstracts of systematic reviews.

Although we found no instances of spin type number 4 regarding safety (Table 1), safety was not mentioned in 61% (22/36) of abstracts. For the purpose of this study, we also considered the discussion of adverse events or treatment side effects as the author inferring safety. A total of 16 studies did not provide a risk of bias assessment (16/36, 44%), and four studies assessed the risk of bias but did not discuss it (4/36, 11%). Of the 11 abstracts that did contain a form of spin, six (55%) did not report a risk of bias assessment or performed a risk of bias assessment but did not discuss it. If an article mentioned a risk of bias assessment but did not provide explanation of risk assessment or data concerning the risk of bias, then we considered this as assessed but not discussed. Spin in the abstracts was not significantly associated with a specific intervention type, funding source, or journal impact factor.

## Discussion

### Principal Findings

Spin is prevalent in the abstracts of systematic reviews and meta-analyses focused on the treatment of acne vulgaris. We identified spin in nearly one-third of the included studies in this study’s sample. The most common type of spin found was type 3, *selective reporting of or overemphasis on efficacy outcomes or analysis favoring the beneficial effect of the experimental intervention*. For example, Koo et al [26] performed a meta-analysis comparing the efficacy of oral antibiotics with that of oral contraceptives for the treatment of acne vulgaris at 3 and 6 months. The oral contraceptive pills were shown to be less effective than oral antibiotics at 3 months but were equivalent to oral antibiotics at 6 months. However, the author suggests in the abstract that oral contraceptive pills may be a better alternative to oral antibiotics while providing no data to support that oral contraceptive pills would be more effective.

We were limited in our assessment of spin type 5, *conclusion claims the beneficial effect of the experimental treatment despite a high risk of bias in primary studies*, as bias was not assessed or was assessed and not discussed, in many of the included studies. This issue is of concern as we noticed six of our 11 studies, which contained spin, either did not assess for the risk of bias or assessed for bias but did not discuss it or further account for it. Providing a risk of bias assessment in systematic reviews is important as it guides the reader in determining the extent to which the included study results should be trusted [27]. Another reason it is essential to perform a risk of bias assessment is that a study may have an important risk of bias, even though the study was performed to the highest possible standards. A few of the reviews included stated that their studies had methodological limitations but did not provide a proper risk of bias assessment. Without a risk of bias assessment, the reader does not have a promising tool to assess the validity of the study. Therefore, it is imperative that systematic reviews perform and discuss a risk of bias assessment of their primary studies.

## Recommendations

To our knowledge, we are the first to investigate the presence of spin in the abstracts of systematic reviews and meta-analyses focused on acne vulgaris treatment options. In the field of dermatology, Motosko et al [28] found that 100% (20/20) of the randomized controlled trials on topical treatments for photoaged skin contained some form of spin, most broadly classified as either inappropriate statistical analyses or inappropriate interpretation of results. Other studies in various fields have previously demonstrated the presence of spin in the abstracts of randomized controlled trials. For example, Boutron et al [29], in the field of oncology, showed that spin in the abstract of a randomized controlled trial can influence clinicians' understanding of the study's results, to the degree that they are more likely to rate a treatment as beneficial, even though the primary outcome is statistically nonsignificant. This finding could have significant clinical implications across all fields of medicine and research in general. A total of four studies conducted by the members of our research team identified spin within abstracts of randomized controlled trials. Austin et al [9] identified spin in the abstracts of nearly half of randomized obesity trials. Cooper et al [10] demonstrated the presence of spin in 70% of the trials published in highly ranked otolaryngology journals. Checketts et al [14] identified spin in 58% of the lower extremity joint trials. Kinder et al [11] reported spin being present in 23% abstracts of the randomized control trials in the field of anesthesiology. Outside of our team's research, others have found similar results. Khan et al [30] reported spin in the abstracts of 53% of the cardiovascular randomized clinical trials. The prevalence of spin in these previously mentioned studies and others led us to ask the question whether spin was also prevalent in the abstracts of systematic reviews. Considering previous studies have shown that abstracts may influence clinical decision making [4-7], the presence of spin in the abstracts of systematic reviews may lead to the dissemination of misinformation and, ultimately, misguided decisions in clinical practice. The steps to minimize spin rest on the shoulders of all research stakeholders, including peer reviewers. Previous studies have demonstrated that reviewers often fail to recognize the misrepresentation of results in study abstracts. For example, Lazarus et al [31] demonstrated that peer reviewers failed to identify spin in abstract conclusions in 76% of the reports reviewed, and 15% of the reviewers actually requested the authors to add some type of spin. Therefore, the task of reducing spin includes the ability to recognize it. Increasing the knowledge of spin could be implemented through education and training so that reviewers, editors, and authors are better equipped to identify and eliminate spin. Guidelines have been developed and published, which will assist reviewers in the detection and interpretation of misleading claims [32,33]. Making use of these guidelines a common practice may help in identifying and eliminating spin.

The PRISMA Statement [34] is an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses, which is widely used by authors, peer reviewers, and editors for reporting and critical appraisal of published systematic reviews [24]. PRISMA-A is an extension to the PRISMA Statement, which provides guidelines for reporting in the abstracts of systematic reviews. Beyond authors and reviewers, we advocate that clinicians, who ultimately make treatment decisions, should be trained to evaluate for the presence of spin in abstracts. Methodologists have performed interesting work to aid in establishing the confidence of the results and conclusions of systematic reviews; however, more work is needed regarding systematic review abstracts. A study performed by Gómez-García et al [35] found that, on average, only 57% of the PRISMA-A items were included in the abstracts of systematic reviews and meta-analyses covering psoriasis treatments. They reported that studies with low-risk bias and high methodological quality had significantly more PRISMA-A items reported compared with studies with high-risk bias and low methodological quality. On the basis of these findings, examining a study's adherence to PRISMA-A may be a useful screening tool for journal editors, reviewers, and clinicians to assess a study's methodological quality and risk of bias. Currently, the PRISMA-A guideline does not specifically address spin; therefore, we recommend mandatory adherence to PRISMA and a revision to the PRISMA extension, PRISMA-A, which focuses on the minimization of spin.

## Strengths and Limitations

This study has several strengths and limitations. Regarding strengths, investigators underwent extensive training to ensure the understanding of the concept of spin in publications. A group consensus was reached on the definition of each of the nine types of spin in an effort to ensure standardized responses. We fostered an atmosphere of reproducibility by posting the full study protocol on OSF before extraction. Each data extractor was blinded during the screening and data extraction process. Data were then extracted in duplicate fashion in accordance with the recommendations of the Cochrane Handbook for Systematic Reviews of Interventions [27]. Despite these efforts, this study is not without limitations. The identification of spin is inherently subjective in nature. To reduce some of this subjectivity, data extractors underwent rigorous training before data extraction and then met to discuss and resolve any discrepancies between extraction responses. A third-party arbitrator was consulted where agreement could not be reached. Another limitation of this study was the low availability of studies that met our inclusion criteria, which limited our sample size. The results of this study are applicable to the field of dermatology with regard to the treatment of acne vulgaris alone. Further studies are needed to investigate the presence of spin in publications regarding other areas of dermatology. Finally, the results of this study should not be generalized across other timelines or publication types other than systematic reviews regarding acne vulgaris treatment options.

## Conclusions

In conclusion, a sizable amount of spin was found in the abstracts of systematic reviews focused on the treatment of acne vulgaris. Moving forward, we recommend that peer reviewers and editors should be educated about the identification of spin. We also recommend the development of strict reporting guidelines for abstracts. Further research is needed to evaluate the types of spin that most significantly affect clinical decision making, which was outside of the purview of this study. We recommend future studies to assess the frequency of spin in other specialties.

## Acknowledgments

All authors of this study had access to all data. This study was not funded by any institution or sponsor.

## Authors' Contributions

RO and TR contributed equally to this paper. RO contributed to data extraction, manuscript writing, manuscript editing, and project design. TR contributed to data extraction, manuscript writing, manuscript editing, and project design. MA contributed to manuscript editing, data collection, analysis, and project design. AJ contributed to data analysis, manuscript editing, and approval. CC contributed to project design, training, and project oversight. MV contributed to project design, manuscript editing, project oversight, and training.

## Conflicts of Interest

None declared.

## References

1. American Academy of Dermatology. Skin Conditions by the Numbers URL: https://www.aad.org/media/stats/conditions/skin-conditions-by-the-numbers [accessed 2019-09-09]
2. Hosthota A, Bondade S, Basavaraja V. Impact of acne vulgaris on quality of life and self-esteem. Cutis 2016 Aug;98(2):121-124. [Medline: 27622255]
3. Gopalakrishnan S, Ganeshkumar P. Systematic reviews and meta-analysis: understanding the best evidence in primary healthcare. J Family Med Prim Care 2013 Jan;2(1):9-14 [FREE Full text] [doi: 10.4103/2249-4863.109934] [Medline: 24479036]
4. Barry HC, Ebell MH, Shaughnessy AF, Slawson DC, Nietzke F. Family physicians' use of medical abstracts to guide decision making: style or substance? J Am Board Fam Pract 2001;14(6):437-442 [FREE Full text] [Medline: 11757886]
5. Marcelo A, Gavino A, Isip-Tan IT, Apostol-Nicodemus L, Mesa-Gaerlan FJ, Firaza PN, et al. A comparison of the accuracy of clinical decisions based on full-text articles and on journal abstracts alone: a study among residents in a tertiary care hospital. Evid Based Med 2013 Apr;18(2):48-53 [FREE Full text] [doi: 10.1136/eb-2012-100537] [Medline: 22782923]
6. Johnson HL, Fontelo P, Olsen CH, Jones MK, Gimbel RW. Family nurse practitioner student perception of journal abstract usefulness in clinical decision making: a randomized controlled trial. J Am Assoc Nurse Pract 2013 Nov;25(11):597-603. [doi: 10.1111/1745-7599.12013] [Medline: 24170534]
7. Haynes RB, McKibbon KA, Walker CJ, Ryan N, Fitzgerald D, Ramsden MF. Online access to MEDLINE in clinical settings. A study of use and usefulness. Ann Intern Med 1990 Jan 1;112(1):78-84. [doi: 10.7326/0003-4819-112-1-78] [Medline: 2403476]
8. Yavchitz A, Ravaud P, Altman DG, Moher D, Hrobjartsson A, Lasserson T, et al. A new classification of spin in systematic reviews and meta-analyses was developed and ranked according to the severity. J Clin Epidemiol 2016 Jul;75:56-65. [doi: 10.1016/j.jclinepi.2016.01.020] [Medline: 26845744]
9. Austin J, Smith C, Natarajan K, Som M, Wayant C, Vassar M. Evaluation of spin within abstracts in obesity randomized clinical trials: a cross-sectional review. Clin Obes 2019 Apr;9(2):e12292. [doi: 10.1111/cob.12292] [Medline: 30576083]
10. Cooper CM, Gray HM, Ross AE, Hamilton TA, Downs JB, Wayant C, et al. Evaluation of spin in the abstracts of otolaryngology randomized controlled trials. Laryngoscope 2018 Dec 21:2036-2040. [doi: 10.1002/lary.27750] [Medline: 30578543]
11. Kinder NC, Weaver MD, Wayant C, Vassar M. Presence of 'spin' in the abstracts and titles of anaesthesiology randomised controlled trials. Br J Anaesth 2019 Jan;122(1):e13-e14 [FREE Full text] [doi: 10.1016/j.bja.2018.10.023] [Medline: 30579417]
12. Reynolds-Vaughn V, Riddle J, Brown J, Schiesel M, Wayant C, Vassar M. Evaluation of spin in the abstracts of emergency medicine randomized controlled trials. Ann Emerg Med 2019 May 14:pii: S0196-0644(19)30230-6. [doi: 10.1016/j.annemergmed.2019.03.011] [Medline: 31101371]

13. Jellison S, Roberts W, Bowers A, Combs T, Beaman J, Wayant C, et al. Evaluation of spin in abstracts of papers in psychiatry and psychology journals. BMJ Evid Based Med 2019 Aug 5:pii: bmjebm-2019-111176. [doi: 10.1136/bmjebm-2019-111176] [Medline: 31383725]
14. Checketts JX, Riddle J, Zaaza Z, Boose MA, Whitener JH, Vassar MB. An evaluation of spin in lower extremity joint trials. J Arthroplasty 2019 May;34(5):1008-1012. [doi: 10.1016/j.arth.2019.01.016] [Medline: 30733070]
15. O'Donohoe TJ, Dhillon R, Bridson TL, Tee J. Reporting quality of systematic review abstracts published in leading neurosurgical journals: a research on research study. Neurosurgery 2019 Jul 1;85(1):1-10. [doi: 10.1093/neuros/nyy615] [Medline: 30649511]
16. Nagendrababu V, Pulikkotil SJ, Sultan OS, Jayaraman J, Peters OA. Methodological and reporting quality of systematic reviews and meta-analyses in Endodontics. J Endod 2018 Jun;44(6):903-913. [doi: 10.1016/j.joen.2018.02.013] [Medline: 29602531]
17. Maticic K, Martinic MK, Puljak L. Assessment of reporting quality of abstracts of systematic reviews with meta-analysis using PRISMA-A and discordance in assessments between raters without prior experience. BMC Med Res Methodol 2019 Feb 14;19(1):32 [FREE Full text] [doi: 10.1186/s12874-019-0675-2] [Medline: 30764774]
18. Tan WK, Wigley J, Shantikumar S. The reporting quality of systematic reviews and meta-analyses in vascular surgery needs improvement: a systematic review. Int J Surg 2014 Dec;12(12):1262-1265 [FREE Full text] [doi: 10.1016/j.ijsu.2014.10.015] [Medline: 25448643]
19. Ottwell RL. OSF. 2019. Evaluation of Spin in the Abstracts of Systematic Reviews Regarding the Treatment of Acne Vulgaris URL: https://osf.io/z59bq/ [accessed 2019-09-18]
20. Cao H, Yang G, Wang Y, Liu JP, Smith CA, Luo H, et al. Complementary therapies for acne vulgaris. Cochrane Database Syst Rev 2015 Jan 19;1:CD009436 [FREE Full text] [doi: 10.1002/14651858.CD009436.pub2] [Medline: 25597924]
21. Barbaric J, Abbott R, Posadzki P, Car M, Gunn L, Layton A, et al. Cochrane systematic review: light therapies for acne. J Am Acad Dermatol 2017;76(6 Suppl 1):AB75 [FREE Full text] [doi: 10.1016/j.jaad.2017.04.307]
22. Rayyan QCRI. URL: https://rayyan.qcri.org/reviews/81224 [accessed 2019-09-10]
23. PRISMA Statement. The PRISMA for Abstracts Checklist URL: http://www.prisma-statement.org/documents/PRISMA%20Abstracts%20Checklist.pdf [accessed 2019-10-09]
24. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
25. Stata: Software for Statistics and Data Science. Version 15. College Station, TX, USA: StataCorp; 2017. URL: https://www.stata.com/ [accessed 2020-02-05]
26. Koo EB, Petersen TD, Kimball AB. Meta-analysis comparing efficacy of antibiotics versus oral contraceptives in acne vulgaris. J Am Acad Dermatol 2014 Sep;71(3):450-459. [doi: 10.1016/j.jaad.2014.03.051] [Medline: 24880665]
27. Higgins J, Green S. Cochrane Handbook For Systematic Reviews Of Interventions. England: Wiley; 2008.
28. Motosko CC, Ault AK, Kimberly LL, Zakhem GA, Gothard MD, Ho RS, et al. Analysis of spin in the reporting of studies of topical treatments of photoaged skin. J Am Acad Dermatol 2019 Feb;80(2):516-22.e12. [doi: 10.1016/j.jaad.2018.04.034] [Medline: 29689326]
29. Boutron I, Altman DG, Hopewell S, Vera-Badillo F, Tannock I, Ravaud P. Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of cancer: the SPIIN randomized controlled trial. J Clin Oncol 2014 Dec 20;32(36):4120-4126. [doi: 10.1200/JCO.2014.56.7503] [Medline: 25403215]
30. Khan MS, Lateef N, Siddiqi TJ, Rehman KA, Alnaimat S, Khan SU, et al. Level and prevalence of spin in published cardiovascular randomized clinical trial reports with statistically nonsignificant primary outcomes: a systematic review. JAMA Netw Open 2019 May 3;2(5):e192622 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.2622] [Medline: 31050775]
31. Lazarus C, Haneef R, Ravaud P, Hopewell S, Altman DG, Boutron I. Peer reviewers identified spin in manuscripts of nonrandomized studies assessing therapeutic interventions, but their impact on spin in abstract conclusions was limited. J Clin Epidemiol 2016 Sep;77:44-51. [doi: 10.1016/j.jclinepi.2016.04.012] [Medline: 27164274]
32. Montori VM, Jaeschke R, Schünemann HJ, Bhandari M, Brozek JL, Devereaux PJ, et al. Users' guide to detecting misleading claims in clinical research reports. Br Med J 2004 Nov 6;329(7474):1093-1096 [FREE Full text] [doi: 10.1136/bmj.329.7474.1093] [Medline: 15528623]
33. Sutherland WJ, Spiegelhalter D, Burgman M. Policy: twenty tips for interpreting scientific claims. Nature 2013 Nov 21;503(7476):335-337. [doi: 10.1038/503335a] [Medline: 24273799]
34. PRISMA Statement. URL: http://www.prisma-statement.org [accessed 2019-05-29]
35. Gómez-García F, Ruano J, Aguilar-Luque M, Alcalde-Mellado P, Gay-Mimbrera J, Hernández-Romero JL, et al. Abstract analysis method facilitates filtering low-methodological quality and high-bias risk systematic reviews on psoriasis interventions. BMC Med Res Methodol 2017 Dec 29;17(1):180 [FREE Full text] [doi: 10.1186/s12874-017-0460-z] [Medline: 29284417]

## Abbreviations

**MEDLINE:** Medical Literature Analysis and Retrieval System Online
**PRISMA:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses
**PRISMA-A:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Abstracts
**OSF:** Open Science Framework

*Edited by G Eysenbach; submitted 08.11.19; peer-reviewed by F Kaliyadan, F Gomez; comments to author 17.12.19; revised version received 04.01.20; accepted 24.01.20; published 20.03.20.*

*Please cite as:*
*Ottwell R, Rogers TC, Anderson JM, Johnson A, Vassar M*
*Evaluation of Spin in the Abstracts of Systematic Reviews and Meta-Analyses Focused on the Treatment of Acne Vulgaris: Cross-Sectional Analysis*
*JMIR Dermatol 2020;3(1):e16978*
*URL: http://derma.jmir.org/2020/1/e16978/*
*doi:10.2196/16978*
*PMID:*

Original Paper

# Cutaneous Bacteria in the Gut Microbiome as Biomarkers of Systemic Malodor and People Are Allergic to Me (PATM) Conditions: Insights From a Virtually Conducted Clinical Trial

Irene S Gabashvili[1], PhD

Aurametrix, MEBO Research, Miami, FL, United States

**Corresponding Author:**
Irene S Gabashvili, PhD
Aurametrix, MEBO Research
13220 SW 35th Terrace
Miami, FL, 33175
United States
Phone: 1 408 341 9355
Email: irene@aurametrix.com

## Abstract

**Background:** The skin is a dynamic ecosystem of microbes and the source of many chemical compounds that affect human health. Skin-microbiome interactions can cause persistent, psychosocially devastating body smell despite good hygiene. Since odor production is often transient, malodors may not be perceptible during medical examinations. Therefore, having odor complaints can be diagnosed as body dysmorphic disorder and referred for psychological evaluations. Development of simple at-home tests and virtual care programs could improve the diagnosis and management of socially debilitating malodor conditions.

**Objective:** The aim of this study was to assess potential effectiveness of at-home gut microbiome testing in the diagnosis and management of idiopathic body and breath odor and in people are allergic to me (PATM) syndrome.

**Methods:** We contacted participants of prior metabolic body odor (MEBO) and PATM studies and online support groups by email or social media. Individuals who consented to participate were mailed test kits for at-home collection of gut microbiome samples. Participants completed an online survey (specially developed for this study) addressing their symptoms and other quality-of-life indicators at baseline and after sampling. Participants collected stool samples after flare-ups or symptom improvements and mailed them to the laboratory to be processed and analyzed. We evaluated between-group differences in symptom severity, as well as symptom improvement observations for the same individuals. For differential abundance testing of microbial taxa, we performed nonparametric statistical analyses using Mann-Whitney *U* tests for unpaired samples and Wilcoxon signed rank test for paired samples.

**Results:** A total of 112 individuals from 21 countries consented to participate. About half the participants had been tested for the metabolic disorder trimethylaminuria, and about half of those tested were diagnosed with the disorder. The levels of bacteria previously associated with cutaneous body odor were significantly elevated in gut samples. For the combination of species from *Anaerococcus*, *Corynebacterium*, *Campylobacter*, and *Propionibacterium* genera, the differences were $P$=.002 for active (73 participants, 182 samples) versus regression or remission groups (30 participants, 51 samples); $P$=.01 for those experiencing symptoms most or all of the time (46 participants, 88 samples) versus those who had symptoms sometimes, rarely, or never (25 participants, 74 samples); and $P$<.001 for improvement of symptoms in the same individuals (22 participants, 43 sets of matched samples). Changes in microbial diversity were significant for between- but not within-participant comparisons.

**Conclusions:** Changes in the gut microbiome composition affect MEBO and PATM severity. In particular, an increase in intestinal bacteria producing odor when in skin flexures was associated with increased intensity of self-reported symptoms. The changes were consistent in the within-group and between-group analyses. Our findings support the feasibility of remote and decentralized clinical studies of malodor conditions. Supplementary sample collection procedures may help to meet established research quality standards.

**Trial Registration:** ClinicalTrials.gov NCT03582826; http://clinicaltrials.gov/ct2/show/NCT03582826

**International Registered Report Identifier (IRRID):** RR2-10.1101/2020.08.21.20179242

***(JMIR Dermatol 2020;3(1):e10508)*** doi:10.2196/10508

**KEYWORDS**

microbiome; idiopathic body odor; systemic malodor; PATM; TMAU; MEBO; bromhidrosis; halitosis; body dysmorphic disorder; virtual care; decentralized clinical trials; diagnosis; management; patient-reported outcome; PRO; at-home test

# Introduction

## Background

The term people are allergic to me (PATM) was coined by a person who believed they were the cause of allergy-like symptoms in surrounding people (see this 2006 post reposted in the PATM support group [1] that received over 8800 responses; Multimedia Appendix 1 shows a screenshot of the webpage). The condition was defined as itchy nose, throat, and eyes in people exposed to the person with PATM, manifesting as sniffing, sneezing, coughing, nose-covering, and throat-clearing behavior. PATM was picked as the name for several online support groups with hundreds of members sharing similar stories (eg, this private Facebook group with over 2100 members [2]). Our molecular diagnostic study conducted in 2009-2012 [3] concluded that PATM may be a subtype of metabolic body odor (MEBO) syndrome. MEBO is another term coined by an affected person to describe idiopathic malodor due to conditions such as trimethylaminuria (TMAU) and other as-yet uncharacterized metabolic disorders. A recent study [4] identified volatile chemicals, eye and nose irritants, toluene, and xylene emitted from the skin of a person with PATM at higher concentrations. Bacterial species recently linked to underarm malodor with sour characteristics in children and teens, *Staphylococcus epidermidis* [5], was found to be overabundant in the nasal cavity of a person with PATM [4]. With the exception of these studies, however, PATM and some cases of self-reported idiopathic malodor are viewed as a dermatological nondisease [6].

The human body can shed a variety of substances that are invisible to the naked eye and are capable of creating discomfort to people in the shedder’s vicinity, depending on their olfactory and immune responses. These substances include odorant molecules stemming from an altered metabolism [7], such as high levels of ketones in diabetic ketoacidosis or buildup of urea in kidney failure. Human body odor is also produced by the microbial degradation of metabolic waste excreted from the body [8,9]. Infections caused by all types of pathogenic organisms—viruses, prokaryotes, and eukaryotes—can produce volatile compounds because of infection-fighting physiological processes in the human body, as well as shifts in the intestinal microflora [9-12]. The odor can linger long after the infection is gone. Microscopic flecks of skin and hair can also become airborne and irritate the respiratory system [13]. Skin flaking can be influenced by the composition of the skin microbiome [14]. Microorganisms are informative indicators of genetic diseases linked to malodor, as well as nongenetic malabsorption and metabolic inefficiencies [15,16].

## Objective

In recent years, cutaneous microbial networks responsible for localized malodors (eg, foot [17], axilla [5,18], neck and head [5]) have been mapped by using next-generation sequencing approaches. These studies used professional assessors to rate odor intensity, and were limited geographically and by numbers of study participants (30 or fewer). This patient-centric study was, to our knowledge, the largest fully virtual microbiome investigation of idiopathic body odor that was open to all of those affected who were interested in participating. Our objective was to assess the potential effectiveness of at-home gut microbiome testing in the diagnosis and management of idiopathic body and breath odor and in PATM syndrome.

# Methods

## Design

The experimental design of this study and the format of the paper follow the Consolidated Standards of Reporting Trials statement for reporting randomized controlled trials and the Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth (CONSORT-EHEALTH) checklist [19]. See Multimedia Appendix 2 for CONSORT-EHEALTH checklist (V 1.6.1). The trial was registered at ClinicalTrials.gov (no. NCT03582826).

## Recruitment

We recruited participants from previous studies (eg, [3,20]), via social media and clinicaltrials.gov (#NCT03582826). Study participants provided electronic informed consent after they were given a complete description of the study, as described in the protocol [21], approved by MEBO Research (Miami, FL, USA) Institutional Review Board on May 11, 2018. Study information and questionnaires were provided in English and Spanish. This research was conducted according to the guidelines established by the Declaration of Helsinki and the International Ethical Guidelines for Biomedical Research Involving Human Subjects. Participants did not receive any incentives.

## Data Collection

Individuals who consented to participate were mailed 3 gut microbiome sampling kits to the address they provided when enrolling. Sampling kits contained 2 sterile polyester swabs, sterile water to prewet the swabs, 2 tubes containing zirconia beads and a lysis and stabilization buffer [22], and sampling instructions (see Multimedia Appendix 1).

Participants completed an online survey (specially developed for this study) addressing their symptoms and other quality-of-life (QoL) indicators at baseline and after sampling. The QoL questionnaire was an open voluntary survey that did not collect personally identifiable information. The survey asked for unique identifiers so the study coordinator could link the data to a prior medical history protected from unauthorized access. The survey also asked for the sample kit ID associated with answers to the survey. Kits were distributed in such a way that researchers from sequence processing facilities did not have access to medical histories and identities of participants. The study coordinator had access to the identifiable information and linked medical histories with test results and answers to the QoL questionnaire.

We assessed MEBO and PATM symptoms in several ways: whether the person thought their condition was in an active state versus regression or remission (asked at the time of enrollment and when collecting a sample), by frequency of symptoms (answering the question “Have you [or your trust buddy] detected any MEBO/PATM symptoms in the past 24 hours or past few days?” on 5-point Likert scale), symptom detection distance (10-point scale), and symptom duration. We added questions assessing the severity of symptoms via related behavior, psychological symptoms (anxiety, stress, depression), social (negative) interactions, physical health, and medical comorbidities to measure internal consistency. We developed the QoL survey in January 2018 and made a few minor iterations incorporating community feedback. We administered the System Usability Scale to a few selected volunteers and found the questionnaire to be acceptable for fielding. Qualitative feedback from participants was continuously obtained through emails and private support groups. A professional health psychologist offered participants personalized behavioral intervention solutions for a wide variety of their mental health needs.

The study’s primary outcome was relative abundance of bacterial classes as measured by operational taxonomic units (OTUs). We categorized these data as active versus regression and remission disease states. Secondary outcomes focused on changes in fecal gut microbiota with symptom improvement.

We used the Strengthening the Reporting of Observational Studies in Epidemiology guidelines and Checklist for Reporting Results of Internet E-Surveys [23] to ensure proper reporting of results (see Multimedia Appendix 3).

### Microbiome Sequencing and Annotation

Participants collected gut microbiome samples after flare-ups or symptom improvements using the kit instructions and mailed the samples in the provided return envelope. Microbial particles were disintegrated, nucleic acid purified, amplified, sequenced in multiplex, demultiplexed, quantified, and assigned to taxonomic membership using previously described methods [24]. In the first step, cells were mechanically disrupted by grinding with glass beads. In the second step, DNA was purified from lysed samples using a liquid-handling robot by a guanidine thiocyanate silica column-based purification method. DNA was amplified with barcoded primers (515F: GTGCCAGCMGCCGCGGTAA; and 806R: GGACTACHVGGGTWTCTAAT) targeting the V4 region of the bacterial 16S ribosomal RNA (rRNA) gene. Indexed polymerase chain reaction) products were pooled by taking the same volume from each reaction, column purified, and size selected through microfluidic DNA fractionation. Consolidated libraries were quantified by quantitative polymerase chain reaction using the Kapa iCycler kit on a MyiQ (Bio-Rad Laboratories, Inc) and sequenced on the NextSeq 500 platform (Illumina, Inc), rendering 2×150 base pair paired-end sequences. After sequencing, demultiplexing of reads according to sample-specific barcodes was performed using Illumina’s BCL2FASTQ algorithm. Reads were filtered using an average Q score greater than 30. After removal of primers and any leading bases, forward and reverse 16S rRNA gene reads were appended together and clustered using version 2.1.5 of the Swarm algorithm, with a distance of 1 nucleotide and the “fastidious” and “usearch-abundance” flags. Depending on the percentage identity, the most abundant forward-reverse read pair per Swarm cluster was assigned taxonomic annotation to the same species, family, order, class, or phylum as in the SILVA V.132 rRNA database [25]. Best hits for the forward and reverse reads with greater than 97% identity to the same sequence in SILVA were annotated to the same species of the hit in SILVA.

### Statistical Analysis

Sequencing data were linked to QoL survey responses using unique identifiers.

We addressed the large variability of the total counts per sample through normalization of raw counts before the analysis, starting from dividing the raw abundances by the total number of counts per sample (number of reads mapped to the taxon divided by the total number of reads mapped to any 16S sequence in the SILVA database). We evaluated several normalization approaches including log ratio. Finally, we applied a centered log-ratio (CLR) transformation to account for the compositional nature of the data. Abundance-weighted phylogenetic diversity was calculated according to McCoy and Matsen [26], by weighting phylogenetic entropy contributed by each lineage by its relative abundance distribution along the rooted phylogenetic tree built for the microbial community. We also estimated other alpha diversity measures, including classical phylogenetic diversity and its partial abundance-weighted extensions.

We identified significant associations between microbial taxa and malodor through a combination of statistical tests, mainly unpaired and paired Mann-Whitney-Wilcoxon tests, herein referred to as Mann-Whitney *U* if used to compare 2 independent groups and Wilcoxon if used for paired dependent samples. OTUs with consistently significant *P* values were adjusted using the Benjamini-Hochberg procedure to avoid type I errors (false positives) and decrease the false discovery rate. All statistical analyses and visualizations were performed using Python v3.7 with the NumPy (v1.16.4), pandas (v0.25.1), scipy (v1.3.1), scikit-learn (v0.22.2), and matplotlib (v3.1.1) toolkits.

## Results

### Participants

Individuals who provided consent received their kits between June and December 2018, answered the QoL questionnaire [21], and sent stool samples for processing between June 2018 and July 2019.

Figure 1 shows the participant flow diagram of the study. A total of 112 individuals consented to participate and were mailed 3 gut microbiome sampling kits to the address they provided when enrolling. Volunteers represented 21 countries: the United States, United Kingdom, Argentina, Brazil, Burkina Faso, Canada, Colombia, France, Hong Kong, Italy, Kenya, Mexico, Morocco, the Netherlands, Nigeria, Pakistan, Peru, Philippines, South Africa, Spain, and Sweden. Of these, 4 participants were not able to receive the kits because of unexpected customs fees or change in address. A total of 38 did not follow up, but 1 of them submitted results of prior microbiome tests, joining 13 others who did not participate in the study but volunteered the data (from the United States, United Kingdom, and Portugal). Of these, 6 did not identify themselves with MEBO or PATM. As a result, we collected 233 samples from 84 participants, including 13 who did not enroll to participate in the study but volunteered the data and 112 who consented to participate, minus the 4 participants who did not receive the kits, 2 who opted out, and 35 who neither followed up on survey responses nor sent their samples.

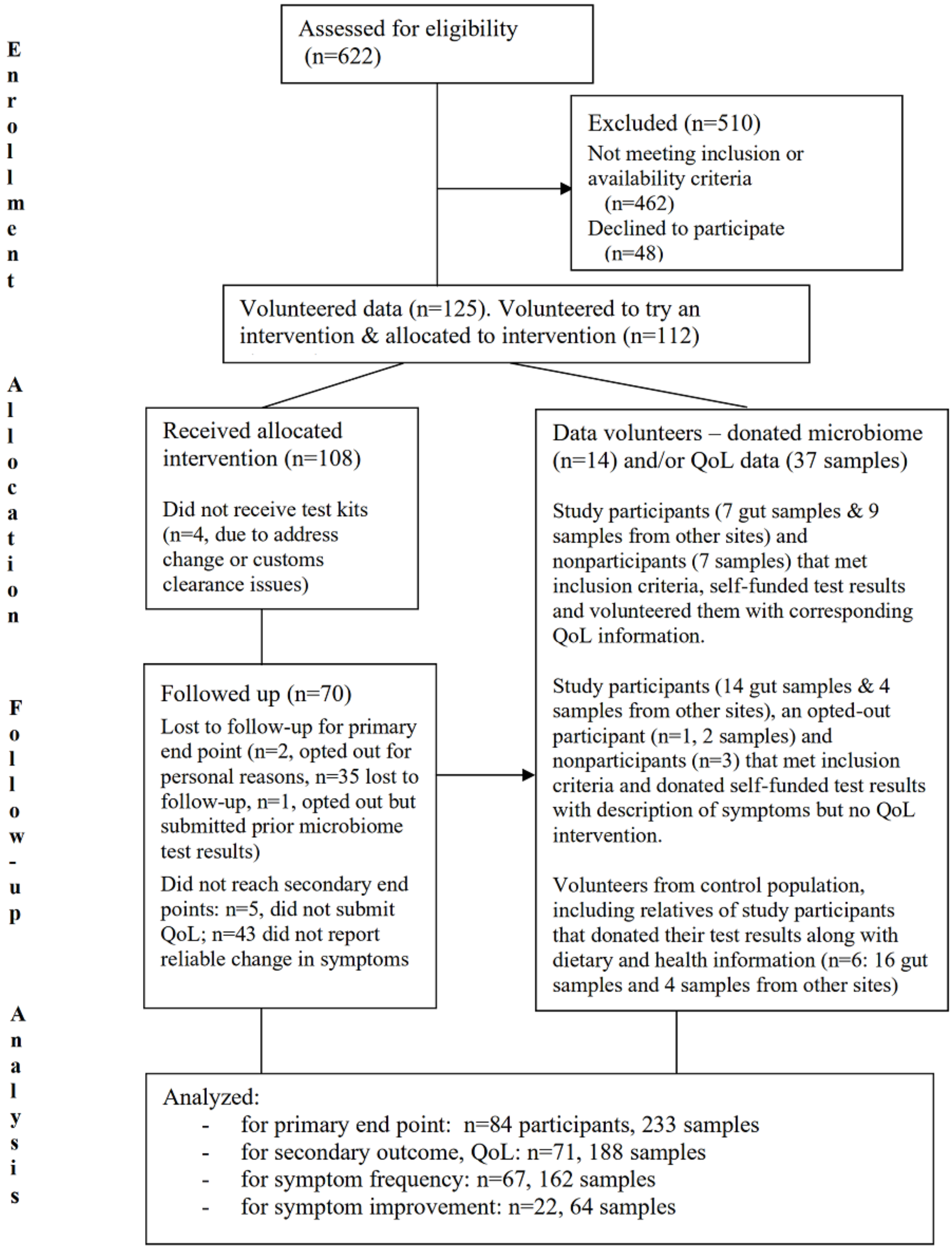


**Figure 1.** Participant flow through the metabolic body odor/people are allergic to me (MEBO/PATM) Microbiome trial. QoL: quality of life.

We analyzed the resulting data by first grouping them into disease states: active versus remission, regression, or healthy controls (never experienced MEBO or PATM). Of the noncontrol participants, 71 submitted answers to the QoL questionnaire [21] for 188 samples. A total of 49 participants could not comment on their symptoms because of their inability to smell, unavailability of a “trust buddy” to objectively evaluate their condition, or no change in their symptoms. The remaining 22 participants reported both flare-ups and improvements in their symptoms that we could use in a paired comparison analysis.

### Baseline Characteristics

Of 336 test kits distributed to 112 consenting individuals, 189 samples were mailed for analysis by 73 participants (73/112, 65.2%; 189/336 distributed kits, 56.3%). Unfortunately, 4 of these samples (3 participants) were not processable, and 8 participants who submitted valid samples did not answer the QoL questionnaire.

Intervention dropout rates were similar to previously reported rates, and more of those who discontinued the study were women [27]. All nonrespondents received several personalized email reminders, but this did not significantly improve survey response and sample submission rates. On the other hand, 15 compliant participants from the intervention group submitted 21 additional samples using kits they purchased on their own. One participant opted out of the intervention but submitted 2 samples using previously purchased kits. In addition, 6 non-MEBO and 8 MEBO volunteers who did not participate in the intervention contributed 27 gut samples to this study. As a result, we collected 233 samples from 84 participants (Figure 1). Table 1 shows the characteristics of these individuals.

**Table 1.** Baseline demographic and clinical characteristics of study participants who submitted valid samples (N=84, unless otherwise stated).

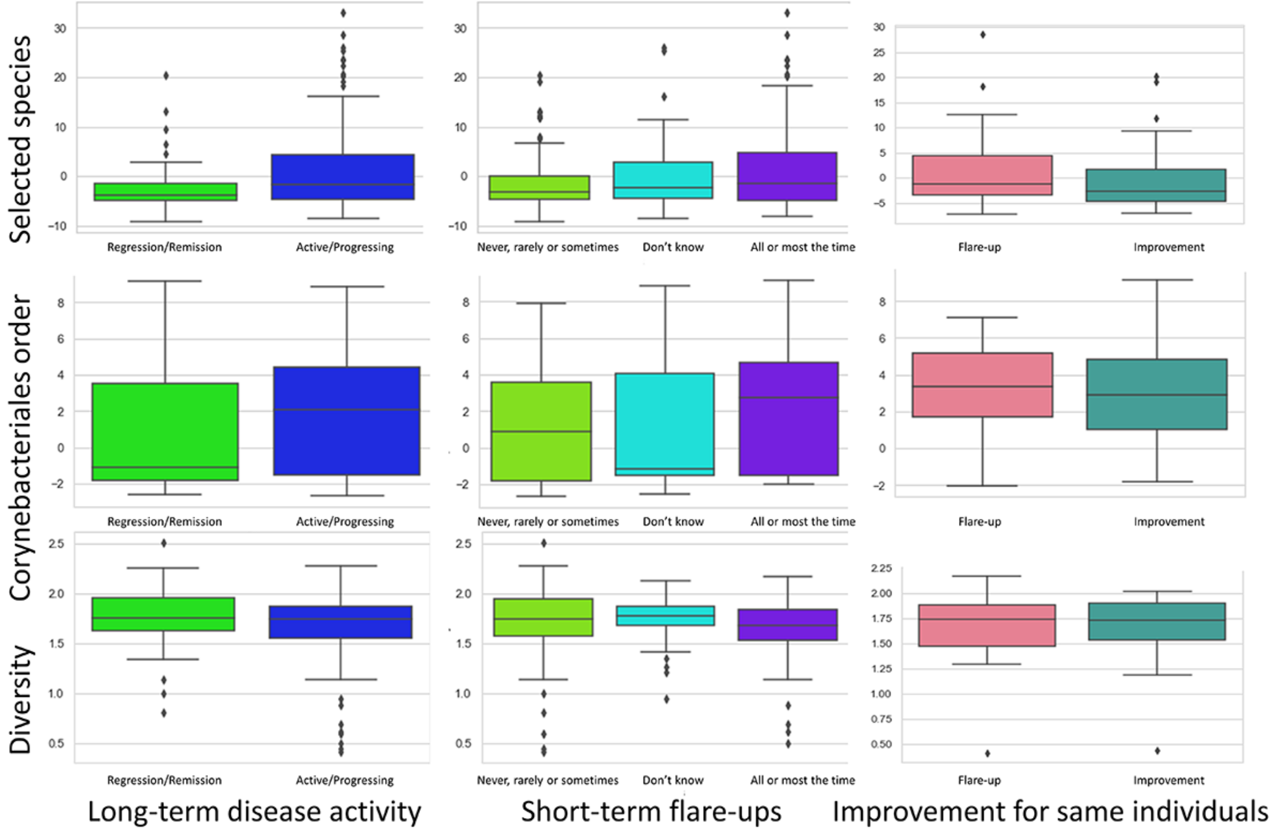
| Characteristics | Values |
|---|---|
| **Age (years), mean (SD)** | |
| All | 40 (12) |
| Female | 42 (11) |
| Male | 37 (12) |
| **Sex, n (%)** | |
| Male | 34 (40) |
| Female | 50 (60) |
| **Trimethylaminuria, n (%) (n=41)** | |
| Positive | 21 (51) |
| Negative | 20 (49) |
| Self-diagnosis of PATM, n (%) (n=77) | 34 (44) |

Of the 84 study volunteers, 41 (49%; 41/78, 53% of MEBO/PATM cohort) tested for the odor-producing disorder TMAU, with 13 positives for primary TMAU and 8 for secondary TMAU, a nongenetic form arising from dysbiosis in the gut bacteria [15]. Of 78 MEBO/PATM participants, only 10 were confident about the primary source of their malodor: in 9 (12%) cases it was nose or mouth, or both, and in 1 (1%) it was the genital area. Of these individuals, 6 thought they also had body odor in other areas. A total of 43 (55%) participants described malodor or PATM toxins emanating from their entire body, not just selected areas. Many participants commented that odor has to build up to be noticed (eg, after spending more than 15 minutes in a closed room with bad air circulation). The remaining 25 (32%) study participants named body sites they thought could be contributing to their malodor: underarm, feet, genitals, scalp, and face (mostly the oily T-zone), in all possible combinations thereof. Several participants were officially diagnosed with hyperhidrosis and bromhidrosis, while in 2 cases bromhidrosis was ruled out by dermatologists. In 2 other cases, patients were treated with botulinum toxin A (Botox) injected into the underarm but they continued to experience body odor after this procedure. In all these instances, patients thought that underarm odor was not their only odor.

### Evaluation of Outcomes

We performed Mann-Whitney *U* tests to identify OTUs that showed significantly different frequencies between distinct groups in our study. Remarkably, microbes known to cause malodor when present on skin were consistently significantly overabundant in active versus regression or remission states of the condition and on acute flare-up versus nonflare-up days. In particular, these included *Corynebacterium* species thought to be the primary causal agents of axillary odor dependent on secretions of the apocrine gland [18,28,29], *Anaerococcus* species also found to correlate with cutaneous odor formation [29], commensal skin bacteria *Cutibacterium* and *Propionibacterium* contributing to foot odor [30], and *Campylobacter* contributing to axillary, in addition to oral, odor in some individuals [18,31]. Figure 2 shows boxplots for microbial diversity, Corynebacteriales, and a combination of selected species from the abovementioned bacterial genera (selected cutaneous species), namely *Anaerococcus* species S9 PR-5, *Anaerococcus hydrogenalis*, *Anaerococcus lactolyticus*, *Anaerococcus octavius*, *Anaerococcus prevotii*, *Anaerococcus provencensis*, *Campylobacter hominis*, *Corynebacterium atypicum*, *Corynebacterium durum*, *Corynebacterium freiburgense*, and *Propionibacterium freudenreichii* species.

**Figure 2.** Changes in centered log-ratio–transformed abundances of selected bacterial species, Corynebacteriales, and microbial diversity with changes in long-term disease state (regression/remission vs active state), short-term flare-ups (24-hour symptoms observed never/rarely or sometimes vs all/most of the time), and improvement of symptoms for the same individuals. The central line in each box is the median, the upper and lower boundaries of the box mark the first and third quartiles, the thin lines show the lowest and largest data points excluding any outliers, and the diamonds show the outliers.



In short-term flare-ups, symptoms were grouped into 3 categories to reduce the noise of subjective self-reporting (all or most of the time vs never, rarely, or sometimes vs those who could not objectively evaluate if they experienced episodes of malodor in the past few days). It is possible to see differences among all answers on a 5-point scale (“all the time,” “most of the time,” “sometimes,” “rarely,” and “never”) as Multimedia Appendix 1 shows.

Table 2 shows the results of paired and unpaired tests for different groups of participants. The difference in the selected cutaneous species defined above is statistically significant independently of analytical approaches used for data preprocessing and normalization and after the false discovery rate correction (adjusted *P*<.001). For pairwise comparison of samples before versus after improvement for the same individuals, the 95% CI is 26-163 for normalized raw counts, 0.14-1.6 for logarithm to base 10 of relative abundance in participants’ stool, and 1.4-56.2 for CLR-transformed counts. For the Corynebacteriales order, the adjusted *P*<.02. Microbial diversity did not significantly change with improvement of symptoms. We note that these 22 participants submitted 73 samples, donating additional kits they purchased on their own. A total of 64 samples out of 73 were accompanied by QoL questionnaires. As a result, 6 participants submitted only 2 samples corresponding to “better” and “worse” states (6 pairs); 12 participants submitted 3 samples and 4 submitted 4 samples that could be paired in 37 different ways, since some individuals were able to recognize their symptoms at a fine-grained level. Hence, we report data for the 43 pairs of samples, although we computed Wilcoxon signed rank tests for several different pairings. *P* values were significant in all scenarios, including the smallest 22-pair set representing the largest improvement observed for each individual (*P*=.001for normalized counts and *P*=.004 for CLR-transformed counts of cutaneous bacteria). We investigated long-term disease activity by comparing participants with different overall MEBO/PATM status, self-reported in the QoL survey. We evaluated short-term effects by comparing 24-hour recalls. The selected cutaneous species index is higher for participants in the active state versus remission or regression and in those who were experiencing symptoms all or most of the time versus never or rarely or sometimes. Adjusted *P* values are higher in these cases, which could be because nonmatched individuals self-assessing their symptoms are more problematic to compare, but the difference is still significant or marginally significant (*P*<.01 for long term and *P*=.06 for short term activity). The 95% CI for CLR-transformed counts is 0.6-3.45 for long-term effects and 0-2.7 for short-term effects. For 24-hour recall, between-group comparisons of changes in Corynebacteriales and microbial diversity are significant even after false discovery rate adjustment. Table 2 also lists results for potential confounding factors, such as females versus males, younger versus older individuals, and those who mentioned underarm odor as one of their problem odors versus those who ruled the axilla out of their odor sources. Differences between these groups were not significant after false discovery rate adjustment and only marginally significant otherwise. We note that individuals who self-diagnosed PATM had significantly higher levels of cutaneous species than those who self-reported malodor only (95% CI 0.5-3.36 for CLR cutaneous species). Yet we see the same patterns for disease severity within each of the groups as within the entire population (Figure 3).

**Table 2.** *P* values for Mann-Whitney *U* tests for unpaired samples and Wilcoxon signed rank test for paired samples.

| Test, groups, and preprocessing | Selected species | Corynebacteriales | Microbial diversity |
|---|---|---|---|
| **Before and after improvement (Wilcoxon test for 43 pairs of matched samples of 22 participants)** | | | |
| Normalized counts | <.001 | <.001 | .6 |
| CLR[a]-transformed | <.001 | .02 | |
| **Active vs regression/remission/no disease status (Mann-Whitney *U* test for 182 samples of 73 participants vs 51 samples of 30 participants)[b]** | | | |
| Normalized counts | .004 | .05 | .3 |
| CLR-transformed | .002 | .008 | |
| **Experienced MEBO[c]/PATM[d] episodes all/most of the time vs never/rarely/sometimes (Mann-Whitney *U* test for 88 samples of 46 participants vs 74 samples of 25 participants)[b]** | | | |
| Normalized counts | .01 | .005 | .02 |
| CLR-transformed | .04 | .002 | |
| **Females vs males: 138 vs 95 samples, 84 participants** | | | |
| Normalized counts | .3 | .2 | .3 |
| CLR-transformed | .09 | .2 | |
| **Individuals with underarm odor vs those who ruled it out: 36 samples/13 participants vs 54 samples/15 participants** | | | |
| Normalized counts | .02 | .07 | .08 |
| CLR-transformed | .07 | .04 | |
| **Participants with MEBO (121 samples/44 participants) vs those with self-diagnosed PATM (96 samples/34 participants)** | | | |
| Normalized counts | .09 | .1 | .2 |
| CLR-transformed | .008 | .2 | |
| **Those who tested negative for trimethylaminuria (58 samples/20 participants) vs positive (56 samples/21 participants)** | | | |
| Normalized counts | .2 | .4 | .3 |
| CLR-transformed | .4 | .5 | |
| **Age <40 years (115 samples/39 participants) vs ≥40 years (118 samples/45 participants)** | | | |
| Normalized counts | .05 | .4 | .1 |
| CLR-transformed | .05 | .4 | |

[a] CLR: centered log-ratio.

[b] The sum of participants in active disease versus remission (73+30>84) and “most of the time” versus “sometimes” groups (46+25>67) is greater than the total number of participants because some individuals changed their answers to surveys associated with follow-up samples.

[c] MEBO: metabolic body odor.

[d] PATM: people are allergic to me.

**Figure 3.** Change in abundance of selected cutaneous species (log scale) with long- and short-term improvement of symptoms in metabolic body odor (MEBO) vs people are allergic to me (PATM) groups of study volunteers. Red boxes represent active state of disease or symptoms observed all or most of the time. Blue boxes show remission or regression and symptoms reported as never, rarely, or sometimes.

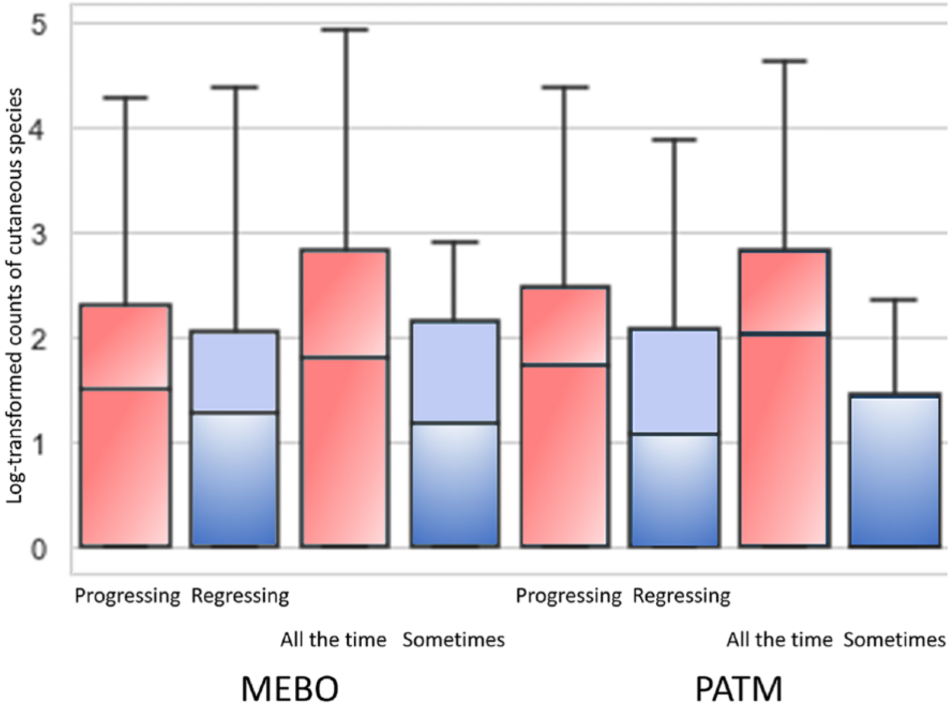


Figure 4 shows similar trends in all other groups, except cases when underarm odor was ruled out as a potential problem. Changes in abundances of cutaneous bacteria are not significant in remission only in the latter case.

Figure 5 shows how the selected cutaneous species index changed with symptom improvements in the 22 participants (12 female and 10 male) who reported flare-ups and improvements. Figure 6 shows changes in genera for 1 of the volunteers, F7, who reported fine-grained symptoms for multiple observations (refer to Multimedia Appendix 1 for more details).

**Figure 4.** Change in selected species of bacteria (centered log-ratio [CLR]–transformed cutaneous species) with long-term improvement of symptoms analyzed separately for different subgroups of study participants: males and females, those including or excluding underarm odor in self-reported symptoms, and participants grouped into 2 age brackets: 20-39 and 40-66 years. Red boxes represent active state of disease. Blue boxes show remission or regression.

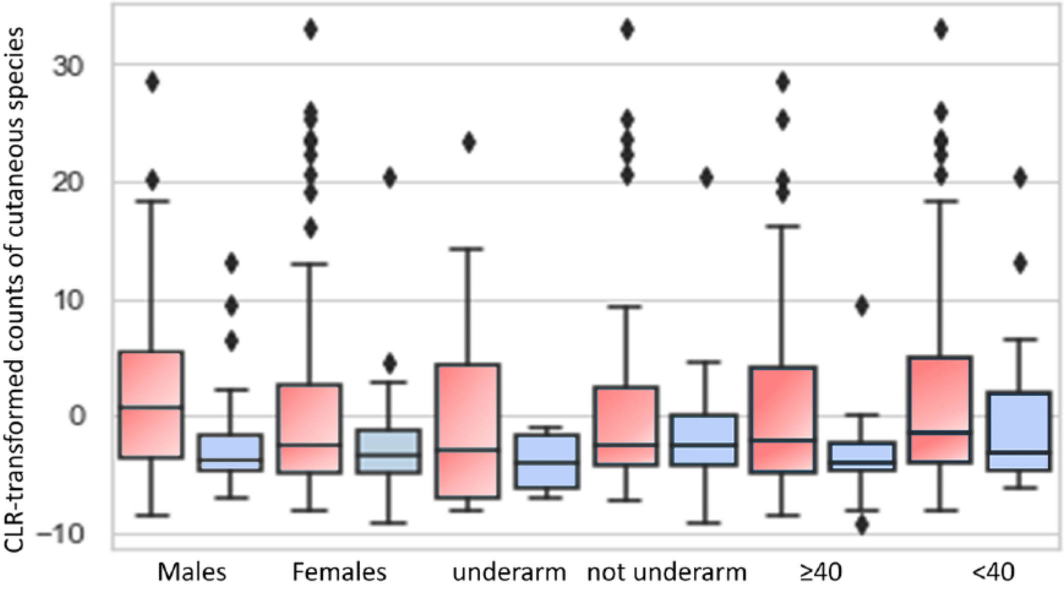


**Figure 5.** Changes in microbial diversity vs abundances of selected bacterial species (CSS) for 12 female (F) and 10 male (M) participants who self-reported both flare-ups (beginning of the arrow) and improvements (end of the arrow).

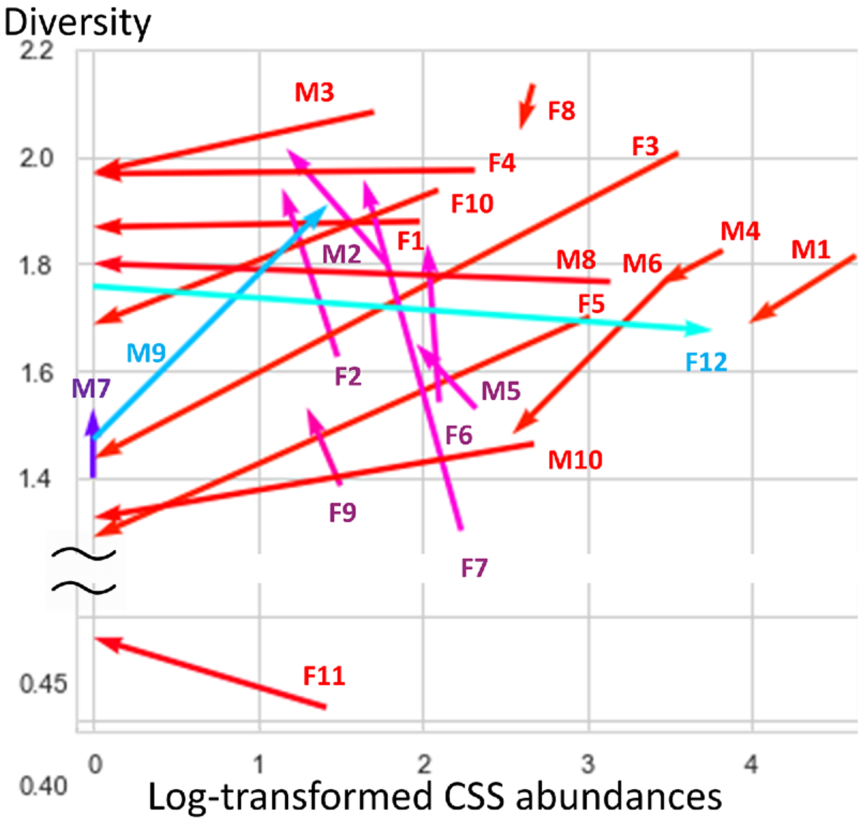
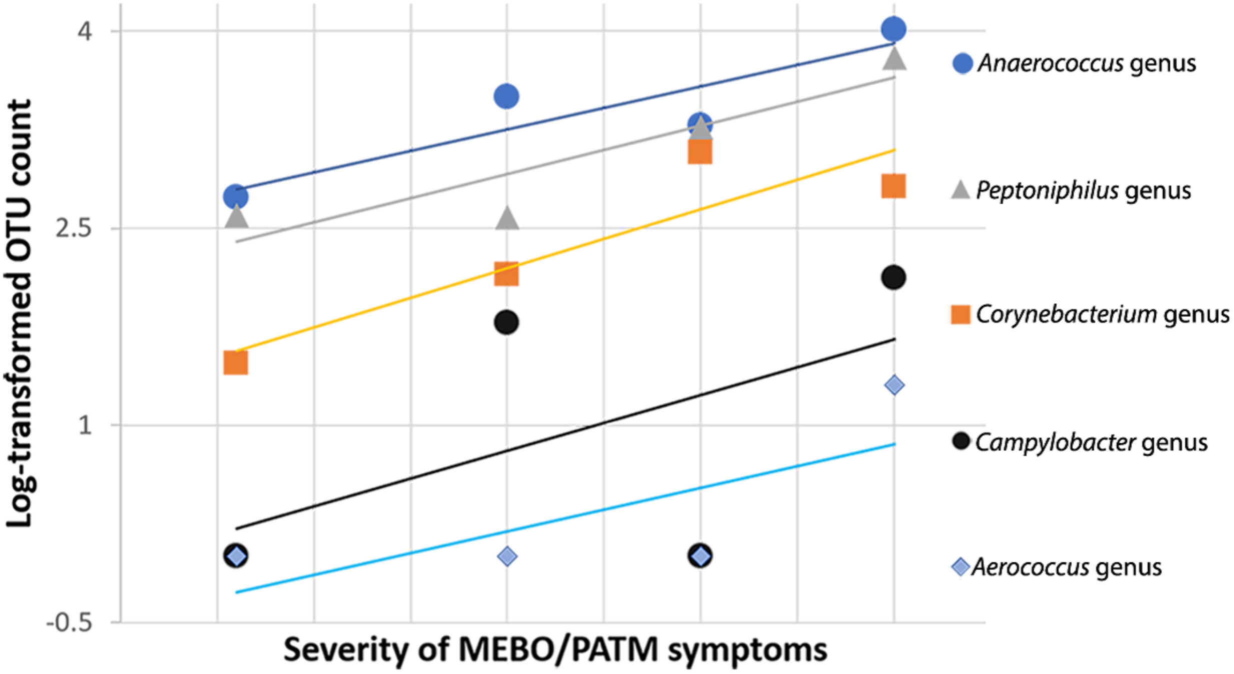


**Figure 6.** Changes in abundances of common dermal bacteria in the gut microbiome of participant F7. Each genus represents the total number of reads mapped to species identified in the gut sample. MEBO: metabolic body odor; OTU: operational taxonomic unit; PATM: people are allergic to me.



## Discussion

### Principal Findings and Prior Work

Previous studies have demonstrated differences in the skin and oral microbiomes of individuals experiencing body odor or halitosis [5,8,17,18,28-31]. The gut microbiota is linked to skin disorders such as psoriasis, rosacea, and atopic dermatitis [32-34] and is a target for the treatment of MEBO [35]. Yet prior studies of malodor had limited sample sizes and were lacking insights into the gut microbial community [36], focusing on skin (including vulvar) and oral microbial compositions. In this study, we observed significant differences in cutaneous gut microbes (able to colonize skin and cause skin malodor) between various groups of MEBO and PATM conditions. Levels of these bacteria were always significantly lower with improvement of symptoms, whether long- or short-term or when observed for the same individuals. There was a moderately positive relationship between answers to the QoL survey about the progression of the MEBO or PATM condition and the intensity of recent symptoms. However, since all individuals answered the question about the state of their condition (84 participants), whereas only 80% (67 participants) reported daily observations and many were not sure about some aspects of their condition, we analyzed multiple illness severity grouping systems. Remarkably, cutaneous malodor bacteria were reproducibly associated with increasing intensity of MEBO and PATM. Based on survey responses, we noticed differences in coping mechanisms between males and females and associations between answers such as “My appearance was affected because of MEBO/PATM” and “I had problems concentrating.” In our future research, we plan to investigate the QoL outcomes and data sets from our past diagnostic studies at a more granular level.

Results of case-by-case studies aligned with overall statistics. As Figure 5 shows, the only exceptions to the conclusion that the fewer cutaneous bacteria in the gut, the fewer skin emanations were M7, M9, and F12. All of them observed very minor if not negligible (and easy to misinterpret) improvement of their condition (flare-ups happening from “all the time” to “most of the time”). M7 was seen by a professional dermatologist, who concluded that a diagnosis of bromhidrosis didn’t seem warranted. F12 had undergone a Botox procedure to treat her hyperhidrosis about 15 years previously. M9 did not report any skin odors and noted only halitosis. It is interesting to note that, whereas those with higher abundances of cutaneous bacteria in the gut benefited from reducing their microbial diversity and bacterial counts in general, those with lower abundances benefited from increasing their microbial diversity. As in a recent study investigating differences in diversity with respect to human personality and other parameters of daily living [37], we, too, found sociability associated with higher diversity, and anxiety, depression, and stress with reduced diversity. We also analyzed differences in fiber, calcium, fat, vegetable, and heterocyclic amine intake, as well as of fermented foods and probiotic supplements. However, the impact of these factors was more profound on microbial diversity than on counts of cutaneous species in the gut microbiome. In some of these cases, such as probiotic supplements, abundances of cutaneous species were significantly lower in both flare-ups and remissions, but the impact of these bacteria remained the same: higher levels with more severe symptoms, similar to the data shown in Figures 3 and 4.

As observed earlier, abundances of common skin bacteria seemed to correlate with each other. Figure 6 shows similar effects on self-reported severity of symptoms observed for genera responsible for axillary and foot odor: *Corynebacterium*, *Anaerococcus*, *Peptoniphilus*, *Campylobacter*, and *Aerococcus*.

### Limitations

The challenges of relying solely on perception of odors by nonprofessionals were previously discussed [5]. In this study [5], “malodor” and “no malodor” groups of children defined by their parents were not statistically significantly different when evaluated by professional odor assessors (Wilcoxon test $P>.1$). The self-reported nature of this study is a limitation, but medical examination may also be prone to errors and the consistency of our findings for different groups and with data from other studies demonstrates that fully remote studies of malodor are feasible.

Measuring the microbiome is not like measuring a single isolated variable pertaining to physiology. Since the data are complex and compositional, daily fluctuations get amplified, requiring multiple longitudinal sampling for patients. Multiomic measurements can improve statistical inferences, avoid false positives, and increase the overall efficiency of clinical trials.

Another limitation of this study was that we did not have data on the time period from sample collection to sequencing for all kits processed. The storage condition of gut samples introduced some biases in microbial diversity and the relative abundance of functional bacteria, even when using the best commercial stabilizers for preserving fecal samples at room temperatures [38]. However, most storage conditions and storage time had minor and acceptable impacts on nucleic acid yields or quality.

An additional limitation of this study was that we did not use incentives for participation.

### Conclusions

The hypothesis of this study was that in spite of genetic heterogeneity, MEBO and PATM conditions display common patterns in the gut microbiome [21]. The collective intestinal microbiome in the study population stratified by disease severity exhibited a relative increase in bacterial species from genera associated with dermal odor. The finding was consistent in active versus remission states, and short-term flare-ups versus relative improvements.

Skin conditions often lead to difficulties in emotional, psychological, and social functioning. Our study showed that web-based research has the potential to not only offer much-needed psychosocial support [39], but also help to develop virtual care solutions for conditions resembling dermatological nondisease [40].

Our work paves the way for the development of cost-effective diagnostics of MEBO and PATM conditions based on an at-home stool test. Current methods (such as trimethylamine N-oxide urine or small intestinal bacterial overgrowth breath diagnostics) rely on tolerance tests that take many hours of preparation, testing, and recovery. External microbiome sampling may require multiple swabs, such as from the oral cavity, forearm, axilla, scalp, and feet. Simple at-home microbiome stool sampling could simplify testing.

Another implication of this work is that individuals with a high abundance of cutaneous malodor bacteria in the gut might benefit from reducing levels of these bacteria. We note that one of the common causes of both halitosis and axillar odor is a zinc deficiency. Zinc oxide is known to decrease populations of bacteria discussed in this work when applied externally [41], attenuating self-perceived malodor. Zinc also contributes to the reduction of halitosis [42]. However, not all people with MEBO and PATM are zinc deficient, and it is important to retain certain levels of cutaneous odor-producing bacteria in the body, as they prevent some skin, ear, and respiratory infections [43].

Future work will focus on incorporating more background knowledge and test results for higher precision data mining. Additional studies are needed to uncover the root causes of socially debilitating malodor and PATM conditions and, most importantly, to connect patient experiences to the development of personalized therapies.

## Acknowledgments

I thank Maria de la Torre, the study coordinator, who worked hard to recruit and educate study volunteers from all over the world, and Crissan Rosalia for offering professional psychological support. We are very grateful to all participants who went the extra mile to ensure the success of this study. We thank uBiome for processing the samples and Melissa Agnello for all the help she gave us to complete this challenging project.

uBiome funded sample processing; MEBO Research funded postage to international participants; and several study volunteers self-funded their microbiome tests. The principal investigator, study coordinator, and support personnel worked pro bono.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials.
[DOCX File , 1274 KB - derma_v3i1e10508_app1.docx ]

Multimedia Appendix 2
CONSORT-EHEALTH checklist V1.6.1.
[PDF File (Adobe PDF File), 3553 KB - derma_v3i1e10508_app2.pdf ]

Multimedia Appendix 3
Checklist for Reporting Results of Internet E-Surveys (CHERRIES).
[DOCX File , 28 KB - derma_v3i1e10508_app3.docx ]

## References

1. MedHelp. People are allergic to me (PATM).: Vitals Consumer Services LLC; 2007 Apr 30. URL: https://www.medhelp.org/posts/People-Allergic-to-Me-PATM/People-Are-Allergic-To-Me-PATM/show/4415 [accessed 2020-10-26]
2. Facebook. MEBO PATM/TMAU. URL: https://www.facebook.com/groups/715913928557803 [accessed 2020-10-26]
3. Gabashvili I. Community-led research discovers links between elusive symptoms and clinical tests. bioRxiv 2017 Jan 01. [doi: 10.1101/139014]
4. Kawakami Y, Sekine Y, Kimura K, Todaka M, Hisayuki ODA. An exploratory study on "people allergic to me" syndrome based on measurement of trace gases released from skin surface and nasal microbial species in nasal cavity of a patient. Indoor Environ 2018;21(1):19-30 [FREE Full text] [doi: 10.7879/siej.21.19]
5. Lam TH, Verzotto D, Brahma P, Ng AHQ, Hu P, Schnell D, et al. Understanding the microbial basis of body odor in pre-pubescent children and teenagers. Microbiome 2018 Nov 29;6(1):213 [FREE Full text] [doi: 10.1186/s40168-018-0588-z] [Medline: 30497517]
6. Bian Y, Ma B, Wang Z, Yang F. People allergic to me and body dysmorphic disorder. Asian J Psychiatr 2019 Aug;44:61-62. [doi: 10.1016/j.ajp.2019.07.031] [Medline: 31325759]
7. Broza YY, Zuri L, Haick H. Combined volatolomics for monitoring of human body chemistry. Sci Rep 2014 Apr 09;4:4611 [FREE Full text] [doi: 10.1038/srep04611] [Medline: 24714440]
8. Shelley WB. Axillary odor. AMA Arch Dermatol Syphilol 1953 Oct 01;68(4):430. [doi: 10.1001/archderm.1953.01540100070012]
9. Elmassry MM, Piechulla B. Volatilomes of bacterial infections in humans. Front Neurosci 2020;14:257 [FREE Full text] [doi: 10.3389/fnins.2020.00257] [Medline: 32269511]

10. Ahmed WM, Lawal O, Nijsen TM, Goodacre R, Fowler SJ. Exhaled volatile organic compounds of infection: a systematic review. ACS Infect Dis 2017 Oct 13;3(10):695-710. [doi: 10.1021/acsinfecdis.7b00088] [Medline: 28870074]
11. Porter SR, Scully C. Oral malodour (halitosis). BMJ 2006 Sep 23;333(7569):632-635 [FREE Full text] [doi: 10.1136/bmj.38954.631968.AE] [Medline: 16990322]
12. Beatty JK, Akierman SV, Motta J, Muise S, Workentine ML, Harrison JJ, et al. Giardia duodenalis induces pathogenic dysbiosis of human intestinal microbiota biofilms. Int J Parasitol 2017 May;47(6):311-326 [FREE Full text] [doi: 10.1016/j.ijpara.2016.11.010] [Medline: 28237889]
13. Hampton SF, Cooke RA. The sensitivity of man to human dander, with particular reference to eczema (allergic dermatitis). J Allergy 1941 Nov;13(1):63-76. [doi: 10.1016/s0021-8707(41)90008-4]
14. Gaitanis G, Magiatis P, Hantschke M, Bassukas ID, Velegraki A. The Malassezia genus in skin and systemic diseases. Clin Microbiol Rev 2012 Jan;25(1):106-141. [doi: 10.1128/CMR.00021-11] [Medline: 22232373]
15. Fennema D, Phillips IR, Shephard EA. Trimethylamine and Trimethylamine N-Oxide, a Flavin-Containing Monooxygenase 3 (FMO3)-Mediated Host-Microbiome Metabolic Axis Implicated in Health and Disease. Drug Metab Dispos 2016 Nov 02;44(11):1839-1850 [FREE Full text] [doi: 10.1124/dmd.116.070615] [Medline: 27190056]
16. Pinheiro de Oliveira F, Mendes RH, Dobbler PT, Mai V, Pylro VS, Waugh SG, et al. Phenylketonuria and gut microbiota: a controlled study based on next-generation sequencing. PLoS One 2016;11(6):e0157513 [FREE Full text] [doi: 10.1371/journal.pone.0157513] [Medline: 27336782]
17. Wulandari NF, Suharna N, Yulinery T, Nurhidayat N. Preliminary study on bacterial diversity causing human foot odor. 2019 Presented at: International Symposium of Innovative Bio-Production Indonesia on Biotechnology and Bioengineering; October 23-24, 2019; Tangerang, Indonesia p. 23-24.
18. Troccaz M, Gaïa N, Beccucci S, Schrenzel J, Cayeux I, Starkenmann C, et al. Mapping axillary microbiota responsible for body odours using a culture-independent approach. Microbiome 2015;3(1):3. [doi: 10.1186/s40168-014-0064-3] [Medline: 25653852]
19. Eysenbach G, Consort-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
20. Gabashvili IS. Identifying subtypes of a stigmatized medical condition. medRxiv 2019 Aug 29. [doi: 10.1101/19005223]
21. Gabashvili IS. Dynamics of the gut microbiota in MEBO and PATM conditions: protocol of a fully remote clinical study. medRxiv 2020 Aug 24. [doi: 10.1101/2020.08.21.20179242]
22. Vera-Wolf P, Cárdenas J, Morton A, Norambuena T, Torres R, Leon L, et al. Measures of reproducibility in sampling and laboratory processing methods in high-throughput microbiome analysis. bioRxiv 2018 May 15. [doi: 10.1101/322677]
23. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
24. Agnello M, Carroll LN, Imam N, Pino R, Palmer C, Varas I, et al. Gut microbiome composition and risk factors in a large cross-sectional IBS cohort. BMJ Open Gastroenterol 2020;7(1):e000345. [doi: 10.1136/bmjgast-2019-000345] [Medline: 32518661]
25. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 2013 Jan;41(Database issue):D590-D596 [FREE Full text] [doi: 10.1093/nar/gks1219] [Medline: 23193283]
26. McCoy CO, Matsen FA. Abundance-weighted phylogenetic diversity measures distinguish microbial community states and are robust to sampling depth. PeerJ 2013;1:e157 [FREE Full text] [doi: 10.7717/peerj.157] [Medline: 24058885]
27. Kannisto KA, Korhonen J, Adams CE, Koivunen MH, Vahlberg T, Välimäki MA. Factors associated with dropout during recruitment and follow-up periods of a mHealth-based randomized controlled trial for Mobile.Net to encourage treatment adherence for people with serious mental health problems. J Med Internet Res 2017 Feb 21;19(2):e46 [FREE Full text] [doi: 10.2196/jmir.6417] [Medline: 28223262]
28. Shehadeh N, Kligman AM. The bacteria responsible for axillary odor. J Invest Dermatol 1963;41:39-43. [doi: 10.1038/jid.1963.60] [Medline: 14043010]
29. Natsch A, Emter R. The specific biochemistry of human axilla odour formation viewed in an evolutionary context. Philos Trans R Soc Lond B Biol Sci 2020 Jun 08;375(1800):20190269 [FREE Full text] [doi: 10.1098/rstb.2019.0269] [Medline: 32306870]
30. Ara K, Hama M, Akiba S, Koike K, Okisaka K, Hagura T, et al. Foot odor due to microbial metabolism and its control. Can J Microbiol 2006 Apr;52(4):357-364. [doi: 10.1139/w05-130] [Medline: 16699586]
31. Callewaert C, Hutapea P, Van de Wiele T, Boon N. Deodorants and antiperspirants affect the axillary bacterial community. Arch Dermatol Res 2014 Oct;306(8):701-710. [doi: 10.1007/s00403-014-1487-1] [Medline: 25077920]
32. Mann EA, Bae E, Kostyuchek D, Chung HJ, McGee JS. The gut microbiome: human health and inflammatory skin diseases. Ann Dermatol 2020;32(4):265. [doi: 10.5021/ad.2020.32.4.265]
33. Hidalgo-Cantabrana C, Gómez J, Delgado S, Requena-López S, Queiro-Silva R, Margolles A, et al. Gut microbiota dysbiosis in a cohort of patients with psoriasis. Br J Dermatol 2019 Dec;181(6):1287-1295. [doi: 10.1111/bjd.17931] [Medline: 30920647]

34. Dei-Cas I, Giliberto F, Luce L, Dopazo H, Penas-Steinhardt A. Metagenomic analysis of gut microbiota in non-treated plaque psoriasis patients stratified by disease severity: development of a new Psoriasis-Microbiome Index. Sci Rep 2020 Jul 29;10(1):12754 [FREE Full text] [doi: 10.1038/s41598-020-69537-3] [Medline: 32728075]
35. Schmidt AC, Leroux JC. Treatments of trimethylaminuria: where we are and where we might be heading. Drug Discov Today 2020 Sep;25(9):1710-1717 [FREE Full text] [doi: 10.1016/j.drudis.2020.06.026] [Medline: 32615074]
36. Rath S, Heidrich B, Pieper DH, Vital M. Uncovering the trimethylamine-producing bacteria of the human gut microbiota. Microbiome 2017 Dec 15;5(1):54 [FREE Full text] [doi: 10.1186/s40168-017-0271-9] [Medline: 28506279]
37. Johnson KV. Gut microbiome composition and diversity are related to human personality traits. Hum Microbiome J 2020 Mar;15:100069. [doi: 10.1016/j.humic.2019.100069]
38. Ma J, Sheng L, Hong Y, Xi C, Gu Y, Zheng N, et al. Variations of gut microbiome profile under different storage conditions and preservation periods: a multi-dimensional evaluation. Front Microbiol 2020;11:972 [FREE Full text] [doi: 10.3389/fmicb.2020.00972] [Medline: 32536906]
39. Heapy C, Montgomery K, Ersser S, Gass M, Goad N, Thompson AR. A psychosocial support website from the British Association of Dermatologists for people living with a skin condition: mixed methods evaluation. JMIR Dermatol 2020 Jun 15;3(1):e17052. [doi: 10.2196/17052]
40. Cotterill JA. Dermatological non-disease: a common and potentially fatal disturbance of cutaneous body image. Br J Dermatol 1981 Jun;104(6):611-619. [doi: 10.1111/j.1365-2133.1981.tb00746.x] [Medline: 7248174]
41. Ågren M, Ghathian K, Frederiksen A, Bjerrum M, Calum H, Danielsen P, et al. Zinc oxide inhibits axillary colonization by members of the genus Corynebacterium and attenuates self-perceived malodour: a randomized, double-blind, placebo-controlled trial. Acta Derm Venereol 2020;100(10):a. [Medline: 32399578]
42. Lynch RJM. Zinc in the mouth, its interactions with dental enamel and possible effects on caries; a review of the literature. Int Dent J 2011 Aug;61 Suppl 3:46-54. [doi: 10.1111/j.1875-595X.2011.00049.x] [Medline: 21762155]
43. Dimitri-Pinheiro S, Soares R, Barata P. The microbiome of the nose-friend or foe? Allergy Rhinol (Providence) 2020;11:2152656720911605. [doi: 10.1177/2152656720911605] [Medline: 32206384]

## Abbreviations

**CLR:** centered log-ratio
**CONSORT-EHEALTH:** Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
**MEBO:** metabolic body odor
**OTU:** operational taxonomic unit
**PATM:** people are allergic to me
**QoL:** quality of life
**rRNA:** ribosomal RNA
**TMAU:** trimethylaminuria

*Edited by G Eysenbach; submitted 16.09.20; peer-reviewed by Z Lin, A Cheudjeu; comments to author 30.09.20; revised version received 05.10.20; accepted 24.10.20; published 04.11.20.*

*Please cite as:*
*Gabashvili IS*
*Cutaneous Bacteria in the Gut Microbiome as Biomarkers of Systemic Malodor and People Are Allergic to Me (PATM) Conditions: Insights From a Virtually Conducted Clinical Trial*
*JMIR Dermatol 2020;3(1):e10508*
*URL: http://derma.jmir.org/2020/1/e10508/*
*doi:10.2196/10508*
*PMID:*

Original Paper

# The Difficulty of German Information Booklets on Psoriasis and Psoriatic Arthritis: Automated Readability and Vocabulary Analysis

Martin Wiesner[1,2*], Dipl-Inform Med; Richard Zowalla[1,2,3], MSc; Monika Pobiruchin[2,4*], Dr sc hum

[1]Department of Medical Informatics, Heilbronn University, Heilbronn, Germany

[2]Consumer Health Informatics SIG, German Association for Medical Informatics, Biometry & Epidemiology (GMDS eV), Cologne, Germany

[3]Center for Machine Learning, Heilbronn University, Heilbronn, Germany

[4]GECKO Institute for Medicine, Informatics & Economics, Heilbronn University, Heilbronn, Germany

[*]these authors contributed equally

**Corresponding Author:**
Martin Wiesner, Dipl-Inform Med
Department of Medical Informatics
Heilbronn University
Max-Planck-Str 39
Heilbronn, 74081
Germany
Phone: 49 71315046947
Email: martin.wiesner@hs-heilbronn.de

## Abstract

**Background:** Information-seeking Psoriasis or Psoriatic Arthritis patients are confronted with numerous educational materials when looking through the internet. Literature suggests that only 17.0%-21.4% of (Psoriasis, Psoriatic Arthritis) patients have a good level of knowledge about psoriasis treatment and self-management. A study from 1994 found that English Psoriasis/Psoriatic Arthritis brochures required a reading level between grades 8-12 to be understandable, which was confirmed in a follow-up study 20 years later. As readability of written health-related text material should not exceed the sixth-grade level, Psoriasis/Psoriatic Arthritis material seems to be ill-suited to its target audience. However, no data is available on the readability levels of Psoriasis/Psoriatic Arthritis brochures for German-speaking patients, and both the volume and their scope are unclear.

**Objective:** This study aimed to analyze freely available educational materials for Psoriasis/Psoriatic Arthritis patients written in German, quantifying their difficulty by assessing both the readability and the vocabulary used in the collected brochures.

**Methods:** Data collection was conducted manually via an internet search engine for Psoriasis/Psoriatic Arthritis–specific material, published as PDF documents. Next, raw text was extracted, and a computer-based readability and vocabulary analysis was performed on each brochure. For the readability analysis, we applied the Flesch Reading Ease (FRE) metric adapted for the German language, and the fourth Vienna formula (WSTF). To assess the laymen-friendliness of the vocabulary, the computation of an expert level was conducted using a specifically trained Support Vector Machine classifier. A two-sided, two-sample Wilcoxon test was applied to test whether the difficulty of brochures of pair-wise topic groups was different from each other.

**Results:** In total, 68 brochures were included for readability assessment, of which 71% (48/68) were published by pharmaceutical companies, 22% (15/68) by nonprofit organizations, and 7% (5/68) by public institutions. The collection was separated into four topic groups: basic information on Psoriasis/Psoriatic Arthritis (G1/G2), lifestyle, and behavior with Psoriasis/Psoriatic Arthritis (G3/G4), medication and therapy guidance (G5), and other topics (G6). On average, readability levels were comparatively low, with FRE=31.58 and WSTF=11.84. However, two-thirds of the educational materials (69%; 47/68) achieved a vocabulary score ≤4 (ie, easy, very easy) and were, therefore, suitable for a lay audience. Statistically significant differences between brochure groups G1 and G3 for FRE (*P*=.0001), WSTF (*P*=.003), and vocabulary measure (L) (*P*=.01) exist, as do statistically significant differences for G2 and G4 in terms of FRE (*P*=.03), WSTF (*P*=.03) and L (*P*=.03).

**Conclusions:** Online Psoriasis/Psoriatic Arthritis patient education materials in German require, on average, a college or university education level. As a result, patients face barriers to understanding the available material, even though the vocabulary used seems appropriate. For this reason, publishers of Psoriasis/Psoriatic Arthritis brochures should carefully revise their educational materials to provide easier and more comprehensible information for patients with lower health literacy levels.

***(JMIR Dermatol 2020;3(1):e16095)*** doi:10.2196/16095

**KEYWORDS**

readability; psoriasis; psoriatic arthritis; health literacy; health education

## *Introduction*

### Overview

Psoriasis (International Classification of Diseases Tenth Edition [ICD-10] code: L40) is one of the most common chronic inflammatory skin disorders in the dermatology field, manifesting as scaly, erythematous plaques. According to Griffiths and Barker [1], “the incidence in white individuals is estimated to be 60 cases per 100 000 head of population per year.” Females and males are equally affected by the disease. Furthermore, this skin disease is associated with a form of inflammatory arthritis known as Psoriatic Arthritis (ICD-10: M07*) [2]. Patients’ health-related quality of life is reduced by both conditions by a considerable amount [3-6], and “is similar to that of other major medical diseases” [7].

The development of Psoriasis and its clinical expression is influenced by several external factors, including smoking, weight, and stressful life events [8]. Moreover, work productivity loss is reported for Psoriatic Arthritis patients with moderate to severe joint symptoms [6].

Self-management plays an important role in coping with the effects of Psoriasis. In this context, it is vital to follow a consistent therapy approach [9]. According to [10], the major reasons for missing treatment were “drinking alcohol, being fed up, forgetfulness, and being too busy.” However, patients require not only a certain degree of knowledge to keep their personal adherence level high, but psychological support [11] and exchange with other patients can also be valuable to improve self-management [12]. Besides consulting health professionals, Psoriasis patients can also seek (emotional) support and therapy advice from other sufferers, such as in online support communities [13]. Still, Renzi et al reported in a study with 240 Italian patients that [14]:

> *The level of knowledge about the disease was not as high, with only 17.0% and 21.4% of patient[s] with [Psoriasis] and [Psoriatic Arthritis], respectively, having a good level of knowledge about psoriasis treatment.*

Information-seeking Psoriasis/Psoriatic Arthritis patients are offered different forms of health education material, such as printed health booklets. In 1994, Feldman et al investigated the readability of such educational material when provided in English [15]. The authors found that the text material required a US education level between grades 8-12, which was above the recommended grade level of text material for health education [16-21]. However, these findings cannot be transferred improvidently to other languages, such as Italian or German, as education systems and language properties differ substantially.

Another major problem of written patient information is the gap between the language of experts and laypeople. Even with a higher level of education, medical vocabulary, such as concepts of diagnosis and treatment, pose problems for those affected by a disease [22]. Furthermore, the medical terms associated with the origin of a disease, as used by health professionals or patients, tend to be different ones [23-28].

To assess the difficulty of written text material, several metrics exist for the English language [29-33]. However, the manual computation of these metrics can be difficult and time-consuming for large document collections and is, therefore, associated with a high demand for human or financial resources. Given the great variety of available Psoriasis/Psoriatic Arthritis brochures on the internet, a manual or semiautomatic approach seems far from practical. In this context, to the best of the authors’ knowledge, no study has previously been published for Psoriasis/Psoriatic Arthritis–specific health education material written in the German language that applies machine learning methods and computes readability levels and vocabulary difficulty in a fully automated approach.

This study presents an automated, computer-based readability and vocabulary analysis of 68 patient information brochures on Psoriasis and Psoriatic Arthritis in German. The difficulty assessment of these brochures was conducted by applying a German adaptation of the Flesch-Reading Ease (FRE) [29] scale [34], the fourth Vienna formula (German: Wiener Sachtextformel, WSTF) [35], and a vocabulary-oriented method that is based on a Support Vector Machine (SVM) [36].

### Related Work

Written or oral patient information should provide scientific evidence on a disease in a way that patients can understand. Individuals must be able to assess the essential chances and risks inherent to available therapeutic strategies and to balance them with their situation in life. In this context, health literacy, according to Ratzen and Parker, describes [37]:

> *The degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions.*

This concept is particularly important as low health literacy is associated with a poorer general health status and increased mortality, especially for higher age groups [38].

To quantify the health literacy level of an individual, the European Health Literacy Survey offers an instrument with a scale ranging from 1 (lowest) to 50 (highest). It was used to compare health literacy levels in different European countries. An analysis by Zok reports an average score of 31.9 for German participants, which was below the European average score (33.8) [39]. In a study from 2016, Schaeffer et al reported that “54.3% of [German study participants] were found to have limited health literacy” (n=2000) [40,41]. These findings support the need for educational materials that meet the capabilities of their readers; that is, those materials must be written at a sufficient readability level. Consequently, expert-centric vocabulary should be avoided as it imposes barriers to patients, hinders understandability of recommended therapy advice, or might lower overall adherence to treatment plans.

In this context, the analysis of health education material plays an important role in text production or for the improvement of existing material. However, several studies found that health education material is often written and published with low readability, which reduces or hinders its understandability for its intended target readers [42-57]. Different medical subdisciplines or diseases have been the subject of readability assessments. These include, among others: (1) cancer; (2) heart diseases; (3) lung diseases; (4) kidney diseases; (5) ophthalmic conditions; or (6) dermatologic conditions. Many other medical subdisciplines have been assessed, and both the previous list and related literature references should not be considered complete. Instead, the selected studies highlight recent studies in the broad field of readability assessment.

In 2004, Friedman et al analyzed cancer education material from 55 websites [42]. They reported a mean FRE score of 41.6; that is, readability of the content presented was at college-level, which corresponds to a US school level of grade 13+. However, their analysis revealed differences between different types of cancer, as “breast cancer sites were written at easier reading grade levels.” A similar study was presented by Basch et al in 2018, where the readability of prostate cancer materials on the internet was assessed using five different metrics [43]. They reported that the “majority of websites had difficult readability” and concluded that a “large majority of information available on the Internet about prostate cancer will not be readable for many individuals.” A recent analysis of printed booklets addressing melanoma patients in the German language found that the median FRE was 43 for nine brochures analyzed manually [44]. The authors reported “low readability in at least half of the booklets” and emphasized the need for content and didactic revision of the educational material.

In 2012, Taylor-Clarke et al studied the suitability and readability of written material (n=18) provided in heart failure clinics and available on the internet [45]. In a non-computer-based analysis, the authors used the Fry readability formula and found that readability levels “ranged between 3rd and 15th grade-level,” and the average readability level was eighth grade level. Similar results were reported by Kher, Johnson, and Griffith [46] in their study, which included health education material on congestive heart failure from 70 websites. Their primary outcome was that “only 5 out of 70 websites were within the limits of the recommended sixth-grade readability level.” The mean FRE score was 48.87.

A recent study on heart failure education via a mobile app [47] analyzed the in-app content with an online readability calculation tool. The authors reported, “although the use of medical terminology in patient educational material is often unavoidable,” which results in many polysyllabic medical terms, the “CHF [congestive heart failure] Info App included fewer polysyllabic terms.” They calculated a mean of sixth grade reading level for the in-app CHF content.

Other studies investigated the readability of educational material provided for patients with lung diseases or their family members. A study from 2016 included 109 patient-directed online information resources and applied ten different readability metrics [48]. Weiss et al found that only “10 articles (9%) were written below a sixth-grade level,” but the “average [FRE] score was 52,” ranging from 18 to 78; the grade level ranged from “9.2 (www.cancer.gov) to 15.2 (www.wikipedia.org)” when grouped by parent website. A study by Hansberry et al [49], assessed the readability of educational material on the “health benefits of lung cancer screening,” which was intended for the general public, using ten readability instruments. The authors reported that of “80 articles, 62.5% required a high school education to comprehend.”

In a similar study, Haas et al reviewed 46 websites on lung cancer screening [50]. The overall mean Flesch-Kincaid grade level was a mean of 10.6 (SD 2.2). In 2017, Fullmann et al [51] assessed consumer information of 26 chronic obstructive pulmonary disease inhalers from the Health Canada Drug Product Database. They concluded that, while the medication information section was on average “difficult to read” or “hard” (FRE=47.8), the instruction section was “easy” or “fairly easy” (FRE=79.0) to read.

For the field of nephrology, Thomas et al [52] analyzed Wikipedia as a resource for patient education, including 69 publicly available articles. The overall mean FRE reported was 19.4, which corresponds to a deficient level of readability. Moreover, the mean Flesch-Kincaid grade level was 15.1, signaling college-level education was required by readers of Wikipedia. A systematic review by Morony et al [53] included 80 patient education materials on chronic kidney disease from the United States, the United Kingdom, and Australia. When evaluated with the Flesch-Kincaid grade level instrument, “most materials required a minimum of grade 9” reading level. The authors emphasized that “cognitive decline in patients” suffering from the effects of this disease resulted in “lower literacy than the average patient,” and content providers should carefully compile text material.

Online ophthalmic patient information was studied by Edmunds et al [54]. They assessed 160 websites, reporting a median FRE score of 52.1. Their analysis found that “83% [..] as being of ‘difficult’ readability.” The authors also reported that “Not-for-profit webpages were of significantly greater length than commercial webpages.” A single-institution study evaluated education materials on glaucoma [55]. The authors checked the handouts’ readability of their institution and found a 10th-grade Flesch-Kincaid reading level. After “applying guidelines on writing easy-to-understand” material and revising the material, readability had improved to “a 6th-grade reading level,” which better suits patients with low health literacy levels.

Tulbert, Snyder, and Brodell [56] compared the readability of “three sources of patient-education material on the internet (WebMD.com, Wikipedia.org, and MedicineOnline.com) […] with materials produced by the American Academy of Dermatology [AAD]”. The educational materials found on Wikipedia.org were more difficult to comprehend than AAD and MedicineOnline. Tulbert et al categorized the retrieved pamphlets by several topics. Psoriasis brochures (no differentiation between Psoriasis/Psoriatic Arthritis) were found with a mean FRE of 39.5 for the AAD materials, and a mean FRE of 53.6 for the WebMD resources.

The readability of education materials designed for patients with Psoriasis was studied in 1994 [15]. The authors found that the text material, written in English, required an education level between grades 8-12, significantly above the recommended grade level for health education. In their analysis, the mean FRE score was 52.7. A follow-up study was conducted 20 years later by Smith [57]. The analysis of these brochures in English revealed that revised, newer online resources on Psoriasis provided by three organizations still “fail to meet the desired 6th grade level” [57].

### Aims of the Study

The authors decided to focus on brochures available for free on the internet and written in German, targeting patients with Psoriasis (Vulgaris) or Psoriatic Arthritis. In this context, the aim of this study was three-fold: (1) to conduct an analysis of the current situation, that is, the volume and scope of information brochures on Psoriasis/Psoriatic Arthritis for (German-speaking) patients; (2) to quantify the level of readability of the text material and the type of vocabulary used in the identified brochures; and (3) to evaluate whether different types of brochures are better suited for citizens with lower health literacy levels. Therefore, this study can provide a baseline for researchers that want to validate their findings.

## Methods

### Study Design

This study of educational material consisted of two stages. First, to answer aim 1, data extraction was conducted manually using an internet search for PDF documents specifically written for and targeting Psoriasis and Psoriatic Arthritis patients. The retrieval was limited to PDF documents. This file type was chosen as the corresponding documents are easily accessible in electronic format (machine-readable) and can also be distributed in printed format. Generally, these documents are highly structured and proofread by publishing institutions.

Next, the subsequent stage used the health education material collected in stage 1 and conducted a computer-based readability and vocabulary analysis. Both analyses were intended to answer research aims 2 and 3.

### Study Setting

Patient information brochures on Psoriasis (Vulgaris) and Psoriatic Arthritis were collected. All booklets had to be freely available on the internet. Print-only booklets or multimedia content were not considered. Documents were eligible for inclusion if they: (1) provided information on Psoriasis and Psoriatic Arthritis for patients; (2) provided information in the German language; and (3) were free to access. If these criteria were not met, then the related documents were excluded from the readability and vocabulary analysis.

For the identification of relevant brochures, the expert term “Psoriasis” was chosen, accompanied by its more layman-friendly German term “Schuppenflechte.” The two terms refer to the same concept, and patients in Germany are familiar with both. The German term “Broschüre” (English: brochure) was included to find educational materials suited for patients rather than other types of PDF files, such as drug package inserts or electronic presentation slides by medical professionals. The DuckDuckGo search engine was utilized to search the Web with the following search terms: +Broschüre +Psoriasis filetype:pdf (search terms A), +Broschüre +Schuppenflechte filetype:pdf (search terms B), +Schuppenflechte filetype:pdf (search terms C), and +Psoriasis filetype:pdf (search terms D).

After the elimination of duplicates, two authors screened the titles and the content of the retrieved information brochures in a joint session to check whether the educational material targeted Psoriasis/Psoriatic Arthritis patients. Therefore, false-positive retrieval results were removed during this manual step.

### Readability Analysis

#### *Definition*

Readability [58] is a term to describe the properties of written text concerning the readers’ competence, motivation, and understanding of a document [59]. It depends on the complexity of a text’s structure, the sentence structure, and the vocabulary used.

#### *Flesch Reading Ease Scale*

A well-established readability scale for the English language is the Flesch Reading Ease metric [29]. The FRE measures the readability of a text via its average sentence length (ASL) and the average number of syllables per word (ASW). It relies on the fact that short words or sentences are usually easier to understand than longer ones. However, for this analysis, we applied the modified FRE for the German language by Toni Amstad [34]:

#### *Vienna Formula*

In contrast to the FRE, the Vienna formula (WSTF) is not an adapted version for the German language. Instead, it relies on work by Bamberger and Vanacek [35], who analyzed the bases of German text material and derived at least five versions of the Vienna formula for prose and nonfiction text. Typically, the fourth WSTF is used for text analyses. This metric also relies on the ASL and the proportion of (complex) words with three or more syllables (MS):

### Vocabulary Classification

For the German language, average words’ length or syllable counts are not a good indicator of whether a term/concept is laypeople compatible, which means it can be easily understood by people with an education level of grades 6-7. This is because German grammer allows the creation and use of many compound words (eg, “Hauterkrankung,” “Hautunverträglichkeit,” “Kontaktallergie”), which are, while lengthy, quite laymen friendly for an average patient. Several machine learning techniques can be leveraged to compensate for the limitations of established readability measures [36,60].

This is why we added the vocabulary-based SVM approach as an extra dimension of text analysis.

In previous work [36], a vocabulary-based computation of an “expert level” using a specially trained SVM for German was presented, which was applied to cancer information brochures [61] and is also applicable to Psoriasis information brochures. To use this pretrained classifier to quantify the vocabulary-based difficulty of medical text material, several preprocessing steps are necessary [62]. As a first step, each text is split into tokens (ie, single word fragments). Second, nonhuman readable markup (eg, XML tags), as well as stop words, are removed (eg, he/she/it). This is important as these kinds of tokens do not influence the difficulty of a text. Next, the remaining tokens are reduced to their stem forms (eg, surgeries becomes surger) to eliminate linguistic variations of the same basic concept. Finally, the text content of a document is transformed into its mathematical representation based on previously selected features, similarly to a study conducted by Keinki et al [63]. In this context, features represent characteristic terms from the medical domain and thereby influence the vocabulary-based difficulty of a text.

To quantify the degree of “expert-centricity” of the text material, the vocabulary measure (L) $\in [1,\ldots,10]$ is defined. It makes use of the SVM classifier above. In this context, higher values of L indicate an academic (medical) background knowledge or working experience in the medical domain is needed; a value of >7 corresponds to a very expert-centric text, a value of 5-6 to a difficult text, a value of 4-5 to a moderate text (laypeople with medical, educational background), a value of 3-4 to an easy text (intermediate level/junior high school), and a value of <3 to a very easy text (elementary level/elementary school).

### Difficulty

The aforementioned instruments make use of different scales to express difficulty, either in terms of readability or vocabulary. Therefore, it seems advisable to map these scales to independent classes that express the difficulty much more simply. The mapping used in this study is presented in Table 1.

**Table 1.** Mapping readability and vocabulary instrument scales to corresponding classes (labels). Adapted according to [61].

| Difficulty | FRE[a] $\in$ [0,100] | WSTF[b] $\in$ [4,15] | L $\in$ [1,10] | Class label |
|---|---|---|---|---|
| Very difficult to read | [0-29] | [14-15] | 9, 10 | VD[b] |
| Difficult to read | [30-49] | [12-14[ | 7, 8 | D[d] |
| Fairly difficult to read | [50-59] | [10-12[ | 6 | D |
| Average readability | [60-69] | [8-10[ | 5 | M[e] |
| Fairly easy to read | [70-79] | [7-8[ | 4 | E[f] |
| Easy to read | [80-89] | [5-7[ | 3 | E |
| Very easy to read | [90-100] | [4-5[ | 1, 2 | VE[g] |

[a] FRE: Flesch Reading Ease.

[b] WSTF: Fourth Vienna Formula (German: Wiener SachTextFormel).

[c] VD: very difficult.

[d] D: difficult.

[e] M: moderate.

[f] E: easy.

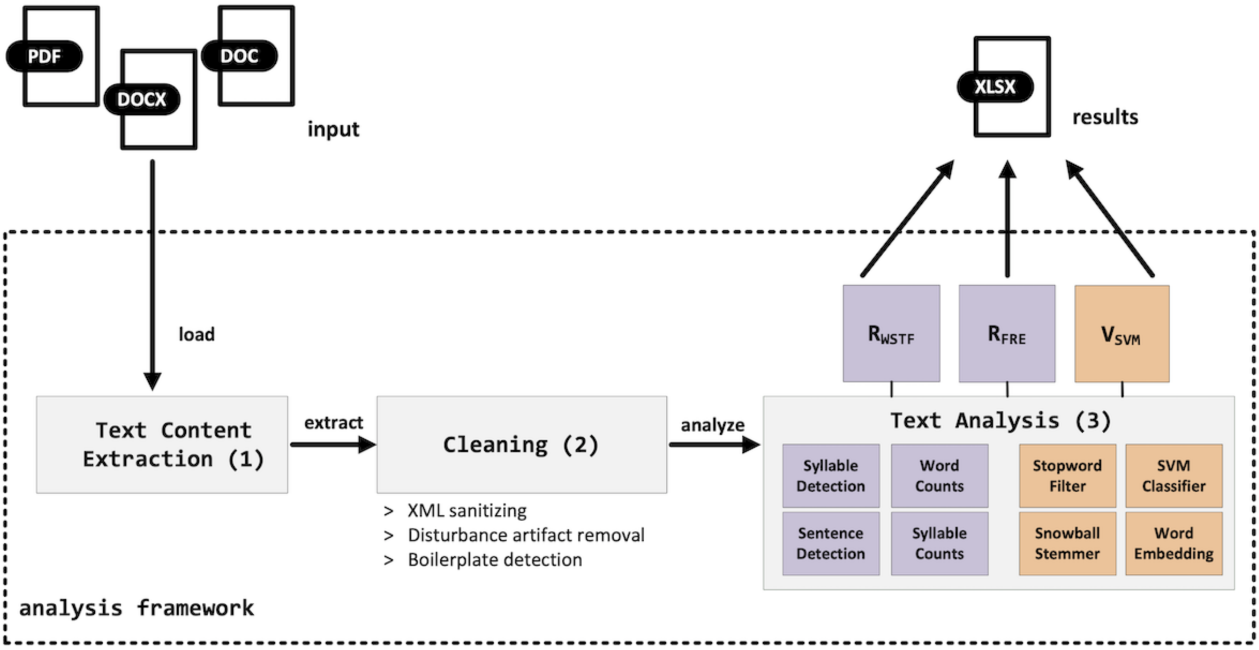
[g] VE: very easy.

### Computational Processing Steps

Parsing a text document is the process of analyzing its structure and fragments according to the rules of a natural language’s grammar. Typically, modern text documents (eg, PDF, DOC, DOCX) include metadata that describes their internal structure or external representation. In this context, text parsers process the descriptive markup structure of such document formats. The primary aim of this process is to extract the raw version of a text without any remaining technical markup which describes structural information. Typically, this includes how a paragraph is oriented, to which section it belongs, if text is formatted bold, if it contains figures or tables, and so on [64] (see chapters 5 and 6 for further details).

Before a parser can extract raw text data, the construction of a document collection is necessary. In the context of this study, all information brochures were downloaded as PDF files. These files were automatically converted to documents in DOCX format and represent the input of our analysis framework. The computational processing steps to compute readability and vocabulary scores for each document follows the workflow depicted in Figure 1.

**Figure 1.** Workflow of the processing steps and involved software components: (1) text content extraction; (2) a collection of data preparation and cleaning tasks; and (3) computation of the readability and vocabulary metrics. The analysis framework processes PDF, DOC, DOCX as input format and outputs a summary Excel spreadsheet for each document processed. SVM: support vector machine; FRE: Flesch Reading Ease; WSTF: Fourth Vienna Formula (German: Wiener SachTextFormel).



First, document parsers from the Apache Tika framework [65] were applied to extract the actual text content. As a second step, the extracted text was cleaned of disturbance artifacts (eg, different hyphen encoding schemes). Finally, the aforementioned readability and vocabulary metrics were computed for every brochure by a self-implemented analysis framework written in Java, which was previously tested against reference material. For sentence detection, the analysis framework relies on the Apache OpenNLP library [66] and their broadly accepted sentence model for the German language [67]. Liang's hyphenation algorithm [68] was used to estimate syllable counts. For stem form reduction, the Snowball Stemmer, according to Porter, was applied [69]. The analysis was conducted on a Mac OS 10.14.6 64bit computer running Java 11.0.4 (Oracle Corporation, Redwood Shores, California, United States) on August 21, 2019.

### Statistical Analysis

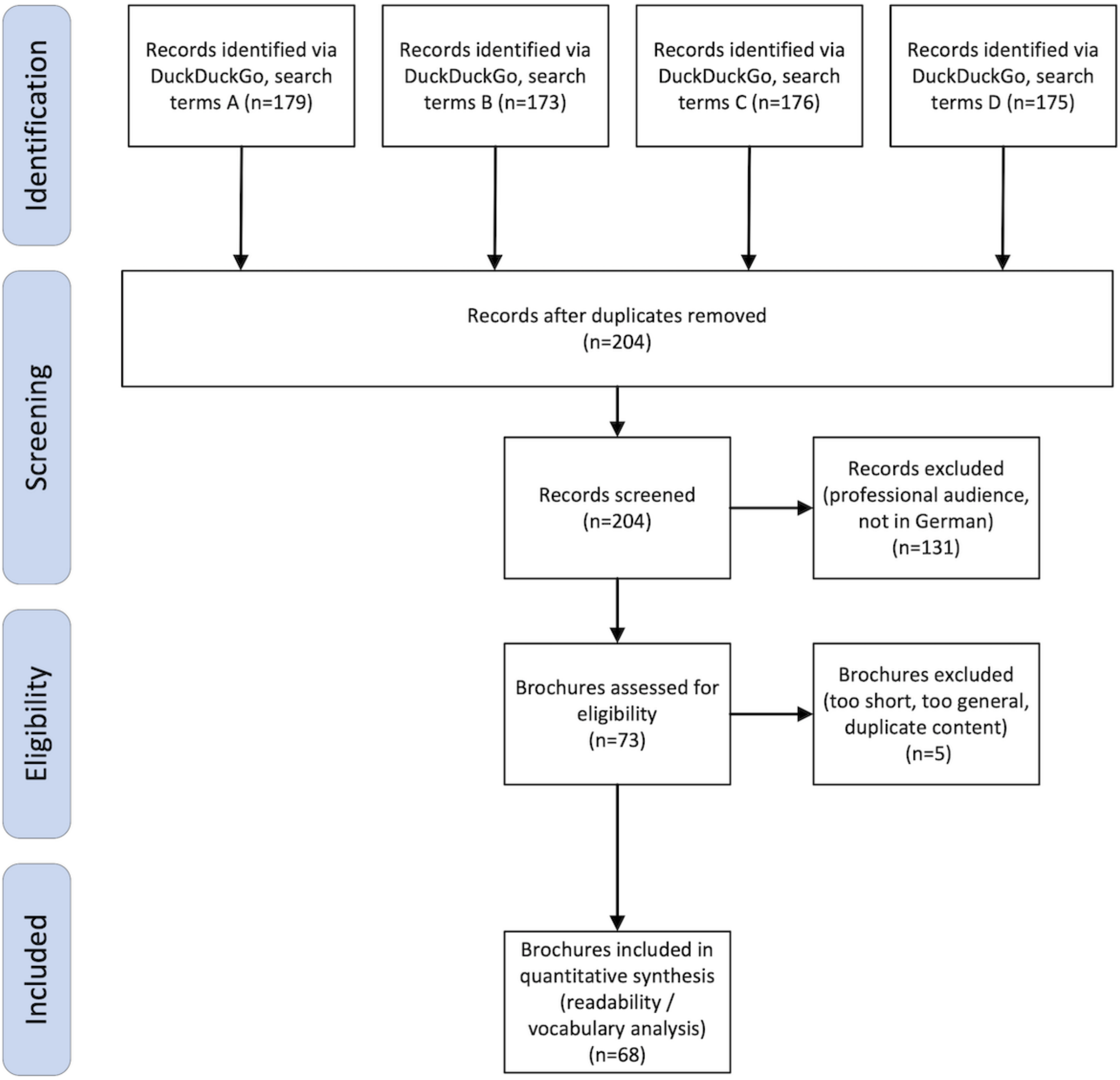
A two-sided, two-sample Wilcoxon test [70], also known as the Mann-Whitney *U* test, was applied to test whether the difficulty of brochures of two topic groups are different to each other ($H_0$: $\mu_1=\mu_2$, $H_1$: $\mu_1\neq\mu_2$, alpha=0.05). If $P<.05$, $H_1$ is accepted, as in there is a significant difference in terms of readability between two groups. The nonparametric *U* test was chosen as the number of brochures for several topic groups was rather small (n<10), and no normal distribution could be assumed. Data were analyzed with the statistics software R (The R Foundation, Vienna, Austria) version 3.6.1, on a Linux, Ubuntu 18.04 LTS/64bit computer.

## Results

### Principal Findings

The acquisition of Psoriasis/Psoriatic Arthritis brochures was carried out on August 19 and 20, 2019, by two of the authors. Given the search terms and the inclusion criteria, 73 brochures were eligible for inclusion, of which five were identified as either duplicate content or as being too general (ie, they were unspecific or covered other dermatology topics). The flowchart in Figure 2 depicts the data acquisition process with all details.

**Figure 2.** Data acquisition process with search terms A-D, as defined in the section "Study Setting".



In total, 68 brochures were included for further readability and vocabulary assessment. While assessing the brochures for eligibility, four categories emerged from the search engine's retrieval results: basic information on the disease (Psoriasis/Psoriatic Arthritis, labeled G1/G2), general advice on coping with Psoriasis/Psoriatic Arthritis in daily life situations (labeled G3/G4), including topics such as stress, diet, smoking, work-life and traveling, medication and therapy guidance (G5), and other topics (G6).

## Sample Characteristics

During the collection, several types of publishers emerged: pharmaceutical company or association, nonprofit organization, and public institution. Of the 68 brochures, 71% (48/68) were published by pharmaceutical companies or associations, 22% (15/68) by nonprofit organizations, and 7% (5/68) by public institutions. A detailed listing, given in Multimedia Appendix 1, includes the original German document title, publisher and type, and publishing year separated into G1-G6.

The included brochures were analyzed in terms of their linguistic characteristics. The number of sentences per brochure ranged from 45-619 (mean 235; SD 147.40) and the number of words from 579-11,430 (mean 3852; SD 2542.58). On average, 16.4 words were used by brochure authors to form a sentence (SD 3.03; minimum=11.5; maximum=27.7). Complex words, which meant ≥3 syllables, ranged from 253-4424 (mean 1284; SD 914.88). The minimal proportion of complex words was 22.85% (995/4354) and the maximum was at 46.9% (441/940), with a mean of 33.57% (1284/3852). A complete listing with data on the number of sentences, words, complex words, and syllables is given in Multimedia Appendix 2 per brochure and group (G1-G6).

## Readability Analysis

All brochure groups (G1-G6) were analyzed according to the readability metrics FRE and WSTF, as outlined in the Methods section. The results are presented in Table 2. The majority of the booklets are difficult (D) (FRE: 66%, 45/68; WSTF: 74%, 50/68), or very difficult (FRE: 34%, 23/68; WSTF: 13%, 9/68), to read.

In G1, the brochure with the lowest readability was PSO_110, with an FRE value of 19.26 and corresponding to the second highest WSTF value of 14.11 (VD). The corresponding Psoriatic Arthritis group G2 showed the lowest FRE value for PSO_210, with FRE=2.71 and WTSF=15 (VD). The third document set (G3) scored higher FRE values, thus signaling higher readability, which is supported by lower WSTF scores in this group. The corresponding Psoriatic Arthritis group (G4) produced similar results to G1. On average, documents about Psoriasis/Psoriatic Arthritis medication or therapy advice (G5) scored lowest, with PSO_502 being the most difficult one in this group (FRE=8.36; WTSF=15; VD). The lowest mean readability levels were $FRE_{G5}$=23.50 and $WSTF_{G5}$=12.95. The highest readability was achieved for G3, with an FRE of 41.39 and a WSTF of 10.27. For G6, no mean was calculated as the sample size was too small. Several selected text fragments with low or high readability levels can be found in Multimedia Appendix 3.

The distributions for both readability metrics, FRE, and the Vienna formula (WSTF), are depicted in Figures 3 and 4.

**Figure 3.** Distribution of achieved readability values on the Flesch Reading Ease scale. Difficulty is indicated by color, with dark green as the highest readability (90-100) and dark red as the lowest readability (0-10).

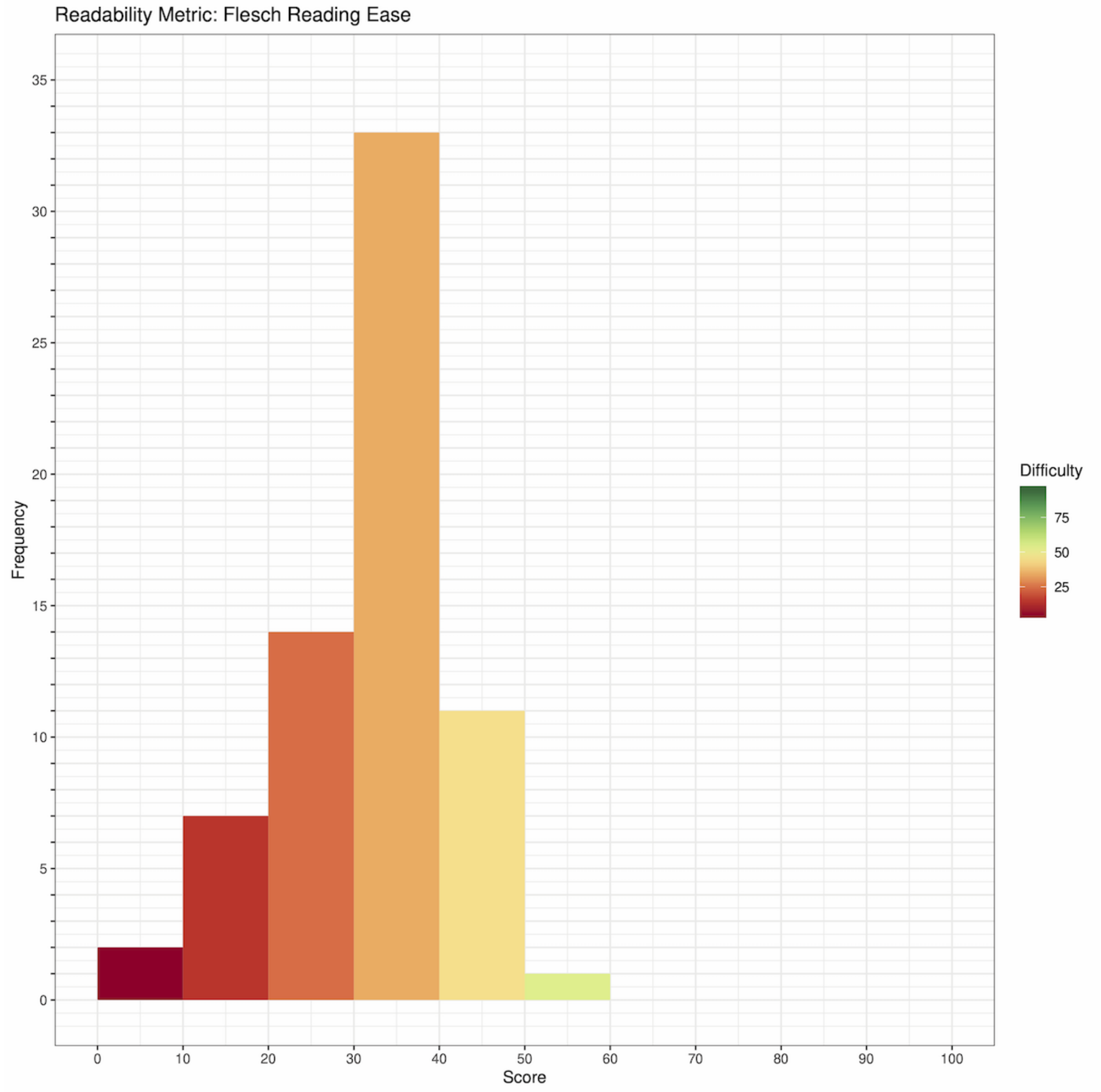


**Figure 4.** Distribution of achieved readability values on the Vienna formula scale. Difficulty is indicated by color, with dark green as the highest readability (4-5) and dark red as the lowest readability (14-15).

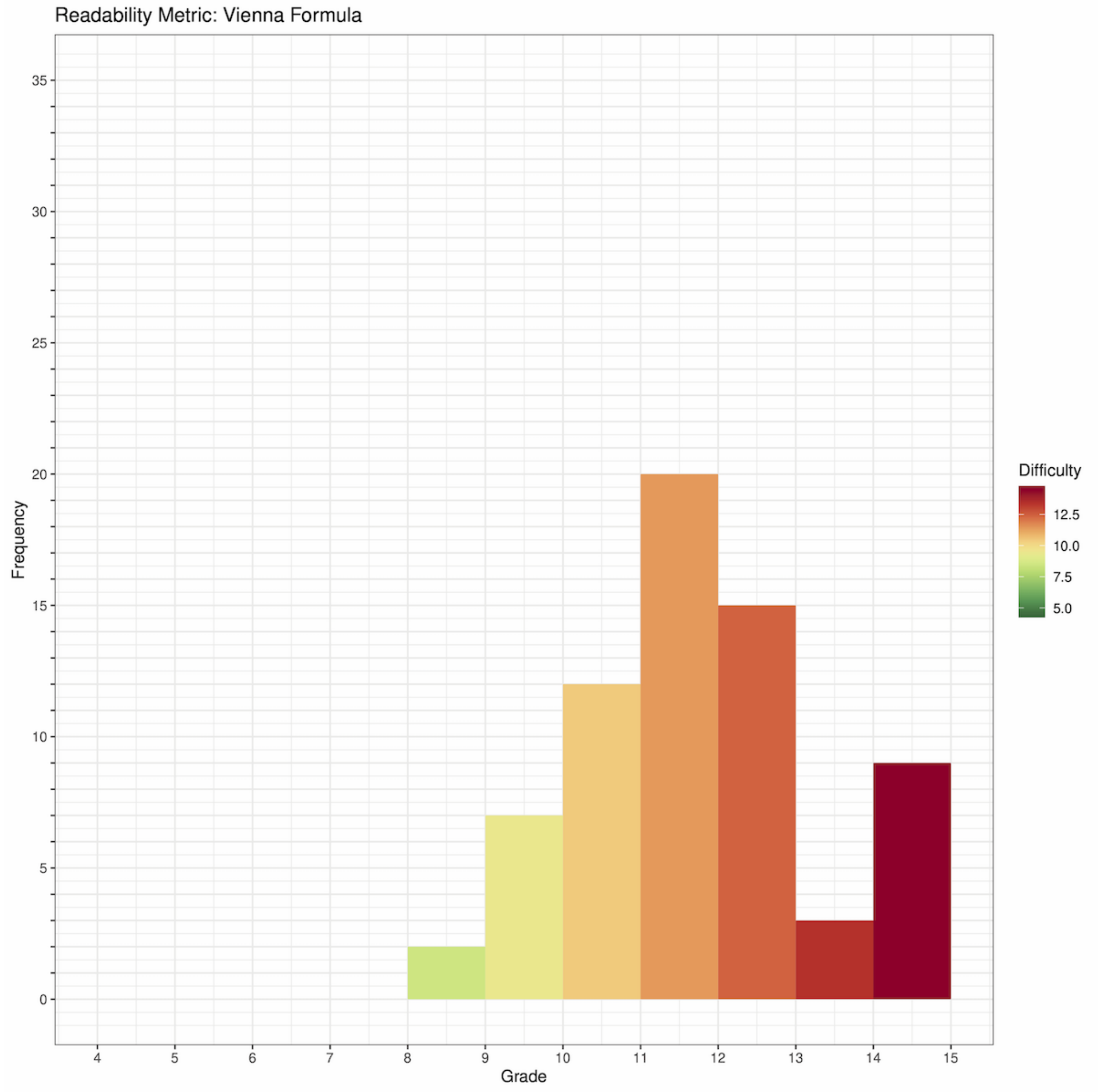


**Table 2.** Listing of readability and vocabulary scores, and associated class labels.

| Group and identifier | FRE[a] | WSTF[b] | L[c] | C[d]$_{FRE}$ | C$_{WSTF}$ | C$_{SVM}$[e] |
|---|---|---|---|---|---|---|
| **G1, Psoriasis, Basic Information (n=20)** | | | | | | |
| PSO_101 | 36.07 | 10.84 | 1 | D[f] | D | VE[g] |
| PSO_102 | 23.17 | 13.30 | 10 | VD[h] | D | VD |
| PSO_103 | 32.09 | 12.25 | 5 | D | D | M[i] |
| PSO_104 | 21.13 | 15.00 | 3 | VD | VD | E[j] |
| PSO_105 | 33.32 | 11.85 | 6 | D | D | D |
| PSO_106 | 38.94 | 10.67 | 3 | D | D | E |
| PSO_107 | 29.48 | 12.43 | 1 | VD | D | VE |
| PSO_108 | 43.55 | 9.89 | 1 | D | M | VE |
| PSO_109 | 31.74 | 11.90 | 1 | D | D | VE |
| PSO_110 | 19.26 | 14.11 | 10 | VD | VD | VD |
| PSO_111 | 29.83 | 12.39 | 6 | VD | D | D |
| PSO_112 | 36.66 | 11.65 | 1 | D | D | VE |
| PSO_113 | 30.57 | 12.07 | 1 | D | D | VE |
| PSO_114 | 30.91 | 12.21 | 8 | D | D | D |
| PSO_115 | 31.65 | 12.28 | 7 | D | D | D |
| PSO_116 | 27.63 | 12.27 | 10 | VD | D | VD |
| PSO_117 | 35.51 | 10.87 | 7 | D | D | D |
| PSO_118 | 37.08 | 11.23 | 3 | D | D | E |
| PSO_119 | 38.25 | 11.39 | 1 | D | D | VE |
| PSO_120 | 45.97 | 9.22 | 1 | D | M | VE |
| Mean | 32.64 | 11.89 | 4.30 | —[k] | — | — |
| **G2, Psoriatic Arthritis, Basic Information (n=15)** | | | | | | |
| PSO_201 | 33.69 | 10.90 | 1 | D | D | VE |
| PSO_202 | 36.95 | 11.68 | 1 | D | D | VE |
| PSO_203 | 30.92 | 11.93 | 3 | D | D | E |
| PSO_204 | 13.49 | 15.00 | 4 | VD | VD | E |
| PSO_205 | 36.40 | 11.07 | 2 | D | D | VE |
| PSO_206 | 30.69 | 11.73 | 1 | D | D | VE |
| PSO_207 | 27.03 | 12.31 | 2 | VD | D | VE |
| PSO_208 | 21.03 | 13.73 | 2 | VD | D | VE |
| PSO_209 | 14.41 | 14.93 | 7 | VD | VD | D |
| PSO_210 | 2.71 | 15.00 | 10 | VD | VD | VD |
| PSO_211 | 38.60 | 10.89 | 1 | D | D | VE |
| PSO_212 | 31.64 | 11.49 | 1 | D | D | VE |
| PSO_213 | 26.93 | 12.63 | 1 | VD | D | VE |
| PSO_214 | 32.08 | 11.90 | 1 | D | D | VE |
| PSO_215 | 23.47 | 12.71 | 9 | VD | D | VD |
| Mean | 26.67 | 12.53 | 3.07 | — | — | — |
| **G3, Psoriasis, Stress, Diet, Travelling, Smoking (n=12)** | | | | | | |
| PSO_301 | 40.16 | 10.18 | 1 | D | D | VE |

| Group and identifier | FRE[a] | WSTF[b] | L[c] | C[d]$_{FRE}$ | C$_{WSTF}$ | C$_{SVM}$[e] |
|---|---|---|---|---|---|---|
| PSO_302 | 39.43 | 10.18 | 1 | D | D | VE |
| PSO_303 | 33.78 | 11.63 | 1 | D | D | VE |
| PSO_304 | 31.22 | 12.81 | 1 | D | D | VE |
| PSO_305 | 38.66 | 10.95 | 1 | D | D | VE |
| PSO_306 | 46.92 | 9.43 | 1 | D | M | VE |
| PSO_307 | 48.48 | 9.55 | 1 | D | M | VE |
| PSO_308 | 53.86 | 8.07 | 1 | D | M | VE |
| PSO_309 | 47.84 | 9.07 | 1 | D | M | VE |
| PSO_310 | 38.03 | 10.64 | 1 | D | D | VE |
| PSO_311 | 32.89 | 11.75 | 10 | D | D | VD |
| PSO_312 | 45.37 | 8.98 | 1 | D | M | VE |
| Mean | 41.39 | 10.27 | 1.75 | — | — | — |
| **G4, Psoriatic Arthritis, Stress, Diet, Travelling, Smoking (n=8)** | | | | | | |
| PSO_401 | 32.90 | 11.12 | 4 | D | D | E |
| PSO_402 | 34.60 | 11.20 | 2 | D | D | VE |
| PSO_403 | 35.54 | 10.69 | 1 | D | D | VE |
| PSO_404 | 44.68 | 9.69 | 1 | D | M | VE |
| PSO_405 | 40.39 | 9.86 | 1 | D | M | VE |
| PSO_406 | 15.25 | 14.96 | 1 | VD | VD | VE |
| PSO_407 | 33.55 | 11.88 | 3 | D | D | E |
| PSO_408 | 40.32 | 10.14 | 1 | D | D | VE |
| Mean | 34.65 | 11.19 | 1.75 | — | — | — |
| **G5, Psoriasis and Psoriatic Arthritis, Medication, Therapy (n=11)** | | | | | | |
| PSO_501 | 24.82 | 11.56 | 3 | VD | D | E |
| PSO_502 | 8.36 | 15.00 | 9 | VD | VD | VD |
| PSO_503 | 14.29 | 15.00 | 10 | VD | VD | VD |
| PSO_504 | 33.88 | 11.68 | 1 | D | D | VE |
| PSO_505 | 10.73 | 14.14 | 10 | VD | VD | VD |
| PSO_506 | 32.74 | 11.02 | 7 | D | D | D |
| PSO_507 | 29.23 | 12.78 | 1 | VD | D | VE |
| PSO_508 | 29.05 | 12.75 | 10 | VD | D | VD |
| PSO_509 | 16.82 | 13.86 | 10 | VD | D | VD |
| PSO_510 | 33.50 | 11.78 | 6 | D | D | D |
| PSO_511 | 25.09 | 12.86 | 6 | VD | D | D |
| Mean | 23.50 | 12.95 | 6.64 | — | — | — |
| **G6, Psoriasis and Psoriatic Arthritis, Other Topics (n=2)** | | | | | | |
| PSO_601 | 29.18 | 12.33 | 2 | VD | D | VE |
| PSO_602 | 41.42 | 10.43 | 4 | D | D | E |

| Group and identifier | FRE[a] | WSTF[b] | L[c] | C[d]$_{FRE}$ | C$_{WSTF}$ | C$_{SVM}$[e] |
|---|---|---|---|---|---|---|
| Total Mean | 31.58 | 11.84 | 3.66 | — | — | — |

[a] FRE: Flesch Reading Ease.

[b] WSTF: Fourth Vienna Formula (German: Wiener SachTextFormel).

[c] L: vocabulary measure.

[d] C: class label.

[e] SVM: support vector machine.

[f] D: difficult.

[g] VE: very easy.

[h] VD: very difficult.

[i] M: moderate.

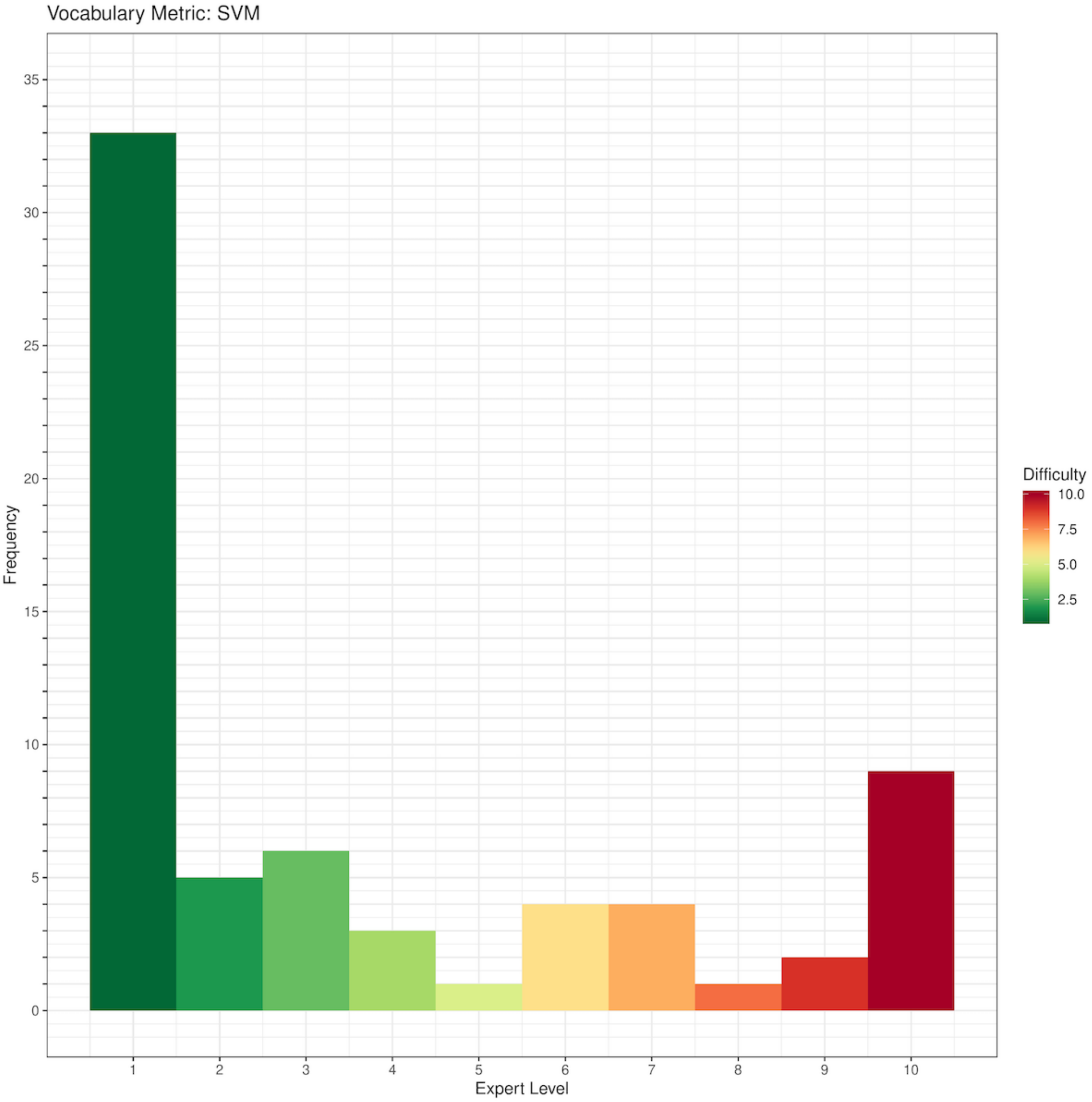
[j] E: easy.

[k] Not applicable.

## Vocabulary Classification

Overall, the brochures had a mean vocabulary measure (L) of L=3.66. As listed in Table 2, two-thirds of the educational materials (69%; 47/68) achieved a score ≤4 (VE+E) and were therefore suitable for a lay audience. A total of 11/68 booklets (16%) had a score ≥9 and are thus only suitable for an academic readership. For the remaining ten booklets (15%, 10/68), a score between >4 and <9 corresponds to a level suitable for persons with medical knowledge or a strong medical background. The groups G3 and G4 scored the lowest vocabulary measure, with L=1.75 for each. The highest vocabulary measure was found for the booklet group on medication and therapy topics (G5), with L=6.64. The distribution of the classification results over all the brochure groups is depicted in Figure 5.

**Figure 5.** Distribution of achieved vocabulary values on the SVM classification scale. Difficulty is indicated by color with dark green as the most laymen friendly (1) and dark red as the highest expert level required (10). SVM: support vector machine.



A comparison of the topic groups was conducted for the pairs G1/G3 and G2/G4. The results of the corresponding Wilcoxon test for two independent samples are presented in Table 3. Negative values originate from the definition of the FRE metric; that is, lower numbers correspond to a higher difficulty. In addition, due to a high number of ties (in the ranks) for the vocabulary metric (L), an exact computation of CI and *P* was not possible. Instead, a normal approximation was used by the statistics software R.

**Table 3.** Comparison of different brochure groups for difficulty.

| Comparison and Metric | Difference of Means | 95% CI | *P* value |
|---|---|---|---|
| **G1 versus G3** | | | |
| FRE[a] | –8.745 | –14.830 to –2.516 | .001 |
| WSTF[b] | 1.622 | 0.597-2.691 | .003 |
| L[c] | 2.55 | 0.00001-5.00 | .01 |
| **G2 versus G4** | | | |
| FRE | –7.985 | –14.513 to –1.256 | .03 |
| WSTF | 1.337 | 0.201-2.622 | .03 |
| L | 1.317 | –0.00004 to 2.00 | .03 |

[a] FRE: Flesch Reading Ease.

[b] WSTF: Fourth Vienna Formula (German: Wiener SachTextFormel). [c] L: vocabulary measure.

The observed differences between the brochure groups G1 and G3 (Psoriasis) for FRE (*P*=.001), WSTF (*P*=.003), and L (*P*=.01) were statistically significant, as were the FRE (*P*=.03), WSTF (*P*=.03), and L (*P*=.03) of G2 and G4 (Psoriatic Arthritis).

## Discussion

### Principal Results

High-quality health information must not only include the best available external evidence, it must also be readable and reflect patients' preferences [71]. In order to comply with these requirements, the application of easy language is essential [42-50,52-55,57,72].

The readability findings show that the majority of the collected material is difficult or very difficult (D+VD) to read, as shown by the WSTF (87%; 59/68). The outcome is more apparent when the German adaption of the FRE scale is applied (100%; 68/68) (Table 2). Thus, educational materials on Psoriasis/Psoriatic Arthritis are not suitable for their intended group of readers. This corresponds to the results of other authors, who also reported the high readability levels of such resources [73-77].

The vocabulary is also of great relevance for comprehensibility and might be even more decisive than the sentence structure [78]. The finding of the vocabulary analysis revealed that two-thirds (69%; 47/68) of the educational materials were well suited for laypeople. This originates from the fact that relatively few medical expert terms have been used during text production, or expert terminology has been actively avoided. With the difficulty assessment of 68 Psoriasis/Psoriatic Arthritis brochures, we demonstrated that a pretrained SVM can analyze text material for its vocabulary. The study findings therefore contribute the first dedicated vocabulary analysis related to the use of expert medical terms in patient educational material for Psoriasis and Psoriatic Arthritis.

### Limitations

Several limitations apply to the study setting. First, a public search engine was utilized to build the data collection used in this study. In this context, the internal mechanisms used to compute and retrieve information from a search engine's index are not fully transparent. For this reason, some potentially relevant documents might have been missed by our data collection process. The retrieval was also limited to PDF documents. The study design included this file type as the corresponding documents are easily accessible in electronic format (machine-readable), can also be distributed in printed format (these documents are, in general, highly structured and proof-read by publishing institutions), and represent a robust, well-known data format to provide information on (chronic) diseases and related treatment options via the internet.

Second, for this study, we analyzed 68 brochures on Psoriasis/Psoriatic Arthritis published by different types of organizations (see Multimedia Appendix 1). Depending on the motivation of an organization, there might be different aims in terms of content, words used, and selected topics. This might have affected our results, as scientific organizations might have used more complex sentence structures to explain Psoriasis/Psoriatic Arthritis concepts, while pharmaceutical companies might tend towards easier vocabulary and sentence structure.

Next, in the preprocessing phase, the included PDF brochures were automatically converted to documents in DOCX format. Nevertheless, disturbance artifacts, that is, different kinds of hyphens or misencoded characters originating from different encoding schemes, may still have been included in the extracted, raw text material.

The adapted FRE metric and Vienna formula are mainly computed on the basis of mean sentence length, the mean number of syllables per word, and language-specific weighting factors. However, detecting syllables is not a trivial task for the German language and does not work reliably in some rare circumstances [79]. For this reason, the computed FRE or WSTF scores can be influenced by the aforementioned inaccuracies. In this context, it should be stressed that this affects all natural language processing analysis tools for German text material.

Furthermore, solely computing the readability of educational materials disregards the individual knowledge and motivation of readers [35]. Aspects related to illustration and design were not included in the analysis of this study. Consequently, the suitability of health information cannot exclusively be judged based on its readability or its used vocabulary [35,80]. In this context, the studies by Taylor-Clarke et al and Tuot et al [45,81], among others, have applied methods that go beyond measures of word and sentence lengths, such as the Suitability Assessment of Materials (SAM) instrument, which reflects other aspects of a brochure's appearance that influence the understandability of (health) information and text comprehension.

However, besides the need for manual efforts, judging quality criteria is a highly subjective task for this instrument. Moreover, a sufficient number of judges are required to ensure an objective assessment of visual and aesthetic aspects in brochure design, which is not met by every study in this field. Even more important: interjudge reliability must be considered, evaluated, and reported properly. Modern approaches use crowd-sourcing techniques for which a large number of judges and related assessments can be obtained more easily [82].

## Comparison With Prior Work

Previous studies investigated the readability of health education materials on Psoriasis/Psoriatic Arthritis written in the English language [15,57]. In both analyses, the outcome was that the materials failed to “meet the desired 6th grade level” [57]. Although no accepted recommendation exists for German health education material, our findings confirm the low readability of Psoriasis/Psoriatic Arthritis brochures for patients. In contrast to the studies by Feldman et al and Smith, this study contributes the first vocabulary-related assessments of materials originating from the dermatology domain. We found that the vocabulary used in Psoriasis/Psoriatic Arthritis brochures is adequate for laypeople; that is, patients and family members who have no professional background in the health sector. A secondary study outcome gives a broad picture over the published materials in German-speaking countries, listed by publisher and year in Multimedia Appendix 1.

In a previous study [61], Keinki et al analyzed information booklets for German cancer patients. In this particular domain, the authors found a mean vocabulary score of $L_{SVM}$=5.09, signaling a higher difficulty for laypeople than in this study ($L_{SVM}$=3.66), that is, Psoriasis/Psoriatic Arthritis brochures make use of less complex medical terminology. This difference might be explained by the fact that Psoriasis/Psoriatic Arthritis brochures are mainly (71%; 48/68) produced and published by pharmaceutical companies or related associations. In contrast, cancer booklets follow a stricter evidence-based text production process in Germany [83], that is, patient guidelines and brochures on cancer topics are written or reviewed by medical professionals.

## Future Directions

This study analyzed static PDF document content for Psoriasis/Psoriatic Arthritis patients. In future work, the authors intend to extend their analyses to other types of online resources. This includes the content of trustworthy health information websites in German or articles in Wikipedia. Given such an analysis, a comparison to the work of Thomas et al [52] would be possible in terms of FRE and grade levels, as the authors reported even lower readability than in this study.

## Conclusions

For 68 German Psoriasis and Psoriatic Arthritis brochures freely available on the internet, the study findings reveal that the readability is low (Figures 3 and 4). Publishing organizations and authors should, therefore, reevaluate existing brochures and reduce sentence complexity, but our findings suggest that the use of vocabulary suits the target audience (Figure 5).

Methods from the field of machine learning can support authors of Psoriasis/Psoriatic Arthritis brochures, as they complement existing readability assessment methodology. For this reason, the assessment of written patient information should preferably be analyzed in terms of sentence structure and vocabulary, such as via the SVM-based classifier used for this study. The authors recommend the use of both sentence dimension and vocabulary dimension as supportive measures to ensure and provide understandable health education materials, independent of the medical domain.

---

## Acknowledgments

The authors thank Psoriasis Netz (psoriasis-netz.de) for hosting a well-maintained, curated collection of Psoriasis/Psoriatic Arthritis brochures on their website.

## Conflicts of Interest

None declared.

Multimedia Appendix 1

List of extracted Pso / PsA brochures. Publisher types: PC: Pharmaceutical Company or Association; NPO: Non-Profit Organization; PI: Public Institution.

[DOCX File , 34 KB - derma_v3i1e16095_app1.docx ]

Multimedia Appendix 2

Linguistic characteristics of analyzed Pso / PsA brochures. Se: Sentences, W: Words, CW: Complex Words, W / Se: Words per Sentence, CW / W: relative share of Complex Words (in per cent), Sy: Syllables, Ch: Characters.

[DOCX File , 38 KB - derma_v3i1e16095_app2.docx ]

Multimedia Appendix 3
Selection of Pso / PsA text fragments (DE/EN).
[DOCX File , 18 KB - derma_v3i1e16095_app3.docx ]

## References

1. Griffiths CE, Barker JN. Pathogenesis and clinical features of psoriasis. Lancet 2007 Jul 21;370(9583):263-271. [doi: 10.1016/S0140-6736(07)61128-3] [Medline: 17658397]
2. Gladman DD, Antoni C, Mease P, Clegg DO, Nash P. Psoriatic arthritis: epidemiology, clinical features, course, and outcome. Ann Rheum Dis 2005 Mar;64 Suppl 2:ii14-ii17 [FREE Full text] [doi: 10.1136/ard.2004.032482] [Medline: 15708927]
3. Krueger G, Koo J, Lebwohl M, Menter A, Stern RS, Rolstad T. The impact of psoriasis on quality of life: results of a 1998 National Psoriasis Foundation patient-membership survey. Arch Dermatol 2001 Mar;137(3):280-284. [Medline: 11255325]
4. Choi J, Koo JY. Quality of life issues in psoriasis. Journal of the American Academy of Dermatology 2003 Aug;49(2):57-61. [doi: 10.1016/s0190-9622(03)01136-8]
5. Mease PJ, Menter MA. Quality-of-life issues in psoriasis and psoriatic arthritis: outcome measures and therapies from a dermatological perspective. J Am Acad Dermatol 2006 Apr;54(4):685-704. [doi: 10.1016/j.jaad.2005.10.008] [Medline: 16546593]
6. Merola JF, Shrom D, Eaton J, Dworkin C, Krebsbach C, Shah-Manek B, et al. Patient Perspective on the Burden of Skin and Joint Symptoms of Psoriatic Arthritis: Results of a Multi-National Patient Survey. Rheumatol Ther 2019 Mar;6(1):33-45 [FREE Full text] [doi: 10.1007/s40744-018-0135-1] [Medline: 30610650]
7. Rapp SR, Feldman SR, Exum M, Fleischer AB, Reboussin DM. Psoriasis causes as much disability as other major medical diseases. Journal of the American Academy of Dermatology 1999 Sep;41(3):401-407. [doi: 10.1016/s0190-9622(99)70112-x]
8. Naldi L, Chatenoud L, Linder D, Belloni Fortina A, Peserico A, Virgili AR, et al. Cigarette Smoking, Body Mass Index, and Stressful Life Events as Risk Factors for Psoriasis: Results from an Italian Case–Control Study. Journal of Investigative Dermatology 2005 Jul;125(1):61-67. [doi: 10.1111/j.0022-202x.2005.23681.x]
9. Richards HL, Fortune DG, O'Sullivan TM, Main CJ, Griffiths CE. Patients with psoriasis and their compliance with medication. J Am Acad Dermatol 1999 Oct;41(4):581-583. [Medline: 10495380]
10. Zaghloul SS, Goodfield MJD. Objective assessment of compliance with psoriasis treatment. Arch Dermatol 2004 Apr;140(4):408-414. [doi: 10.1001/archderm.140.4.408] [Medline: 15096368]
11. Fortune DG, Richards HL, Griffiths CE. Psychologic factors in psoriasis: consequences, mechanisms, and interventions. Dermatol Clin 2005 Oct;23(4):681-694. [doi: 10.1016/j.det.2005.05.022] [Medline: 16112445]
12. Seng TK, Nee TS. Group therapy: a useful and supportive treatment for psoriasis patients. Int J Dermatol 1997 Feb;36(2):110-112. [doi: 10.1046/j.1365-4362.1997.00011.x] [Medline: 9109006]
13. Idriss SZ, Kvedar JC, Watson AJ. The role of online support communities: benefits of expanded social networks to patients with psoriasis. Arch Dermatol 2009 Jan 01;145(1):46-51. [doi: 10.1001/archdermatol.2008.529] [Medline: 19153342]
14. Renzi C, Di PC, Tabolli S. Participation, satisfaction and knowledge level of patients with cutaneous psoriasis or psoriatic arthritis. Clin Exp Dermatol 2011 Dec;36(8):885-888. [doi: 10.1111/j.1365-2230.2011.04126.x] [Medline: 21790723]
15. Feldman SR, Vanarthos J, Fleischer AB. The readability of patient education materials designed for patients with psoriasis. Journal of the American Academy of Dermatology 1994 Feb;30(2):284-286. [doi: 10.1016/s0190-9622(08)81932-9]
16. Doak LG, Doak CC. Patient comprehension profiles: recent findings and strategies. Patient Counselling and Health Education 1980 Jul;2(3):101-106. [doi: 10.1016/s0738-3991(80)80049-8]
17. Kirsch IS, Jungeblut A, Jenkins L, Kolstad A. National Center for Education Statistics. Washington, DC: U.S. Department of Education; 2002. Adult Literacy in America: A First Look at the Findings of the National Adult Literacy Survey URL: https://nces.ed.gov/pubs93/93275.pdf [accessed 2020-01-16]
18. Davis TC, Mayeaux EJ, Fredrickson D, Bocchini JA, Jackson RH, Murphy PW. Reading ability of parents compared with reading level of pediatric patient education materials. Pediatrics 1994 Mar;93(3):460-468. [Medline: 8115206]
19. Centers for Disease Control and Prevention. Atlanta, Georgia, United States: U.S. Department of Health and Human Services; 2009 Apr. Simply Put: A guide for creating easy-to-understand materials URL: https://www.cdc.gov/healthliteracy/pdf/Simply_Put.pdf [accessed 2020-01-16]
20. The Nation's Report Card. Washington, DC: U.S. Department of Education; 2012. NAEP - 2012 Long-term Trend: Summary of Major Findings URL: https://www.nationsreportcard.gov/ltt_2012/summary.aspx [accessed 2020-01-16]
21. U.S. National Library of Medicine. How to Write Easy-to-Read Health Materials. 2017. URL: https://medlineplus.gov/etr.html [accessed 2020-01-16] [WebCite Cache ID 6zBeCFhPU]
22. Ownby RL. Influence of vocabulary and sentence complexity and passive voice on the readability of consumer-oriented mental health information on the Internet. AMIA Annu Symp Proc 2005:585-589 [FREE Full text] [Medline: 16779107]
23. Bourhis RY, Roth S, MacQueen G. Communication in the hospital setting: a survey of medical and everyday language use amongst patients, nurses and doctors. Soc Sci Med 1989;28(4):339-346. [doi: 10.1016/0277-9536(89)90035-x] [Medline: 2705006]

24. Hume MA, Kennedy B, Asbury AJ. Patient knowledge of anaesthesia and peri-operative care. Anaesthesia 1994 Aug;49(8):715-718 [FREE Full text] [doi: 10.1111/j.1365-2044.1994.tb04408.x] [Medline: 7943707]
25. Chapple A, Campion P, May C. Clinical terminology: anxiety and confusion amongst families undergoing genetic counseling. Patient Educ Couns 1997;32(1-2):81-91. [doi: 10.1016/s0738-3991(97)00065-7] [Medline: 9355575]
26. Koch-Weser S, Dejong W, Rudd RE. Medical word use in clinical encounters. Health Expect 2009 Dec;12(4):371-382 [FREE Full text] [doi: 10.1111/j.1369-7625.2009.00555.x] [Medline: 19709316]
27. Wittenberg-Lyles E, Goldsmith J, Oliver DP, Demiris G, Kruse RL, Van Stee S. Using medical words with family caregivers. J Palliat Med 2013 Sep;16(9):1135-1139 [FREE Full text] [doi: 10.1089/jpm.2013.0041] [Medline: 23937064]
28. Wittenberg E, Goldsmith J, Ferrell B, Platt CS. Enhancing Communication Related to Symptom Management Through Plain Language. J Pain Symptom Manage 2015 Nov;50(5):707-711. [doi: 10.1016/j.jpainsymman.2015.06.007] [Medline: 26162506]
29. Flesch R. A new readability yardstick. J Appl Psychol 1948 Jun;32(3):221-233. [doi: 10.1037/h0057532] [Medline: 18867058]
30. Gunning R. The Technique of Clear Writing. New York City, New York, United States: McGraw Hill; Jun 01, 1968.
31. Mc Laughlin GH. SMOG Grading-a New Readability Formula. J Read 1969;12(8):639-646.
32. Coleman M, Liau TL. A computer readability formula designed for machine scoring. Journal of Applied Psychology 1975;60(2):283-284. [doi: 10.1037/h0076540] [Medline: 28425725]
33. Fry E. Fry's Readability Graph: Clarifications, Validity, and Extension to Level 17. J Read 1977;21(3):242-252.
34. Amstad T. Wie verständlich sind unsere Zeitungen? [How readable are our newspapers?]. Zurich, Switzerland: University of Zurich; 1978.
35. Bamberger R, Vanecek E. Lesen - Verstehen - Lernen - Schreiben. Die Schwierigkeitsstufen von Texten in deutscher Sprache [Reading - Understanding - Learning - Writing. The difficulty levels of German texts]. Vienna, Austria: Jugend u. Volk Sauerlaender; 1984.
36. Zowalla R, Wiesner M, Pfeifer D. Automatically Assessing the Expert Degree of Online Health Content using SVMs. Stud Health Technol Inform 2014;202:48-51. [Medline: 25000012]
37. Ratzan S, Parker R. Introduction. In: Seldon C, Zorn M, Ratzan S, Parker R. editors. Natl Libr Med Curr Bibliogr Med Health Lit 1st edition. Washington, DC: National Institutes of Health, US Department of Health and Human Services; 2000.
38. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011 Jul 19;155(2):97-107. [doi: 10.7326/0003-4819-155-2-201107190-00005] [Medline: 21768583]
39. Zok K. Unterschiede bei der Gesundheitskompetenz - Ergebnisse einer bundesweiten Repräsentativ-Umfrage unter gesetzlich Versicherten [Differences of Health Literacy - Results of a nation-wide Representative Survey among Statutory Health Insurees]. WIdO-monitor 2014;11(2):1-12 ISSN: 1614-8444 [FREE Full text]
40. Schaeffer D, Berens E, Vogt D. Health Literacy in the German Population: Results of a Representative Survey. Dtsch Arztebl Int 2017 Jan 27;114(4):53-60 [FREE Full text] [doi: 10.3238/arztebl.2017.0053] [Medline: 28211318]
41. Schaeffer D, Vogt D, Berens EM, Hurrelmann K. Gesundheitskompetenz der Bevölkerung in Deutschland: Ergebnisbericht [Health Literacy of the Population in Germany: Final Report]. Bielefeld, Germany: University of Bielefeld, Faculty of Health Sciences; 2017.
42. Friedman DB, Hoffman-Goetz L, Arocha JF. Readability of cancer information on the internet. J Cancer Educ 2004;19(2):117-122. [doi: 10.1207/s15430154jce1902_13] [Medline: 15456669]
43. Basch CH, Ethan D, MacLean SA, Fera J, Garcia P, Basch CE. Readability of Prostate Cancer Information Online: A Cross-Sectional Study. Am J Mens Health 2018 Sep 09;12(5):1665-1669 [FREE Full text] [doi: 10.1177/1557988318780864] [Medline: 29888641]
44. Brütting J, Reinhardt L, Bergmann M, Schadendorf D, Weber C, Tilgen W, NVKH. Quality, Readability, and Understandability of German Booklets Addressing Melanoma Patients. J Cancer Educ 2019 Aug 7;34(4):760-767. [doi: 10.1007/s13187-018-1369-x] [Medline: 29736794]
45. Taylor-Clarke K, Henry-Okafor Q, Murphy C, Keyes M, Rothman R, Churchwell A, et al. Assessment of commonly available education materials in heart failure clinics. J Cardiovasc Nurs 2012;27(6):485-494 [FREE Full text] [doi: 10.1097/JCN.0b013e318220720c] [Medline: 21743339]
46. Kher A, Johnson S, Griffith R. Readability Assessment of Online Patient Education Material on Congestive Heart Failure. Adv Prev Med 2017;2017:9780317. [doi: 10.1155/2017/9780317] [Medline: 28656111]
47. Athilingam P, Jenkins B, Redding BA. Reading Level and Suitability of Congestive Heart Failure (CHF) Education in a Mobile App (CHF Info App): Descriptive Design Study. JMIR Aging 2019 Apr 25;2(1):e12134 [FREE Full text] [doi: 10.2196/12134] [Medline: 31518265]
48. Weiss KD, Vargas CR, Ho OA, Chuang DJ, Weiss J, Lee BT. Readability analysis of online resources related to lung cancer. J Surg Res 2016 Nov;206(1):90-97. [doi: 10.1016/j.jss.2016.07.018] [Medline: 27916381]
49. Hansberry DR, White MD, D'Angelo M, Prabhu AV, Kamel S, Lakhani P, et al. Lung Cancer Screening Guidelines: How Readable Are Internet-Based Patient Education Resources? AJR Am J Roentgenol 2018 Jul;211(1):W42-W46. [doi: 10.2214/AJR.17.19042] [Medline: 29708784]

50. Haas K, Brillante C, Sharp L, Elzokaky AK, Pasquinelli M, Feldman L, et al. Lung cancer screening: assessment of health literacy and readability of online educational resources. BMC Public Health 2018 Dec 07;18(1):1356 [FREE Full text] [doi: 10.1186/s12889-018-6278-8] [Medline: 30526544]
51. Fullmann K, Blackburn DF, Fenton ME, Mansell H. Readability and Suitability of COPD Consumer Information. Can Respir J 2017;2017:2945282-2945288 [FREE Full text] [doi: 10.1155/2017/2945282] [Medline: 28947874]
52. Thomas GR, Eng L, de Wolff JF, Grover SC. An evaluation of Wikipedia as a resource for patient education in nephrology. Semin Dial 2013;26(2):159-163. [doi: 10.1111/sdi.12059] [Medline: 23432369]
53. Morony S, Flynn M, McCaffery KJ, Jansen J, Webster AC. Readability of Written Materials for CKD Patients: A Systematic Review. Am J Kidney Dis 2015 Jun;65(6):842-850. [doi: 10.1053/j.ajkd.2014.11.025] [Medline: 25661679]
54. Edmunds MR, Barry RJ, Denniston AK. Readability assessment of online ophthalmic patient information. JAMA Ophthalmol 2013 Dec;131(12):1610-1616. [doi: 10.1001/jamaophthalmol.2013.5521] [Medline: 24178035]
55. Williams AM, Muir KW, Rosdahl JA. Readability of patient education materials in ophthalmology: a single-institution study and systematic review. BMC Ophthalmol 2016 Aug 03;16:133 [FREE Full text] [doi: 10.1186/s12886-016-0315-0] [Medline: 27487960]
56. Tulbert BH, Snyder CW, Brodell RT. Readability of Patient-oriented Online Dermatology Resources. J Clin Aesthet Dermatol 2011 Mar;4(3):27-33 [FREE Full text] [Medline: 21464884]
57. Smith GP. The readability of patient education materials designed for patients with psoriasis: what have we learned in 20 years? J Am Acad Dermatol 2015 Apr;72(4):737-738. [doi: 10.1016/j.jaad.2014.12.018] [Medline: 25773418]
58. Klare GR. Assessing Readability. Reading Research Quarterly 1974;10(1):62. [doi: 10.2307/747086]
59. Klare G. The formative years. In: Zakaluk BL, Samuels SJ, editors. Readability: Its past, present, and future. Newark, Delaware, United States: International Reading Association; 1988:14-34.
60. Leroy G, Miller T, Rosemblat G, Browne A. A balanced approach to health information evaluation: A vocabulary-based naïve Bayes classifier and readability formulas. J. Am. Soc. Inf. Sci 2008 Jul;59(9):1409-1419. [doi: 10.1002/asi.20837]
61. Keinki C, Zowalla R, Pobiruchin M, Huebner J, Wiesner M. Computer-Based Readability Testing of Information Booklets for German Cancer Patients. J Cancer Educ 2019 Aug 12;34(4):696-704. [doi: 10.1007/s13187-018-1358-0] [Medline: 29651761]
62. Joachims T. Text categorization with Support Vector Machines: Learning with many relevant features. In: Nédellec C, Rouveirol C, editors. Machine Learning: ECML-98. Berlin, Heidelberg, Germany: Springer; 1998:137-142.
63. Keinki C, Zowalla R, Wiesner M, Koester MJ, Huebner J. Understandability of Patient Information Booklets for Patients with Cancer. J Cancer Educ 2018 Jun 10;33(3):517-527. [doi: 10.1007/s13187-016-1121-3] [Medline: 27726109]
64. Mattmann CA, Zitting JL. Tika In Action. Shelter Island, New York, United States: Manning Publications; 2020.
65. Apache Software Foundation. Apache Tika. 2019. URL: https://tika.apache.org/ [accessed 2020-01-16]
66. Apache Software Foundation. Apache OpenNLP. 2019. URL: https://opennlp.apache.org/ [accessed 2020-01-16]
67. Apache Software Foundation. OpenNLP. 2019. URL: http://opennlp.sourceforge.net/models-1.5/ [accessed 2020-01-16]
68. Liang F. Word Hy-phen-a-tion by Com-put-er. Stanford, California, United States: Stanford University; 1983.
69. Porter M. An algorithm for suffix stripping. Program 1980 Mar;14(3):130-137. [doi: 10.1108/eb046814]
70. Hollander M, Wolfe D, Chicken E. Nonparametric statistical methods. Third edition. Hoboken, New Jersey, United States: John Wiley and Sons, Inc; 2014.
71. Hoefert HW, Klotter C. In: Hoefert HW, Flick U, Härter M, editors. Wandel der Patientenrolle. Neue Interaktionsformen im Gesundheitswesen [Change of the Patient role. New Forms of Interaction in the Healthcare System]. Göttingen, Germany: Hogrefe Verlag; 2011.
72. Fagerlin A, Zikmund-Fisher BJ, Ubel PA. Helping patients decide: ten steps to better risk communication. J Natl Cancer Inst 2011 Oct 05;103(19):1436-1443 [FREE Full text] [doi: 10.1093/jnci/djr318] [Medline: 21931068]
73. Cooley ME, Moriarty H, Berger MS, Selm-Orr D, Coyle B, Short T. Patient literacy and the readability of written cancer educational materials. Oncol Nurs Forum 1995 Oct;22(9):1345-1351. [Medline: 8539175]
74. Garcia SF, Hahn EA, Jacobs EA. Addressing low literacy and health literacy in clinical oncology practice. J Support Oncol 2010;8(2):64-69 [FREE Full text] [Medline: 20464884]
75. Nicholls S, Hankins M, Hooley C, Smith H. A survey of the quality and accuracy of information leaflets about skin cancer and sun-protective behaviour available from UK general practices and community pharmacies. J Eur Acad Dermatol Venereol 2009 May;23(5):566-569. [doi: 10.1111/j.1468-3083.2008.03017.x] [Medline: 19175488]
76. Singh J. Reading Grade Level and Readability of Printed Cancer Education Materials. Oncology Nursing Forum 2007 Feb 8;30(5):867-870. [doi: 10.1188/03.onf.867-870]
77. Weintraub D, Maliski SL, Fink A, Choe S, Litwin MS. Suitability of prostate cancer education materials: applying a standardized assessment tool to currently available materials. Patient Educ Couns 2004 Nov;55(2):275-280. [doi: 10.1016/j.pec.2003.10.003] [Medline: 15530765]
78. Hasan M, Kotov A, Carcone A, Dong M, Naar S, Hartlieb KB. A study of the effectiveness of machine learning methods for classification of clinical interview fragments into a large number of categories. J Biomed Inform 2016 Aug;62:21-31 [FREE Full text] [doi: 10.1016/j.jbi.2016.05.004] [Medline: 27185608]

79. Müller K. Automatic Detection of Syllable Boundaries Combining the Advantages of Treebank and Bracketed Corpora Training. In: Proc 39th Annu Meet Assoc Comput Linguist. Stroudsburg, Pennsylvania: Association for Computational Linguistics; 2001 Presented at: Annual Meeting on Association for Computational Linguistics; July 6-11; Toulouse, France p. 410-417. [doi: 10.3115/1073012.1073065]
80. Friedman DB, Hoffman-Goetz L. A systematic review of readability and comprehension instruments used for print and web-based cancer information. Health Educ Behav 2006 Jun 30;33(3):352-373. [doi: 10.1177/1090198105277329] [Medline: 16699125]
81. Tuot DS, Davis E, Velasquez A, Banerjee T, Powe NR. Assessment of printed patient-educational materials for chronic kidney disease. Am J Nephrol 2013;38(3):184-194 [FREE Full text] [doi: 10.1159/000354314] [Medline: 23970127]
82. Carvalho A, Dimitrov S, Larson K. How many crowdsourced workers should a requester hire? Ann Math Artif Intell 2016 Jan 6;78(1):45-72. [doi: 10.1007/s10472-015-9492-4]
83. Schaefer C, Zowalla R, Wiesner M, Siegert S, Bothe L, Follmann M. Patientenleitlinien in der Onkologie: Zielsetzung, Vorgehen und erste Erfahrungen mit dem Format [Patient guidelines in oncology: objectives, procedures and first experiences with this format]. Z Evid Fortbild Qual Gesundhwes 2015 Jan;109(6):445-451. [doi: 10.1016/j.zefq.2015.09.013] [Medline: 26474649]

---

## Abbreviations

**AAD:** American Academy of Dermatology
**ASL:** Average Sentence Length
**ASW:** Average Number of Syllables per Word
**CHF:** Congestive Heart Failure
**FRE:** Flesch Reading Ease
**ICD-10:** International Classification of Diseases Tenth Edition
**L:** vocabulary measure
**MS:** Words with three or More Syllables
**SAM:** Suitability Assessment of Materials
**SVM:** support vector machine
**WSTF:** Fourth Vienna Formula (German: Wiener SachTextFormel)

---

*Edited by G Eysenbach; submitted 02.09.19; peer-reviewed by J Papadakos, J Lander; comments to author 09.11.19; revised version received 17.12.19; accepted 19.12.19; published 11.03.20.*

*Please cite as:*
*Wiesner M, Zowalla R, Pobiruchin M*
*The Difficulty of German Information Booklets on Psoriasis and Psoriatic Arthritis: Automated Readability and Vocabulary Analysis*
*JMIR Dermatol 2020;3(1):e16095*
*URL: http://derma.jmir.org/2020/1/e16095/*
*doi:10.2196/16095*
*PMID:*

Original Paper

# Assessing YouTube as an Educational Tool for Shingles: Cross-Sectional Study

Teevit Dunnsiri[1], BS; Takumi Kawashita[1], BS; Sharon C Lee[1], BS; Aaron Kumar Monga[1], BS; Benjamin K P Woo[2], MD

[1]Western University of Health Sciences, Pomona, CA, United States
[2]Olive View-UCLA Medical Center, Sylmar, CA, United States

**Corresponding Author:**
Teevit Dunnsiri, BS
Western University of Health Sciences
309 E 2nd St
Pomona, CA
United States
Phone: 1 909 623 6116
Email: teevit@g.ucla.edu

## Abstract

**Background:** YouTube is a popular platform with many videos, which have potential educational value for medical students. Due to the lack of peer review, other surrogates are necessary to determine the content quality of such educational videos. Few studies have analyzed the research background or academic affiliation of the physicians associated with the production of YouTube videos for medical education. The research background or academic affiliations of those physicians may be a reflection of the content quality of these educational videos.

**Objective:** This study identifies physicians associated with the production of educational YouTube videos about shingles and analyzes those physicians based on their research background or academic affiliation, which may be good surrogates for video content quality.

**Methods:** Using the YouTube search engine with default settings, the term “shingles” was searched on May 8, 2020. A cross-sectional study was performed using the first 50 search results. A search on Scopus for each identified physician was performed, and data regarding their research background and academic affiliation were recorded.

**Results:** Of the 50 YouTube videos, 35 (70%) were categorized as academic. Of the 35 academic videos, 24 (71%) videos featured physicians, totaling 25 physicians overall. Out of these 25 physicians, 5 (20%) had at least 1 shingles-related publication and 8 (32%) had an h-index >10. A total of 21 (84%) physicians held an academic affiliation.

**Conclusions:** These results ensure to a certain degree the quality of the content in academic videos on YouTube for medical education. However, further evaluation is needed for this growing platform.

***(JMIR Dermatol 2020;3(1):e20338)*** doi:10.2196/20338

**KEYWORDS**

YouTube; medical education; online; shingles; herpes zoster; peer review; dermatology; research background; content quality; views

## Introduction

Shingles, also referred to as herpes zoster, classically presents as a painful vesicular rash distributed along a dermatomal pattern. Symptoms typically last for 2-4 weeks. However, pain can persist even after the rash resolves in a complication known as postherpetic neuralgia. Shingles is caused by reactivation of the varicella-zoster virus, which lies dormant in nerve tissue after resolution of the initial infection known as chickenpox. Reactivation is typically due to stress or a weakened immunity. Approximately 1 million cases of shingles occur each year in the United States [1]. Despite the introduction of a shingles vaccine in 2006, recent studies have found no significant decline in the incidence of cases [2].

As online media platforms become more pervasive in our society, medical students are finding different methods to incorporate these resources into their education [3]. YouTube is a popular platform with many videos, which have potential educational value [4-6]. However, a study on YouTube videos as an educational resource on psoriasis found that misleading videos had a similar number of views to high-quality videos [7]. Although other studies have supported the educational value of YouTube videos, they have also acknowledged the need for more evaluation, particularly via peer-review [8,9].

A different study on psoriasis content on YouTube found that only 7.1% of the reviewed videos were associated with a medical institution or physician. The study concluded that this emphasized the need for more medical professionals to participate on YouTube in order to raise the quality of information [10]. However, few studies have thoroughly analyzed the physicians associated with the production of YouTube videos for dermatological medical education. Due to the lack of peer-review for YouTube videos, some studies have postulated that the research background of physicians featured in these videos for medical education could be a surrogate for content quality [9]. This study identifies physicians associated with the production of academic YouTube videos about shingles and analyzes their research background and academic affiliation, which may be good surrogates for video content quality.

## Methods

Using the YouTube search engine with default settings, the term “shingles” was searched on May 8, 2020. A cross-sectional study was conducted using the first 50 search results, which were sorted by relevance. Recorded data on the YouTube videos included the category (academic, video blog or “vlog”, patient interview, or miscellaneous), number of views, type of health professional associated with the video, and the name of the physician if available. Videos categorized as miscellaneous included commercials and local news reports. An academic video was defined as a video having scientifically correct information about the causes, symptoms, treatment, or prevention of the disease. The accuracy of information from a video was verified by medical students using UpToDate, an evidence-based medical resource. If the presence of a health professional was not clearly discernible in the video, the YouTube channel descriptions of the video producers were investigated.

Scopus is a large abstract and citation database for research. A search on each identified physician was performed on Scopus. Data was recorded from Scopus regarding the searched physicians’ area of specialty, h-index, number of publications, academic affiliations, and number of shingles-related publications. The h-index is a measure of the research output and citation impact of an author [11]. The h-index was categorized as either 0-10 or >10.

Physician variables such as having a shingles-related publication, h-index >10, or academic affiliation were recorded. Difference in the average number of views between videos associated with a physician who had a shingles-related publication, an h-index >10, or an academic affiliation was analyzed. Welch's unpaired *t* test was utilized due to unequal sample sizes between different groups.

## Results

From the first 50 “shingles” search results on YouTube, there were 35 academic videos, 6 patient interviews, 3 vlogs, and 6 miscellaneous videos. From the 35 academic videos, there were 25 physicians, 1 nurse practitioner, 1 pharmacist, 1 dentist, and 1 medical student identified. A total of 7 academic videos were produced by educational media outlets without an identified health professional. A physician was featured in 2 different academic videos, and 2 academic videos had more than one physician featured. From the 25 physicians identified from academic videos, there were 4 internists, 4 ophthalmologists, 4 family physicians, 3 infectious disease physicians, 3 anesthesiologists, 2 dermatologists, 2 obstetrician-gynecologists, 1 plastic surgeon, 1 urologist, and 1 rheumatologist.

The average number of publications for these 25 physicians was 62 (SD 130). The number of publications ranged from 0 to 529. Among these 25 physicians, 5(20%) had at least 1 shingles-related publication and 8 (32%) had an h-index >10. The h-index had an average of 13 (SD 22) and a range of 0 to 86. A total of 84% (21/25) of the physicians had an academic affiliation (Table 1).

**Table 1.** Physicians sorted by qualities.

| Physician quality | n (%), N=25 |
|---|---|
| Shingles-related publication | 5 (20) |
| H-index >10 | 8 (32) |
| Academic affiliation | 21 (84) |

For academic YouTube videos associated with a verified physician, the average number of views was 9676 (SD 16,888) if there was a shingles-related publication, 59,439 (SD 118,706) if there was an academic affiliation, and 42,191 (SD 184,823) if there was an h-index >10. The difference in the average number of views compared to those of videos without these characteristics was not significant (Table 2). Only 4 videos featured physicians who authored a shingles-related publication and only 3 videos featured physicians without an academic affiliation. The total number of views for the 24 academic videos with a verified physician was 1,325,693. The number of views had a median of 20,554 and an average of 55,237 (SD 111,890). The numbers of views ranged from 231 to 537,390. The video with the most views was associated with an MD-PhD dermatologist who held an academic affiliation. This video alone accounted for 41% of the total number of views for academic videos with a verified physician.

**Table 2.** Comparison of the average number of views by physician characteristics.

| Physician quality | No. of videos | Average no. of views, Mean (SD) | *P* value |
|---|---|---|---|
| **Shingles-related publication** | | | .07 |
| Yes | 4 | 9676 (16,888) | |
| No | 20 | 64,349 (120,771) | |
| **Academic affiliation** | | | .36 |
| Yes | 21 | 59,439 (118,706) | |
| No | 3 | 25,826 (39,739) | |
| **h-index** | | | .57 |
| 0-10 | 16 | 42,191 (52,065) | |
| >10 | 8 | 81,329 (184,823) | |

## *Discussion*

YouTube is one of the world's most popular online media platforms, which many medical students utilize as a resource to supplement their education. Prior studies have suggested the potential benefit that YouTube has for improving medical education. However, those studies also mentioned the need for more evaluations in order to ensure the quality of YouTube videos [8,9]. The aim of this study was to analyze the physicians associated with the production of YouTube videos that could be incorporated into medical education. This study was performed around shingles because the unique presentation of this infection makes it well-suited to be studied on online media platforms.

A significant majority of identified physicians had an affiliation with a renowned academic institution with many having produced publications. The substantial research background and academic affiliations of these physicians ensure to a certain degree the accuracy of information in educational YouTube videos about shingles. However, these surrogates for content quality cannot replace the critical analysis that peer-review provides. Preferably, an external rating system should be developed to enable critical analysis by physicians and avoid the limitations that traditional forms of peer-review would impose.

The quality of academic content in a YouTube video may contribute to its popularity. In fact, the video with the most views (537,390) was associated with an MD-PhD dermatologist who has an academic affiliation. This may provide reassurance to viewers given that the associated physician is trained in a specialty that manages patients with shingles. Although videos with physicians who held an academic affiliation or an h-index >10 did trend towards more views, this study did not find a significant difference between the number of views and research background of the physicians associated with a video. This suggests that there may be other more significant factors that contribute to the popularity of a video such as the length of time since the video was posted and the number of subscribers of the YouTube channel.

This study has a few limitations that should be considered. The most significant limitation is that the study was unable to characterize the audience of these YouTube videos. Therefore, the proportion of views that could be attributed to medical students could not be gauged. Future studies should utilize YouTube Analytics to survey the audience [12]. Other limitations which also prevented characterization of the audience include YouTube channels turning off the feedback and commenting feature. Moreover, the small sample size in certain categories was a limitation in this study. This limitation was weighed against the decreased relevance of videos as the overall sample size of the study increased. Outside of the first 50 search results were less relevant videos about other topics such as rooftop shingles.

Although this study to a certain degree supports the quality of YouTube videos on shingles for medical education, more evaluation is needed for this growing platform. Future studies should examine how medical students may utilize social media platforms to share such videos. Some studies have suggested that platforms such as WhatsApp and Instagram were highly efficacious in the dissemination of educational information linked to YouTube videos [13,14].

YouTube is an efficient and accessible resource for medical students to supplement their learning. Half of the videos reviewed were associated with verified physicians most of which had an academic affiliation. This study supports to a certain degree the academic quality of YouTube videos about shingles produced for medical education. This provides reassurance to viewers that the information they are receiving is accurate. However, more studies in the future will be needed to assess the utility of YouTube as a tool for medical education.

## Conflicts of Interest

None declared.

## References

1. Insinga RP, Itzler RF, Pellissier JM, Saddier P, Nikas AA. The incidence of herpes zoster in a United States administrative database. J Gen Intern Med 2005 Aug;20(8):748-753 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.0150.x] [Medline: 16050886]
2. Johnson BH, Palmer L, Gatwood J, Lenhart G, Kawai K, Acosta CJ. Annual incidence rates of herpes zoster among an immunocompetent population in the United States. BMC Infect Dis 2015 Nov 06;15:502 [FREE Full text] [doi: 10.1186/s12879-015-1262-8] [Medline: 26546419]
3. Koya KD, Bhatia KR, Hsu JTS, Bhatia AC. YouTube and the expanding role of videos in dermatologic surgery education. Semin Cutan Med Surg 2012 Sep;31(3):163-167 [FREE Full text] [doi: 10.1016/j.sder.2012.06.006] [Medline: 22929353]
4. Lam NHT, Tsiang JT, Woo BKP. Exploring the Role of YouTube in Disseminating Psychoeducation. Acad Psychiatry 2017 Dec;41(6):819-822 [FREE Full text] [doi: 10.1007/s40596-017-0835-9] [Medline: 29022242]
5. Lam NHT, Woo BKP. YouTube as a New Medium for Dementia Education Among Chinese Americans. Community Ment Health J 2020 Apr;56(3):435-439 [FREE Full text] [doi: 10.1007/s10597-019-00493-7] [Medline: 31641910]
6. Guo JZ, Chong KPL, Woo BKP. Utilizing YouTube as platform for psychiatric emergency patient outreach in Chinese Americans. Asian J Psychiatr 2020 Apr;50:101960 [FREE Full text] [doi: 10.1016/j.ajp.2020.101960] [Medline: 32086173]
7. Qi J, Trang T, Doong J, Kang S, Chien AL. Misinformation is prevalent in psoriasis-related YouTube videos. Dermatol Online J 2016 Nov 15;22(11). [Medline: 28329562]
8. Sunderland N, Camm CF, Glover K, Watts A, Warwick G. A quality assessment of respiratory auscultation material on YouTube. Clin Med (Lond) 2014 Aug;14(4):391-395 [FREE Full text] [doi: 10.7861/clinmedicine.14-4-391] [Medline: 25099841]
9. Shires CB, Wilson CD, Sebelik M. Thyroid surgery YouTube videos: estimating quality by surgeon characteristics and view rate. Gland Surg 2019 Jun;8(3):207-211 [FREE Full text] [doi: 10.21037/gs.2018.10.01] [Medline: 31328098]
10. Lenczowski E, Dahiya M. Psoriasis and the Digital Landscape: YouTube as an Information Source for Patients and Medical Professionals. J Clin Aesthet Dermatol 2018 Mar;11(3):36-38 [FREE Full text] [Medline: 29606999]
11. Gray R, Hassanein E, Thompson DR. Journal editors and their h-index. J Adv Nurs 2017 Sep;73(9):2031-2034 [FREE Full text] [doi: 10.1111/jan.13070] [Medline: 27400154]
12. Woo BKP, Chung JOP. Using YouTube analytics to evaluate a Chinese video-based lecture regarding Parkinson's disease. J Clin Neurosci 2018 Apr;50:300 [FREE Full text] [doi: 10.1016/j.jocn.2018.01.001] [Medline: 29361404]
13. Shu S, Woo BKP. The Roles of YouTube and WhatsApp in Dementia Education for the Older Chinese American Population: Longitudinal Analysis. JMIR Aging 2020 Apr 13;3(1):e18179 [FREE Full text] [doi: 10.2196/18179] [Medline: 32281940]
14. Lam NH, Woo BK. Efficacy of Instagram in Promoting Psychoeducation in the Chinese-Speaking Population. Health Equity 2020;4(1):114-116 [FREE Full text] [doi: 10.1089/heq.2019.0078] [Medline: 32258963]

*Edited by G Eysenbach; submitted 16.05.20; peer-reviewed by S Shu, J Chung, H Namba, N Lam, E Khoong; comments to author 29.06.20; revised version received 02.07.20; accepted 26.07.20; published 03.09.20.*

*Please cite as:*
*Dunnsiri T, Kawashita T, Lee SC, Monga AK, Woo BKP*
*Assessing YouTube as an Educational Tool for Shingles: Cross-Sectional Study*
*JMIR Dermatol 2020;3(1):e20338*
*URL: http://derma.jmir.org/2020/1/e20338/*
*doi:10.2196/20338*
*PMID:*

Original Paper

# Development of an Innovative Real-World Evidence Registry for the Herpes Simplex Virus: Case Study

Michelle Helena van Velthoven[1], BSc, MSc, PhD; Ching Lam[1], MEng; Caroline de Cock[1], BSc; Terese Stenfors[2], BSc, MSc, PhD; Hassan Chaudhury[3], BSc; Edward Meinert[1,3], MA, MSc, MBA, MPA, PhD

[1]Digitally Enabled Preventative Health Research Group, Department of Paediatrics, University of Oxford, Oxford, United Kingdom
[2]Division of Surgery, Department of Clinical Sciences, Danderyd Hospital, Karolinska Institutet, Stockholm, Sweden
[3]Department of Primary Care and Public Health, Imperial College London, London, United Kingdom

**Corresponding Author:**
Edward Meinert, MA, MSc, MBA, MPA, PhD
Digitally Enabled Preventative Health Research Group
Department of Paediatrics
University of Oxford
Headley Way, Headington
John Radcliffe Hospital
Oxford, OX3 9DU
United Kingdom
Phone: 44 7824446808
Email: e.meinert14@imperial.ac.uk

## Abstract

**Background:** Infection with the herpes simplex virus (HSV) is common but not well understood. Furthermore, there remains a social stigma surrounding HSV that can have psychosocial implications for those infected. Despite many patients infected with HSV experiencing mild-to-severe physical symptoms, only one subeffective treatment is available. A registry collecting real-world data reported by individuals potentially infected with HSV could help patients to better understand and manage their condition.

**Objective:** This study aimed to report on the development of a registry to collect real-world data reported by people who might be infected with HSV.

**Methods:** A case study design was selected as it provides a systematic and in-depth approach to investigating the planning phase of the registry. The case study followed seven stages: plan, design, prepare, collect, analyze, create, and share. We carried out semistructured interviews with experts, which were thematically analyzed and used to build use cases for the proposed registry. These use cases will be used to generate detailed models of how a real-world evidence registry might be perceived and used by different users.

**Results:** The following key themes were identified in the interviews: (1) stigma and anonymity, (2) selection bias, (3) understanding treatment and outcome gaps, (4) lifestyle factors, (5) individualized versus population-level data, and (6) severe complications of HSV. We developed use cases for different types of users of the registry, including individuals with HSV, members of the public, researchers, and clinicians.

**Conclusions:** This case study revealed key considerations and insights for the development of an appropriate registry to collect real-world data reported by people who might be infected with HSV. Further development and testing of the registry with different users is required. The registry must also be evaluated for the feasibility and effectiveness of collecting data to support symptom management. This registry has the potential to contribute to the development of vaccines and treatments and provide insights into the impact of HSV on other conditions.

***(JMIR Dermatol 2020;3(1):e16933)*** doi:10.2196/16933

**KEYWORDS**

case study; herpes simplex; herpes simplex virus vaccines; real-world evidence; patient registry

# Introduction

## Background

Globally in 2012, an estimated 3.7 billion people younger than 50 years were infected with herpes simplex virus (HSV) type 1 (HSV1) and 417 million with HSV type 2 (HSV2) [1]. HSV can be contracted through oral or genital contact and is associated with symptoms of varying severity in assorted body parts, for example, painful sores on the face (*cold sores*), genitals (*herpes*), or hands [2].

HSV lies dormant in the nerves until shedding occurs, typically between 2 and 14 days after infection; however, it can take years before noticeable symptoms appear. Many individuals are unaware that they are infected with HSV as the virus often causes no or mild symptoms. The first shedding episode is generally associated with severe symptoms (eg, painful sores). Patients are typically diagnosed when symptoms are present as many methods for the diagnosis of HSV are reliant on viral shedding. The techniques for diagnosis in both symptomatic and asymptomatic periods vary in accuracy [3]. The only HSV treatment currently available is antiviral therapy (acyclovir or Zovirax), which is prescribed to patients with sores [4]. Symptomatic episodes vary in frequency between patients; some may experience one shedding episode, whereas others can have frequent painful outbreaks triggered by impaired immune responses.

Currently, no Food and Drug Administration–approved therapeutic vaccine is available for HSV [5] as the biology and biological interactions of HSVs are not comprehensively understood. It is known that HSV2 can interact biologically with HIV1; however, the implications for vaccine development require further elucidation [6]. Previous in vivo studies of HSV vaccines have been seen to prevent transmission of HSV in pigs, rodents, and primates, but they are yet to be successful in humans [7,8]. One therapeutic vaccine, GEN-003, was shown to reduce viral shedding and outbreaks of sores in human participants with genital HSV2 infection, but the bioscience company, Genocea, did not pursue this further [9].

The lack of available, up-to-date, high-quality, standardized data is a key barrier to efficient research, management, and treatment of HSV [10]. Clinical registries can be used to collect data about individuals with a specific diagnosis or condition on a voluntary basis. Typically, this information is contributed by health care providers following consultations with patients [11]. Clinical registries are used for data collection globally; there are more than 60 clinical registries in the United Kingdom [12] and the United States [13] and approximately 100 in Sweden [14] that have contributed to substantial research outputs and quality improvement [15].

Although clinical registries are valuable to research and treatment, clinical improvements are primarily long term, with little direct patient benefit [16]. Other limitations of current registries include the lack of privacy, anonymity, informed consent, stakeholder feedback, and awareness of existing standards and processes when building or maintaining a patient registry. In addition, clinicians may be unable or unwilling to contribute to registries because of time or technical constraints [17]. In recent decades, there has been an increasing focus on incorporating patient-reported outcomes and real-world data into clinical registries to overcome these issues [11]. Real-world evidence (RWE) is a subset of evidence-based medicine that refers to health care information gathered through means outside of typical clinical research settings. The increasing ubiquity of the internet has made it more convenient for internet-literate patients to participate in data provision and to link and analyze data for planning and managing care [18].

## Aim and Objectives

Currently, there is no registry for the collection of real-world HSV data. A patient registry whereby people submit information themselves, providing real-world data on HSV, could provide more standardized and higher quality data than currently available. A patient registry is "an organised system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves one or more predetermined scientific, clinical, or policy purposes" [19]. A patient registry could help individuals to understand the triggers and indications of HSV recurrences and make changes to their lifestyle to mitigate these symptoms where possible. A real-world registry could facilitate research into the development of vaccines and new, effective treatments, and examine the impact of HSV on other conditions. The aim of this study was to develop a registry to collect real-world data reported by people who might be infected with HSV. The objectives were to obtain expert opinions on important aspects of a registry and to provide insights into how specific users might use the registry.

# Methods

## Case Study Overview

A case study design was selected to provide a structured means to generate in-depth evidence to contribute to the further development of the proposed registry. We conducted semistructured interviews with experts and analyzed their opinions on an HSV registry to establish themes. The results of the thematic analysis were used as the basis for the development of use cases (a specific user of the registry) that will be applied to the development of an RWE registry. A traceability log was maintained in Microsoft Excel linking the research questions to data sources and the study findings.

This research project brings researchers in contact with people who gave their views on the value and development of an HSV registry. The interviewees did not provide information about themselves and the opinions that they offer were not themselves the subject of research. Therefore, in line with the University of Oxford guidance, ethical approval was not required.

## Case Study Framework

The case study development followed seven stages, which are outlined in Table 1.

**Table 1.** Case study framework.

| Number | Stage | Outcome |
|---|---|---|
| 1 | Plan | Description of case and linking of approach to outcomes. |
| 2 | Design | Construction of research design and linkage of research questions, data, and criteria for evaluation and synthesis. |
| 3 | Prepare | Draft, execution, and approval of study protocols. |
| 4 | Collect | Conduction of semistructured interviews. |
| 5 | Analyze | Thematic analysis of data. |
| 6 | Create | Development of use cases for a clinical registry website. |
| 7 | Share | Presentation of the findings in a report for publication in a peer-reviewed journal (this paper) and sharing with registry design and production team. |

### *Step 1: Plan*

The research question centered on investigating the potential of a registry for collecting real-world data contributed by people who might be infected with HSV. The focus of the registry was selected because evidence from the literature suggests that although infection with HSV is common, it is not well understood. Furthermore, there is no vaccine or effective treatment for HSV, leading to morbidity for individuals suffering symptoms of viral shedding. An RWE registry would help people with HSV understand their condition and contribute valuable information to the development of vaccines and treatment. A case study can be used to inform the development of an appropriate real-world registry.

### *Step 2: Design*

We designed a case study that involved semistructured interviews with experts. The findings of these interviews were thematically analyzed. On the basis of these findings, we developed use cases that are reported in this study and will be used to design a prototype registry.

### *Step 3: Prepare*

We produced a list of 11 experts with different backgrounds, selected by iterative discussions with a health care expert (HC, Digital Health Lead at Healthcare UK, Department for International Trade). We chose four experts who were most relevant to this case study, represented different views, and were available for an interview. The interviewees included a male and female representative (experts 1 and 2) of the Herpes Viruses Association, which provides information to people infected with HSV [19], a sexual health consultant with long-standing interest and published research in HSV therapeutics (expert 3), and a scientist working on HSV and herpes simplex encephalitis (HSE; expert 4). Participants provided verbal consent to involvement in the study, note-taking during conversations, and, in one case, the conversation being recorded. The list of interview questions was discussed in detail in a steering committee meeting and then developed through repeated review (HC; Yusuf Ermak, Director KPMG Advisory; and EM, Research Fellow in Digital Health). The final questionnaire had 48 questions and was piloted in the first interview (Multimedia Appendix 1).

### *Step 4: Collect*

The interviews took place at a location convenient for the interviewees and lasted up to an hour. In the interviews, one researcher asked the questions and followed these up, and another researcher took notes (female research assistant Abrar Alturkistani, MPH, and male analyst Ben Southwood, MA). The researchers were experienced in conducting qualitative interviews and were knowledgeable about HSV. Although not all interviews were directly transcribed, there was a high level of agreement between the separate sets of notes of the interviews. The notes mostly comprised direct quotations and paraphrases. The notes were consolidated into a single document to code the responses into themes. Transcripts were not sent back to the participants.

### *Step 5: Analyze*

We thematically analyzed the content of the interviews. One researcher built the themes from the first interview and reviewed, edited, and refined the themes as the interview process went on. A second independent researcher worked on the data following completion of the thematic analysis. We used this independent coder system to increase the reliability of the coding [20]. Later, we compared the two thematic analyses to determine consistent, unified themes. The final themes were not centered around any of the questions from the interviews and were instead constructed from the data [21].

Several reviews, discussions, and reconfigurations took place before finalization of a set of themes that justified the source material. After initial coding, the two researchers met on two occasions and had several phone calls to discuss differences in their coding. The themes and subthemes identified were repeatedly passed through the data until the smallest coherent and consistent set of themes that captured all the insights was reached. After agreeing on themes, the entire interview text was passed through our agreed themes. This allowed us to generate a full count of the number of times themes and subthemes appeared in the interview texts. The findings were not reviewed by the interviewees.

### *Step 6: Create*

Having reviewed and finalized the thematic analysis of the interviews, we developed use cases for the prototype real-world registry. These were centered around four types of users with several subtypes. These use cases will inform design work regarding potential users, their motivations for use, how they would access features, and what they would achieve from interacting with this registry.

### *Step 7: Share*

This paper shares the findings of the case study.

## Results

### Semistructured Interview Findings

A high-level summary of the key themes and subthemes identified in the interviews is provided in Table 2, followed by a detailed explanation of the themes with quotes. The key themes were stigma and anonymity, selection bias problems, understanding treatment gaps and outcomes, lifestyle factors and causes, individualized versus population-level data, and severe complications of HSV1/2.

**Table 2.** Interview themes, count, and subthemes.

| Theme | Subthemes | Count, n |
|---|---|---|
| Stigma and anonymity | • Root causes<br>• Impact<br>• Anonymity<br>• Careful design | 12 |
| Selection bias problems | • Being unaware of having an HSV[a] infection<br>• Available population exaggerates seriousness<br>• Other selection bias problems<br>• Ways to ameliorate | 21 |
| Understanding treatment and outcome gaps | • Trouble getting treatment<br>• Spatial<br>• Differences by provider<br>• Treatment over time | 15 |
| Lifestyle factors and causes | • Health and fitness<br>• Sex and dating | 5 |
| Individualized versus population-level data | • Genetics of HSV<br>• Broad-based studies | 10 |
| Severe complications of HSV1/2 | • HIV<br>• Herpes simplex encephalitis<br>• Pregnancy and neonatal herpes<br>• Pain<br>• Other severe outcomes<br>• Understanding severe complications | 23 |

[a] HSV: herpes simplex virus.

### *Theme 1: Stigma and Anonymity*

Interviewees mentioned that HSV is stigmatized for different reasons, including being considered a sexually transmitted infection (STI) and misinterpretation by the public. This was thought to start from sex education when people are told that catching a sexually transmitted disease is *bad* and *disgusting*. In addition, HSV1/2 has been stigmatized by the media and by many jokes about herpes from the early 1980s onward. Furthermore, people think they know that HSV is incurable, but other conditions that stay in the body like herpes are not called incurable, including thrush and chicken pox. These findings imply that better education on HSV is required to reduce the psychosocial implications of HSV infection:

> *General public and media need to know more—could lessen psychological impact a great deal.* [Expert 1]

For stigma-related reasons, anonymity will be an important factor to consider when designing the registry as participants will likely not want their HSV infection status to be revealed. It is important not to discourage potential users; therefore, any registry must have credible security, options for anonymity, and explain data access:

> *Would have to be extremely discreet to get people to submit info about themselves. People join our association under assumed names, that's perfectly fine.* [Expert 2]

The registry needs to be designed with the user and their experience in mind (user centricity), meaning that the data collection process serves the expectations and needs of users. In the development frameworks of existing registries, patient centricity is limited, thereby reducing utility. Lack of anonymity will result in user concerns about privacy and control over data when sensitive information is requested. In addition, it was thought that terms such as *incurable*, *chronic*, *attack*, *victim*, and dramatic red colors should be avoided.

### *Theme 2: Selection Bias Problems*

The interviewees considered currently available data collection methods to be subject to selection bias problems. There are epidemiological herpes studies doing similar things as a registry but their data are not detailed, longitudinal, and broad enough. It was thought that herpes gets overrepresented by how much it bothers people and, therefore, studies might overestimate the symptoms by representing the worst cases. A patient registry could be used to register more people who are affected and people with minor symptoms who do not visit a health care professional:

> *But I think that the beauty of a registry like this would be to catch everybody affected, even those with occasional cold sores that doesn't visit their GP or a clinic and isn't diagnosed with the condition.* [Expert 4]

A registry needs to represent a diverse body of people infected with HSV, with variable symptom frequency. Interviewees expected a registry to attract mostly those experiencing severe morbidity from HSV infection. A registry must also include those with HSV and from diverse sociodemographic backgrounds and age groups. There are an increasing number of older patients, which means that a purely app-based registry could be biased toward younger generations:

> *App would bias things to younger people. Actually we do have quite a number of over 50 year olds who suffer from the more severe forms of HSV1—tentative links to Alzheimer's and dementia. I have a feeling we might lose out to these. True of even websites—but maybe websites an easier form. App/website combo might capture with largest swathe.* [Expert 4]

Future tools to collect HSV data must consider and address all the sources of bias. These tools must engage everyone who may have the condition and not just those with frequent recurrences, complications, pain, or psychological impacts.

Selection bias can become more profound during direct data collection as there would be a subset of users who are more likely to complete an extensive questionnaire, for example, more computer-literate people who have more time, patients with a confirmed diagnosis, and/or those with more frequent and/or severe symptoms keen on being informed on relevant clinical trials.

### *Theme 3: Understanding Treatment and Outcome Gaps*

Interviewees indicated gaps in current HSV treatment, management, and outcomes. In the United Kingdom, if you get diagnosed at a sexual health clinic, you get treated free. This has resulted in sexual health clinics being shifted to local councils and privatized. Given the limited funds at public clinics, people with HSV are often pushed back to their general practitioner after diagnosis and early treatment. This means that ongoing treatment and follow-up can be difficult to get as patients are shunted between services. Furthermore, general practitioners are often less knowledgeable on herpes than sexual health care professionals:

> *It can be difficult to get hold of because of the messy nature of the treatment landscape. Some patients can get thru GP and some patients can't. Some patients can repeatedly go to their sexual health clinic, some can't.* [Expert 1]

A registry that collects more data could work out who is having trouble getting treatment. In addition, it could collect follow-up data as, for example, people who receive acyclovir typically do not return to their health care professional. Efficacy of treatment is another issue, and there have been anecdotal reports on ineffectiveness or misdiagnosis. A registry should take into consideration how these current gaps could be improved by real-world data collection:

> *At the moment all you have is patient reflection, it has to be better than that, maybe a weekly diary. otherwise it becomes reflective and three months later they say it was like this...so it has to be real-time.* [Expert 3]

A key benefit of a real-world registry would be the ability to collect data on people who are not seen in the health care system. Data on these people may reveal new insights into the transmission of HSV and what adds value for people with HSV symptoms:

> *Very little data on unmet need, since they don't contact us or anybody else after they've been diagnosed. "I've been sitting in my house for 10 years not having relationships—thank God I found you." Wide scale research across the general population.* [Expert 1]

### *Theme 4: Lifestyle Factors and Causes*

The interviewees mentioned that HSV episodes are related to a low immune response, which can be influenced by lifestyle factors; however, the existing evidence is insufficient to detail the specific lifestyle factors triggering these episodes and guidance for mitigation:

> *If people are getting too many outbreaks they may want to change their lifestyle. Cold sores is like that because you get them when you have a cold. Understanding underlying health/fitness is potentially useful.* [Expert 2]

Lifestyle factors play an important role in the spread and management of HSV. The use of collecting information regarding interaction between partners is important to consider. A registry could be used to obtain data on how long they have been living with a partner and how many times they had sex:

> *Don't really understand how often patients infect partners. Have indirect evidence of how quickly people get infected. Think infection diminishes in longer relations. Would be useful to know how long in relationship, how often they have sex.* [Expert 3]

In addition, it was noted that dating apps are starting to move into this area, for example, reminding people about taking STI testing or telling other people who do not use a condom to get screened.

### *Theme 5: Individualized Versus Population-Level Data*

The interviewees highlighted the importance of collecting data that will be useful for research. It was mentioned that there is no reliable national data on HSV1 and HSV2 estimates. The focus has been on individualized data, and it is resource intensive to obtain swabs from individual patients. Collecting general population–level data would be useful; a registry would be able to provide individualized and broad RWE. However, it was thought to not be sufficiently informative for researchers. It was thought that a registry should also include relevant family history, for example, infectious disease, immunodeficiency, ethnicity, and anything running in the family:

> *I think it would give you more real world angle on how effective therapies are. You would be able to work out more quickly if its effective or not.* [Expert 3]

The interviewees anticipated that the medical and research community may not be interested in only information that a traditional clinical registry may provide. Links with health care services data or genetic databases may provide a more useful long-term source and breadth of data:

> *I think it would be great to know the genomic consequences of herpes infection, at the moment we use the natural history.* [Expert 3]

### *Theme 6: Severe Complications of Herpes Simplex Virus 1/2*

The interviewees reported that the majority of HSV cases are mild and easily manageable, and HSV is usually not especially harmful to those affected. However, severe HSV cases can be debilitating, are not well understood, and require further research to improve understanding of the pathophysiology behind their occurrence. In rare cases, HSV can have costly and damaging complications (eg, HSE). HIV status will also need to be considered as HSV infection increases HIV transmission. It was thought to be gratifying for the patient to have information around why they suffer such a life-threatening disease when so many others do not:

> *Why do some get very strongly affected? Only 5% is explained by us. Not all due to genetics but we don't know any of the other factors. There is no way to predict the final outcome of the patient from things we currently observe—we don't understand what the parameters are that determine the severity of the infection. We only know that as soon as patients present with a brain infection we need to give them acyclovir FAST- and severity correlates with this but not perfectly.* [Expert 4]

Furthermore, newborn babies get neonatal herpes usually from the mother during birth. If a man infects a woman toward the end of pregnancy, she may actually have a baby born with severe neurological damage. It was thought that there is a rise in neonatal herpes, but it was uncertain exactly how much and not enough to be a priority in the UK health care system.

## Use Cases

After completing the thematic analysis, we developed use cases, which show potential journeys of users (Table 3). The use cases are drawn from the themes, explaining how people may use the registry and database. These use cases can broadly be categorized into members of the public, patients, clinicians, and researchers. Several issues were found after an analysis of both the themes and different use cases.

People who have HSV can be categorized into three groups: those with frequent recurrences, those with mild or rare recurrences, and pregnant women at risk of transmitting the virus vertically during parturition. Those with mild or rare recurrences might be less motivated to log their data, particularly, if they have concerns around the stigma attached to HSV. To overcome selection bias, a wide body of patients needs to be reached. As the proposed HSV registry collects data directly from users, considering ways to ensure anonymity is important. It is additionally important to ask as few questions as possible to minimize the time and effort required from the participant for data collection while ensuring that the data are comprehensive.

Lifestyle factors play a major role in the spread and management of HSV. People with frequent recurrences of HSV symptoms are most likely to contribute to an HSV registry that can help them reduce the frequency and severity of symptoms that they experience compared with a registry where the data collected do not support their personal health. Those that are not infected with HSV may be interested in engaging with an HSV registry to understand the risk factors for acquiring the virus; this is particularly important for immunodeficient individuals such as those with HIV.

Researchers will be interested in using these data to advance the development of a vaccine and better treatment. To enable a population-level view, the registry needs to provide interoperability across systems and adhere to the common and widely accepted data. Clinicians will be interested in using this to improve their clinical knowledge and practice, particularly for patients with severe HSV symptoms.

**Table 3.** Use cases.

| Use case | I am a... | I want to... | So that I can... |
|---|---|---|---|
| Person with HSV[a] infection | Patient with HSV experiencing frequent recurrences | Keep track of my recurrences and treatments | Reduce the number and severity of recurrences |
| | Patient with HSV experiencing mild and/or rare recurrences | Add information to the database | Help the scientific community advance knowledge |
| | Pregnant patient with HSV | Learn about neonatal HSV | Avoid transmission of HSV to my baby |
| Researcher | HSV1/2 researcher | Elucidate interactions between HSV and other conditions or demographic factors | Contribute to novel research to advance the understanding of HSV among the scientific and medical community |
| | Herpes simplex encephalitis researcher | Understand the factors related to severe HSV | Prevent HSV infection leading to severe conditions such as encephalitis |
| Members of the public | General public | Learn basic facts and risk factors for HSV | Reduce risk of infection, understand and adopt safe practices, and assuage fear and stigma about HSV |
| | HIV-positive member of public | Learn how HIV and HSV1/2 interact and implications for personal health | Reduce risk transmission of HIV to others |
| Clinicians | General practitioner | Improve clinical insights about HSV | Optimize management of HSV |
| | Sexually transmitted infection specialist clinician | Track patients with HSV and related conditions | Ensure optimal treatment of patients and mitigate the risk of severe complications |

[a]HSV: herpes simplex virus.

## Discussion

### Principal Findings

The key themes identified in the interviews were (1) stigma and anonymity of HSV, (2) selection bias, (3) understanding treatment and outcome gaps, (4) lifestyle factors and causes, (5) individualized versus population-level data, and (6) severe complications of HSV. On the basis of these themes, we developed use cases that show the potential needs and motivations of users. These can be used to explain how people may use the registry. These use cases can broadly be categorized into members of the public, individuals with HSV, clinicians, and researchers.

The themes and use cases identified in this study are consistent with the literature. The interviewees reported that HSV can have a considerable negative psychosocial impact and affect people's relationships because of the stigma surrounding HSV [22]. Other reports corroborate that although people are generally supportive of sharing their data for research, they expect their data to be anonymized [23]. However, it has been stipulated that people may be more comfortable with sharing data if they are asked by a trusted organization (eg, universities, hospitals, or disease foundations) [24]. Addressing privacy concerns is critical in the development of an HSV registry, and anonymization is required to provide a secure environment for data sharing [25].

A second issue raised by the interviewees was different sources of selection bias. Although the internet is increasingly accessible and utilized, it is not ubiquitous. Previous research has shown limited adoption in certain sociodemographic groups related to language barriers, levels of internet literacy, and geographical locations [26]. These need to be addressed to reduce selection bias and maximize the breadth of the data obtained. Another source of selection bias is that people without symptoms are generally unaware of having HSV and those with rare or mild symptoms may be less interested in enrolling and contributing data to a registry than people with severe HSV. To address these limitations, a registry must provide a meaningful return for all those contributing data to ensure the breadth of reach and minimize selection biases. A common problem in digital data collection is the inverse correlation between engagement and number of questions/steps in the user journey. This conflicts with the overarching goal of the registry to obtain a breadth and depth of demographic information from the users to maximize the potential contribution of the registry to research and trial enrollment. It will be important to ascertain the optimal balance between collecting meaningful amounts of data and the effort required from users.

A potential use of the proposed registry highlighted by the interviewees was the potential for real-world data to address current treatment and knowledge gaps. Only one treatment for HSV is currently available—prescribed at physicians' discretion, the optimal timing, duration, and effectiveness of which is uncertain for both the first onset and subsequent recurrences [27]. A registry collecting information on treatment prescription, adherence and efficacy, and timing and severity of recurrence could help to address these unknowns. Furthermore, the collection of data on the comorbidities of those infected and their HSV journey could help to shed light on the interaction of HSV with other conditions, such as Alzheimer disease and general cognitive ability [28,29].

Furthermore, recurrences of HSV can be influenced by lifestyle factors; however, these are not comprehensively understood or well communicated to infected individuals. The frequency and severity of symptoms vary among those with HSV and depend on a person's age, gender, immune system, and the route of transmission and type of virus [4]. A registry could help people with HSV and researchers to identify the links between life events and recurrences and change their lifestyle to mitigate or prevent recurrences. It is additionally not well known why an HSV infection can lead to more serious, often debilitating complications, such as encephalitis, in rare cases [30]. A registry could be a tool to facilitate research into these areas.

The interviewees speculated that the medical and research community may not be interested in engaging with a registry that provides information comparable with a traditional registry; therefore, links with health care services data and/or genetic databases may provide a more useful long-term source of data and give it a more diverse range of use cases.

### Strengths and Limitations of the Study

The case study design is a strength of this study as it supports a systematic means of observing the subject of investigation. This research provides a structured means to generate evidence to subsequently evaluate such claims by collecting baseline data for further evaluation.

The limitations of this study are that it comprises a single case study conducted in a UK context, which limits its generalizability to other settings. Only four experts were interviewed; however, they reflected different users and provided an in-depth and rigorous insight into the key issues related to an HSV real-world registry. We conducted a thorough thematic analysis of the findings of these semistructured interviews. Future work can build on our rigorous case study. We built a step-by-step working methodology, according to the principles of case study research, and consolidated all our source materials together. The findings were shared with the registry design and production team and were presented in this case study report.

### Further Research

Work to further develop the user personas identified in this case study is underway. Using the agile delivery project management framework [31], the project will be delivered in four stages: discovery phase, in which we try to better understand the users and use cases, constraints and improvement opportunities; alpha phase, in which we will develop a prototype of the solution; beta phase, in which we will build the HSV registry and test and evaluate the registry with a small patient population; and the live phase, in which the registry will be available to the public and continuous iterations and improvement will be made where necessary.

We have completed the discovery phase where we identified the key factors that impact data collection quality, technological constraints, and potential improvements that can be made for HSV registries. The alpha phase is ongoing, and a prototype of the registry is being built. The results of the discovery and alpha phases will be published subsequently.

Further research is needed on testing this registry with different users and implementing and evaluating its feasibility and effectiveness of collecting data to support the development of vaccines and treatments as well as addressing numerous knowledge gaps in this field.

### Conclusions

This case study reports insights from interviews for the development of an appropriate registry for real-world data contributed by people with HSV. In existing registries, user centricity is limited, which reduces people's use of the system. The motivations and expectations of users need to be taken into account when designing the system to overcome the limitations of traditional registries. The stigma associated with HSV elicits critical privacy concerns. People who might be infected with HSV need to have the option to be anonymous, and transparency regarding the use of the data is paramount. A patient registry could help those affected by HSV to better understand their condition and reduce or mitigate recurrences. Furthermore, a registry could provide researchers and clinicians with a tool to systematically collect higher quality data for studying HSV epidemiology and to evaluate the efficacy and cost-effectiveness of treatments and vaccines.

## Acknowledgments

This work was funded by the European Institute of Innovation and Technology Health (grant 18654).

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview questions.
[DOCX File , 17 KB - derma_v3i1e16933_app1.docx ]

## References

1. Looker KJ, Magaret AS, May MT, Turner KM, Vickerman P, Gottlieb SL, et al. Global and regional estimates of prevalent and incident herpes simplex virus type 1 infections in 2012. PLoS One 2015 Oct 28;10(10):e0140765 [FREE Full text] [doi: 10.1371/journal.pone.0140765] [Medline: 26510007]
2. Li C, Li Y, Yang Y, Wang J, Zhu C, Tang S, et al. The detection and characterization of Herpes Simplex Virus Type 1 in confirmed measles cases. Sci Rep 2019 Sep 4;9(1):12785 [FREE Full text] [doi: 10.1038/s41598-019-48994-5] [Medline: 31484944]

3. Arshad Z, Alturkistani A, Brindley D, Lam C, Foley K, Meinert E. Tools for the diagnosis of Herpes Simplex Virus 1/2: systematic review of studies published between 2012 and 2018. JMIR Public Health Surveill 2019 May 23;5(2):e14216 [FREE Full text] [doi: 10.2196/14216] [Medline: 31124465]
4. Johnston C, Corey L. Current concepts for genital Herpes Simplex Virus infection: diagnostics and pathogenesis of genital tract shedding. Clin Microbiol Rev 2015 Nov 11;29(1):149-161. [doi: 10.1128/cmr.00043-15]
5. Srivastava R, Roy S, Coulon PG, Vahed H, Prakash S, Dhanushkodi N, et al. Therapeutic mucosal vaccination of Herpes Simplex Virus 2-infected guinea pigs with ribonucleotide reductase 2 (RR2) protein boosts antiviral neutralizing antibodies and local tissue-resident CD4 and CD8 T cells associated with protection against recurrent genital herpes. J Virol 2019 May 01;93(9):803-805 [FREE Full text] [doi: 10.1128/JVI.02309-18] [Medline: 30787156]
6. Schiffer JT, Gottlieb SL. Biologic interactions between HSV-2 and HIV-1 and possible implications for HSV vaccine development. Vaccine 2019 Nov 28;37(50):7363-7371 [FREE Full text] [doi: 10.1016/j.vaccine.2017.09.044] [Medline: 28958807]
7. Johnston C, Koelle D, Wald A. HSV-2: in pursuit of a vaccine. J Clin Invest 2011 Dec;121(12):4600-4609 [FREE Full text] [doi: 10.1172/JCI57148] [Medline: 22133885]
8. Johnston C, Gottlieb S, Wald A. Status of vaccine research and development of vaccines for herpes simplex virus. Vaccine Internet(26)?2952 2016;34:2948 [FREE Full text] [doi: 10.1016/j.vaccine.2015.12.076]
9. Flechtner J, Long D, Larson S, Clemens V, Baccari A, Kien L, et al. Immune responses elicited by the GEN-003 candidate HSV-2 therapeutic vaccine in a randomized controlled dose-ranging phase 1/2a trial. Vaccine 2016 Oct 17;34(44):5314-5320. [doi: 10.1016/j.vaccine.2016.09.001] [Medline: 27642130]
10. Jørgensen LK, Dalgaard LS, Østergaard LJ, Nørgaard M, Mogensen TH. Incidence and mortality of herpes simplex encephalitis in Denmark: a nationwide registry-based cohort study. J Infect 2017 Dec;74(1):42-49. [doi: 10.1016/j.jinf.2016.09.004] [Medline: 27717782]
11. Trotter JP. Patient registries: a new gold standard for 'real word' research. Ochsner J 2002;4(4):211-214 [FREE Full text] [Medline: 22826660]
12. National Health Service. Clinical audits and registries URL: https://digital.nhs.uk/data-and-information/clinical-audits-and-registries [accessed 2020-02-04]
13. US National Institutes of Health. List of Registries URL: https://www.nih.gov/health-information/nih-clinical-research-trials-you/list-registries [accessed 2020-02-04]
14. Swedish Research Council. Registerforskning URL: https://www.registerforskning.se/en/ [accessed 2020-02-04]
15. Feltelius N, Gedeborg R, Holm L, Zethelius B. Utility of registries for post-marketing evaluation of medicines. A survey of Swedish health care quality registries from a regulatory perspective. Ups J Med Sci 2017 Mar 03;122(2):136-147. [doi: 10.1080/03009734.2017.1285837]
16. Al-Rubeaan KA, Youssef AM, Subhani SN, Ahmad NA, Al-Sharqawi AH, Ibrahim HM. A web-based interactive diabetes registry for health care management and planning in Saudi Arabia. J Med Internet Res 2013 Sep 09;15(9):e202 [FREE Full text] [doi: 10.2196/jmir.2722] [Medline: 24025198]
17. Wright A, McGlinchey EA, Poon EG, Jenter CA, Bates DW, Simon SR. Ability to generate patient registries among practices with and without electronic health records. J Med Internet Res 2009 Aug 10;11(3):e31 [FREE Full text] [doi: 10.2196/jmir.1166] [Medline: 19674961]
18. Ashley L, Jones H, Thomas J, Newsham A, Downing A, Morris E, et al. Integrating patient reported outcomes with clinical cancer registry data: a feasibility study of the electronic Patient-Reported Outcomes From Cancer Survivors (ePOCS) system. J Med Internet Res 2013 Oct 25;15(10):e230 [FREE Full text] [doi: 10.2196/jmir.2764] [Medline: 24161667]
19. Gliklich RE, Dreyer NA. In: Leavy MB, editor. Registries for Evaluating Patient Outcomes: A User's Guide. Rockville, Maryland, United States: Agency for Healthcare Research and Quality; Apr 2014.
20. Herpes Viruses Association. URL: https://herpes.org.uk/ [accessed 2020-02-04]
21. Gorden R. Basic Interviewing Skills. Itasca, Illinois: Peacock Publishers; 1992.
22. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
23. Merin A, Pachankis JE. The psychological impact of genital herpes stigma. J Health Psychol 2010 Jul 23;16(1):80-90. [doi: 10.1177/1359105310367528]
24. El Emam K, Rodgers S, Malin B. Anonymising and sharing individual patient data. Br Med J 2015 Mar 20;350:h1139 [FREE Full text] [doi: 10.1136/bmj.h1139] [Medline: 25794882]
25. Nass SJ, Levit LA, Gostin LO. Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research. Bethesda, Maryland, United States: National Institutes of Health; 2009.
26. European Medicines Agency. Data anonymisation -a key enabler for clinical data sharing URL: https://www.ema.europa.eu/en/documents/report/report-data-anonymisation-key-enabler-clinical-data-sharing_en.pdf [accessed 2020-02-04]
27. Blank G, Duton W. University of Oxford. Perceived Threats to Privacy Online: The Internet in Britain Internet URL: https://oxis.oii.ox.ac.uk/wp-content/uploads/sites/43/2019/09/OxIS-report-2019-final-digital-PDFA.pdf [accessed 2020-02-04]
28. Arduino PG, Porter SR. Oral and perioral herpes simplex virus type 1 (HSV-1) infection: review of its management. Oral Dis 2006 May;12(3):254-270. [doi: 10.1111/j.1601-0825.2006.01202.x] [Medline: 16700734]

29. Looker KJ, Elmes JA, Gottlieb SL, Schiffer JT, Vickerman P, Turner KM, et al. Effect of HSV-2 infection on subsequent HIV acquisition: an updated systematic review and meta-analysis. Lancet Infect Dis 2017 Dec;17(12):1303-1316 [FREE Full text] [doi: 10.1016/S1473-3099(17)30405-X] [Medline: 28843576]
30. Itzhaki RF, Lathe R. Herpes viruses and senile dementia: first population evidence for a causal link. J Alzheimers Dis 2018;64(2):363-366. [doi: 10.3233/JAD-180266] [Medline: 29889070]
31. Whitley RJ, Kimberlin DW. Herpes simplex encephalitis: children and adolescents. Semin Pediatr Infect Dis 2005 Jan;16(1):17-23. [doi: 10.1053/j.spid.2004.09.007] [Medline: 15685145]

## Abbreviations

**HSE:** herpes simplex encephalitis
**HSV:** herpes simplex virus
**HSV1:** herpes simplex virus type 1
**HSV2:** herpes simplex virus type 2
**RWE:** real-world evidence
**STI:** sexually transmitted infection

*Edited by G Eysenbach; submitted 06.11.19; peer-reviewed by J Farzi, N Lindström; comments to author 28.11.19; revised version received 04.12.19; accepted 24.01.20; published 12.03.20.*

*Please cite as:*
*van Velthoven MH, Lam C, de Cock C, Stenfors T, Chaudhury H, Meinert E*
*Development of an Innovative Real-World Evidence Registry for the Herpes Simplex Virus: Case Study*
*JMIR Dermatol 2020;3(1):e16933*
*URL: http://derma.jmir.org/2020/1/e16933/*
*doi:10.2196/16933*
*PMID:*

Original Paper

# Understanding Social Media Use and Engagement Among Dermatology Patients to Inform Dermatological Prevention and Care in Vietnam: Cross-sectional Study

Sau Huu Nguyen[1,2], MD; Giang Thu Vu[3], MSc; Long Hoang Nguyen[4], MPH; Cuong Tat Nguyen[5,6], MPH; Thu Hoai Thi Le[1,2], MD; Tung Hoang Tran[7], MD; Bach Xuan Tran[8,9], PhD; Carl A Latkin[9], PhD; Wilson W S Tam[3,10], PhD; Cyrus S H Ho[11], MBBS; Roger C M Ho[4,12,13], MBBS

[1]National Hospital of Dermatology and Venereology, Hanoi, Vietnam
[2]Department of Dermatology and Venereology, Hanoi Medical University, Hanoi, Vietnam
[3]Center of Excellence in Evidence-based Medicine, Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam
[4]Center of Excellence in Behavioral Medicine, Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam
[5]Institute for Global Health Innovations, Duy Tan University, Da Nang, Vietnam
[6]Faculty of Medicine, Duy Tan University, Da Nang, Vietnam
[7]Institute of Orthopaedic and Trauma Surgery, Vietnam-Germany Hospital, Hanoi, Vietnam
[8]Institute for Preventive Medicine and Public Health, Hanoi Medical University, Hanoi, Vietnam
[9]Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States
[10]Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore
[11]Department of Psychological Medicine, National University Hospital, Singapore, Singapore
[12]Department of Psychological Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore
[13]Institute for Health Innovation and Technology (iHealthtech), National University of Singapore, Singapore, Singapore

**Corresponding Author:**
Bach Xuan Tran, PhD
Institute for Preventive Medicine and Public Health
Hanoi Medical University
No 1 Ton That Tung Street
Dong Da District
Hanoi, 100000
Vietnam
Phone: 84 982228662
Email: bach.ipmph@gmail.com

## Abstract

**Background:** Social media has emerged as a common source of dermatological information. Monitoring the patterns of social media use and engagement is important to counteract the limitations of social media. However, evidence in Vietnamese dermatology patients is lacking.

**Objective:** This study aimed to explore social media use and engagement by dermatology patients and to identify factors associated with social media use and engagement.

**Methods:** A cross-sectional study was conducted with 519 participants at the Vietnam National Hospital of Dermatology and Venereology during September to November 2018. Data about sociodemographic characteristics, social media use, and social media engagement were collected. Multivariate logistic and tobit regression models were used to identify factors associated with social media use and engagement.

**Results:** Interest in information about “cosmetic, beauty, and skincare techniques” was the greatest (184/519, 46.2%). The mean engagement score was 8.4 points (SD 2.4 points). Female patients were more likely to use social media (odds ratio [OR] 2.23, 95% CI 1.23-4.06) and be interested dermatological information on social media (OR 3.09, 95% CI 1.35-7.09). Women also had higher social media engagement scores (coefficient=0.68, 95% CI 0.17-1.18). Higher social media engagement scores were related with Instagram use (coefficient=0.58, 95% CI 0.00-1.15) and higher credibility scores for “family members” (coefficient=0.15, 95% CI 0.03-0.26) and “dermatology companies” (coefficient=0.22, 95% CI 0.04-0.39).

**Conclusions:** This study discovered high social media usage among dermatology patients. However, only moderate utilization and credibility levels were reported regarding the use of social media as a source of dermatological information. More efforts should focus on involving dermatologists in the development of individualized information on social media targeting specific groups of dermatology patients.

***(JMIR Dermatol 2020;3(1):e13424)*** doi:10.2196/13424

**KEYWORDS**

dermatology; social media; engagement; prevention; Vietnam

## Introduction

Dermatological diseases are popular health issues and the leading causes of disability and mortality worldwide. Skin diseases contributed to 1.79% of the global burden of diseases in 2017; of the skin diseases, dermatitis, acne vulgaris, psoriasis, and urticaria represented the greatest burden [1]. Dermatological diseases are also the most common reasons for health care service utilization [2,3], accounting for 8.4% of primary care visits [4]. People with dermatological diseases often have low quality of life [5] and self-esteem as well as feelings of stigmatization [6]. However, lack of access to dermatology services, particularly in low and middle-income countries [1], still challenges the provision of appropriate care for dermatology patients. As a result, patients seek other ways for self-treatment, including finding information on the internet and social media platforms.

Internet has played a key role in health care by reducing the cost of health service delivery [7], providing health information for specific disorders [8,9], delivering behavioral interventions [10], reducing harm for adverse health habits [11,12], facilitating physical rehabilitation [13], providing access to potential patients during health care crises [14], and supporting caregivers [15,16]. Social media is a popular method of communication [17] that enables internet users to generate information and share their opinions or media (eg, photos, clips) [18,19]. The term social media encompasses a diversity of platforms such as social networking sites (eg, Facebook, Instagram), blogs (eg, Twitter, Tumblr), or media sharing (eg, YouTube) [20-22]. It is estimated that more than 3.1 billion people worldwide were active social media users in 2018 [23]. The use of social media by patients to access medical information has accelerated with overall internet use [24].

Social media is well-recognized as a useful aid for health care providers to communicate and support patients [25-27]. Moreover, it empowers patients by encouraging active information seeking about disease prevention and treatment [28,29], providing support among peers [29,30], improving self-efficacy and self-management [27], supplementing information provided by health professionals [28], and facilitating the patient-health care professional relationship [31]. Social media is a dynamic teaching tool for health professionals [32,33] and has modernized dermatology training [34]. Although social media is a convenient source for health information, there is a paucity of literature that has evaluated the quality [35] and credibility [27] of the information on social media. Therefore, patient misinformation through social media use is a major concern among health professionals [25].

Social media is frequently used by patients to solicit advice related to dermatological concerns, has the potential to advance professional training in dermatology, and can facilitate new research methods [36]. Additionally, for dermatology, social media has the advantages of enabling the sharing of visual components such as images and videos that are important for diagnosis and consultancy. Nonetheless, little is known about the impact of social media on dermatology patients in developing countries.

Vietnam is among the countries with a high burden of dermatological diseases. A recent estimate indicates that skin disorders accounted for 2.3% of the disease burden in Vietnam in 2017 [1]. In addition, the growth in internet and social media use in Vietnam is substantial. In 2015, Vietnam had 44.4 million internet users, and this number grew to 55.8 million in 2018 [37,38]. Social media has been used widely in Vietnam, with about 46 million active users of some of the dominant social media platforms including Facebook, YouTube, Zalo, and Instagram [23]. Despite a call to understand social media use and engagement for seeking health information among dermatology patients or people interested in dermatological issues [39-43], no studies have determined social media use in this population in Vietnam. Hence, this study aimed to explore social media use and engagement by dermatology patients and to identify factors associated with social media use and engagement.

## Methods

### Study Design and Sampling Method

A cross-sectional study was performed at the Vietnam National Hospital of Dermatology and Venereology (NHDV) from September to November 2018. The NHDV was chosen due to the diversity of dermatology illnesses and the background of treated patients. As a leading hospital for dermatology and venereology diagnosis and treatment in Vietnam, the NHDV receives patient referrals from health facilities at various levels. A convenient sampling technique was adopted to recruit participants for this study. Sample size calculation was conducted using inputs determined based on a study of a similar topic in 13 European countries [44]: an expected mean social engagement score of 0.70, expected standard deviation of 0.20, and a confidence level of 95%. The calculation resulted in a minimum sample size of 385 participants. We selected from the pool of patients attending the outpatient department at the NHDV using the pre-determined eligibility criteria of age ≥18 years, a diagnosis of any dermatology disease, receiving services in the outpatient clinic, ability to provide coherent answers to the interview questions, and agreeing to participate by providing written consent. A total of 519 participants was successfully recruited for the study. The response rate was 100%.

### Measurements and Instruments

We built a structured questionnaire to collect data concerning the sociodemographic status of participants and their social media use and engagement regarding dermatological issues and information. The questionnaire also assessed various aspects of health status and service utilization, based on data from other studies exploring these topics [45]. The questionnaire was first piloted in 10 patients to validate the language and logic of each item. After revising the questionnaire based on the patient feedback, face-to-face interviews were conducted by undergraduate medical students from the Hanoi Medical University, who were well-trained in conducting community interviews. A private room at the hospital was used to hold the interviews to ensure the confidentiality and comfort of the participants.

#### *Sociodemographic Characteristics*

Data about age, gender, education, marital status, occupation, and living location were collected.

#### *Social Media Use*

In this study, we collected information about the type of social network sites the participants visited and used frequently for dermatology care (eg, What are the social media sites that you use frequently for dermatology care?) and the types of dermatological information in which they were interested to find online. In addition, we asked the participants to rate the credibility of the dermatological information sources on social media including family members, friends/relatives, celebrities (including medical professionals/clinics), television programs/magazines, dermatological product retailers (people who sell dermatology-related products), and dermatology companies (companies manufacturing and selling dermatology-related products). To obtain the credibility score, each source was rated on a scale from 0 points (Totally not credible) to 10 points (Totally credible).

#### *Engagement With Social Media*

Data on social media engagement were collected using three questionnaire items: searching for dermatological information on social media, sharing dermatological information on social media, and applying dermatological information obtained from social media. Each item used a Likert scale with five response levels: “always” (5 points) to “never” (1 point). Then, we calculated the engagement score by summing the scores of the three items. The highest possible score was 15 points, and the lowest possible score was 3 points. This approach was adapted from the Social Media Engagement theory [46].

### Statistical Analysis

We analyzed the data using Stata version 15.0 (Stata Corp. LP, College Station, TX). Multivariable logistic regression analysis was used to determine the factors correlated with social media use and interest in any dermatological information on social media. Multivariable tobit regression analysis was used to identify the factors associated with the engagement score. Potential explanatory variables included sociodemographic characteristics (age, gender, education, marital status, and occupation), social network platforms, and the credibility score. Stepwise forward selection strategies were combined with the multivariable regressions to reduce the models. A value of 0.2 of the log-likelihood’s $P$ value was considered the threshold of variable selection. The results of a previous study indicated that engagement depended heavily on specific contexts, ie, specific social network platforms as well as the credibility of the information from each platform [47]. $P<.05$ was considered statistically significant.

### Ethical Approval

The Institutional Review Board of the NHDV approved the study protocol (document number 855/HDDDBVDLTU dated September 7, 2018).

## Results

Of the 519 dermatology patients participating in the study, 62.8% (326/519) resided in urban areas. The mean age was 35.7 years (SD 13.7 years), and the greatest proportion of participants was 18-30 years old. Participants were mostly women (282/506, 55.7%), had completed at least a vocational education (341/517, 66.0%), and were married or had a partner (323/518, 62.4%). The most commonly reported occupation was freelancing (176/518, 34.0%), followed by white-collar job (145/518, 28.0%). Atopic dermatitis accounted for the greatest proportion of dermatology diseases (127/519, 24.4%), followed by contact dermatitis (75/519, 14.5%) and skin fungal infections (57/519, 11.0%; Table 1).

Facebook (359/392, 91.6%) and Zalo (247/392, 63.2%) were the most commonly used social media platforms. Information about “cosmetic, beauty, and skin care techniques” was sought by the most participants (184/397, 46.3%), followed by “general information about dermatology diseases” (168/397, 42.5%) and “medical institutions for dermatology disease treatment” (132/397, 32.6%). Regarding the credibility score of information sources on social media, “family members” and “television programs/magazines” scored the highest (mean 6.9 points, SD 2.3 points; mean 6.9 points, SD 2.2 points, respectively). Information from dermatological product retailers scored the lowest (mean 4.8 points, SD 2.3 points). No differences were found between men and women regarding the credibility of information sources or the social network platforms used (Table 2).

**Table 1.** Sociodemographic and clinical characteristics of the participants, N=519.

| Characteristics | n (%) |
|---|---|
| **Living location** | |
| Urban | 326 (62.8) |
| Rural | 193 (37.2) |
| **Age group (n=507)** | |
| 18-30 years | 239 (47.1) |
| 31-40 years | 114 (22.5) |
| 41-50 years | 72 (14.2) |
| 51-60 years | 46 (9.1) |
| >60 years | 36 (7.1) |
| **Gender (n=506)** | |
| Male | 224 (44.3) |
| Female | 282 (55.7) |
| **Education (n=517)** | |
| Up to secondary school | 73 (14.1) |
| Upper secondary school | 103 (19.9) |
| Vocational education and higher | 341 (66.0) |
| **Marital status (n=518)** | |
| Single | 195 (37.6) |
| Having a partner/married | 323 (62.4) |
| **Occupation (n=518)** | |
| Unemployed | 13 (2.5) |
| Freelancer | 176 (34.0) |
| White-collar worker | 145 (28.0) |
| Blue-collar worker | 70 (13.5) |
| Student | 62 (12.0) |
| Other | 52 (10.0) |
| **Dermatology diseases (n=519)** | |
| Atopic dermatitis | 127 (24.4) |
| Contact dermatitis | 75 (14.5) |
| Psoriasis | 29 (5.6) |
| Skin infections | 13 (2.5) |
| Skin fungal infections | 57 (11.0) |
| Urticaria | 44 (8.5) |
| Warts | 16 (3.1) |
| Zona | 31 (6.0) |
| Age (n=507), years | 35.7 (13.7)[a] |

[a] Mean (SD).

**Table 2.** Social media use among dermatology patients, with comparisons between genders.

| Characteristics | Total sample | Men | Women | *P* value |
|---|---|---|---|---|
| **Uses a social network, n (%) (n=493)** | | | | |
| Yes | 393 (79.7) | 156 (71.9) | 237 (85.9) | .00[a] |
| No | 100 (20.3) | 61 (28.1) | 39 (14.1) | |
| **Social network platform, n (%)** | | | | |
| Facebook (n=392) | 359 (91.6) | 144 (92.3) | 215 (91.1) | .67[a] |
| Instagram (n=391) | 103 (26.3) | 35 (22.4) | 68 (28.8) | .16[a] |
| Zalo (n=392) | 247 (63.2) | 89 (57.4) | 158 (66.9) | .06[a] |
| Other (n=392) | 18 (4.6) | 5 (3.2) | 13 (5.5) | .29[a] |
| **Type of dermatological information sought on social network sites, n (%)** | | | | |
| Cosmetic, beauty, and skin care techniques (n=397) | 184 (46.3) | 30 (19.1) | 154 (64.2) | <.001[a] |
| Medical institution for cosmetic surgery or beauty salons (n=397) | 95 (23.9) | 18 (11.5) | 77 (32.1) | <.001[a] |
| Beauty or cosmetic surgery experts (n=397) | 60 (15.1) | 10 (6.4) | 50 (20.8) | .10[a] |
| General information about dermatology diseases (n=397) | 168 (42.3) | 68 (43.3) | 100 (41.7) | .75[a] |
| Preventive methods for dermatology disease (n=397) | 90 (22.7) | 41 (26.1) | 49 (20.4) | .19[a] |
| Treatment methods for dermatology diseases (n=397) | 95 (23.9) | 41 (26.1) | 54 (22.5) | .41[a] |
| Medical institutions for dermatology disease treatment (n=397) | 132 (33.2) | 51 (32.5) | 81 (33.8) | .79[a] |
| Medication for dermatology disease treatment (n=397) | 128 (32.2) | 54 (34.4) | 74 (30.8) | .46[a] |
| Treatment experiences from people who had undergone dermatology disease treatment (n=397) | 80 (20.2) | 28 (17.8) | 52 (21.7) | .35[a] |
| Others (n=396) | 8 (2.0) | 3 (1.9) | 5 (2.1) | .91[a] |
| **Credibility score regarding dermatological information sources on social media, mean (SD)** | | | | |
| Family members (n=397) | 6.9 (2.3) | 7.1 (2.1) | 6.8 | .33[b] |
| Friends/relatives (n=392) | 6.5 (2.1) | 6.5 (2.1) | 6.5 | .60[b] |
| Celebrities (including famous medical professionals) (n=387) | 5.7 (2.4) | 5.7 (2.4) | 5.7 | .66[b] |
| Television programs/magazines (n=399) | 6.9 (2.3) | 7.1 (2.1) | 6.8 | .19[b] |
| Dermatological product retailers (n=386) | 4.8 (2.3) | 4.9 (2.3) | 4.8 | .65[b] |
| Dermatology companies (n=390) | 5.2 (2.4) | 5.1 (2.3) | 5.2 | .92[b] |

[a] Chi-squared test.

[b] Mann-Whitney test.

Table 3 outlines the social media engagement by dermatology patients. Actively searching for, actively sharing, and actively applying dermatological information were reported by 95.2% (380/399), 81.4% (323/397), and 85.4% (339/397) of the patients, respectively. Dermatological information on social media was perceived as useful or very useful by 57% (229/399) of the patients. The mean engagement score was moderate (mean 8.4, SD 2.4). The social media engagement score was significantly different between male and female patients ($P<.001$).

**Table 3.** Engagement with social media among dermatology patients, with comparisons between genders.

| Characteristics | Total sample | Men | Women | *P* value |
|---|---|---|---|---|
| **Actively searching for dermatological information on social media, n (%) (n=399)** | | | | |
| No | 19 (4.8) | 10 (6.3) | 9 (3.8) | .24[a] |
| Yes | 380 (95.2) | 149 (93.7) | 231 (96.3) | |
| **Actively sharing dermatological information on social media, n (%) (n=397)** | | | | |
| No | 74 (18.6) | 36 (22.8) | 38 (15.9) | .09[a] |
| Yes | 323 (81.4) | 122 (77.2) | 201 (84.1) | |
| **Actively applying dermatological information found on social media, n (%) (n=397)** | | | | |
| No | 58 (14.6) | 30 (18.9) | 28 (11.8) | .05[a] |
| Yes | 339 (85.4) | 129 (81.1) | 210 (88.2) | |
| **Perceived usefulness of dermatological information on social media, n (%) (n=399)** | | | | |
| Very useful | 47 (11.8) | 17 (10.8) | 30 (12.4) | .39[a] |
| Useful | 182 (45.6) | 73 (46.2) | 109 (45.2) | |
| Neutral | 154 (38.6) | 58 (36.7) | 96 (39.8) | |
| Not useful | 14 (3.5) | 9 (5.7) | 5 (2.1) | |
| Completely not useful | 2 (0.5) | 1 (0.6) | 1 (0.4) | |
| Social media engagement score, mean (SD) (n=399) | 8.4 (2.5) | 8.0 (2.7) | 8.7 (2.2) | .00[b] |

[a] Chi-squared test.

[b] Mann-Whitney test.

The results of three multivariable regression models are displayed in Table 4. Among dermatology patients, women were more likely to use social media (odds ratio [OR] 2.23, 95% CI 1.23-4.06), while older patients were less likely to use social media. Compared with male patients, female patients were also more likely to be interested in dermatological information on social media (OR 3.09, 95% CI 1.35-7.09) and have higher social media engagement scores (coefficient 0.68, 95% CI 0.17-1.18).

Respondents who were married/had a partner had a higher likelihood of being interested in dermatology information (OR 2.37, 95% CI 1.01-5.53). Regarding social media engagement, higher scores were present for patients using Instagram and who rated the credibility of “family members” and “dermatology companies” higher. Meanwhile, lower social media engagement scores were related with higher credibility scores for dermatological product retailers.

**Table 4.** Factors associated with social media use and engagement in multivariable regression models.

| Characteristics | Uses social media | | | Interest in any dermatological information on social media | | | Social media engagement score | | |
|---|---|---|---|---|---|---|---|---|---|
| | OR[a] | 95% CI | *P* value | OR | 95% CI | *P* value | Coefficient | 95% CI | *P* value |
| **Gender** | | | | | | | | | |
| Male | ref[b] | —[c] | — | ref | — | — | ref | — | — |
| Female | 2.23 | 1.23-4.06 | .008 | 3.09 | 1.35-7.09 | .008 | 0.68 | 0.17-1.18 | .009 |
| **Age** | | | | | | | | | |
| 18-30 years | ref | — | — | — | — | — | — | — | — |
| 31-40 years | 0.45 | 0.16-1.27 | .132 | — | — | — | — | — | — |
| 41-50 years | 0.07 | 0.03-0.19 | .000 | — | — | — | — | — | — |
| 51-60 years | 0.02 | 0.01-0.06 | .000 | — | — | — | — | — | — |
| >60 years | 0.01 | 0.00-0.04 | .000 | — | — | — | — | — | — |
| **Occupation** | | | | | | | | | |
| Unemployed | ref | — | — | — | — | — | — | — | — |
| Freelancer | 1.48 | 0.16-13.61 | .729 | — | — | — | — | — | — |
| White-collar worker | 1.68 | 0.18-15.73 | .648 | — | — | — | — | — | — |
| Blue-collar worker | 0.40 | 0.04-3.81 | .428 | — | — | — | — | — | — |
| Student | 0.85 | 0.07-9.81 | .894 | — | — | — | — | — | — |
| Other | 1.74 | 0.16-18.57 | .648 | — | — | — | — | — | — |
| **Location of residence** | | | | | | | | | |
| Urban | — | — | — | ref | — | — | — | — | — |
| Rural | — | — | — | 2.24 | 0.74-6.83 | .156 | — | — | — |
| **Marital status** | | | | | | | | | |
| Single | — | — | — | ref | — | — | — | — | — |
| Married | — | — | — | 2.37 | 1.01-5.53 | .047 | — | — | — |
| **Uses Facebook** | | | | | | | | | |
| No | — | — | — | ref | — | — | ref | — | — |
| Yes | — | — | — | 2.82 | 0.85-9.37 | .090 | 0.64 | –0.26-1.55 | .163 |
| **Uses Instagram** | | | | | | | | | |
| No | — | — | — | — | — | — | ref | — | — |
| Yes | — | — | — | — | — | — | 0.58 | 0.00-1.15 | .049 |
| **Credibility score regarding dermatological information sources on social media** | | | | | | | | | |
| Family members | — | — | — | — | — | — | 0.15 | 0.03-0.26 | .012 |
| Dermatological product retailers | — | — | — | — | — | — | –0.19 | –0.38-0.00 | .049 |
| Dermatology companies | — | — | — | — | — | — | 0.22 | 0.04-0.39 | .016 |

[a] Odds ratio.

[b] Reference group.

[c] Not applicable.

## Discussion

### Principal Findings

Our study found that, despite social media use by a high proportion of dermatology patients, utilizing this platform as a source of credible information for dermatological issues remains modest. Women were more likely to be interested in dermatological information on social media and be engaged with social media. The findings suggest that increased dermatologist involvement in contributing to online dermatological content and more effort to develop targeted, individualized information should be considered to take advantage of social media platforms.

Previous studies have documented the modest use of social media to find or share health-related content and particularly dermatology-related information. Despite reportedly high general social media engagement (85.0-99.3% of the participants reported regular access of at least one social media platform), only 19.0-31.7% of the participants reported accessing dermatology-related information [29,48]. We found that younger patients (aged 18-30 years) were significantly more likely to use social media. Young people in Asia are influenced more by social media than young people in Western countries because young people in Asia have a higher rate of smartphone use [49].

The higher percentage of use found in our study may reflect the increased popularity of social media in health care, due to technological advances and recent changes in the public perception of social media [50]. Our findings support the argument that there has been a desire among patients to use social media as an additional medium to traditional platforms for obtaining medical information [48]. Thus, social networking sites can be utilized as a platform to distribute educational information regarding dermatological issues for a wider reach of the population and at a potentially lower cost, especially in resource-poor settings in developing countries [51,52].

Regarding the sources of dermatological information on social media sites, our findings showed that the highest credibility scores were given to information received from family members, friends, televisions programs, and magazines. This popularity of informal information sources further highlights the lack of active contribution from dermatological health care providers and dermatologists to health information online. Dermatologists have expressed concern that patients might be misled by other social media indicators (eg, number of followers) without checking the educational background and clinical experience of information providers [53]. Increasing the involvement of medical professionals is crucial, as contributions by those with certified expertise would enhance the accuracy and reliability of information available online, reducing the possibility of incidents occurring as a result of misleading, inaccurate information [29,54]. Indeed, concerns over the quality of dermatology-related knowledge obtained online are a main barrier to enhancing the use of social media as a source of health information for health care service users [29]. Such concerns among service users may explain the medium level of engagement with social media for dermatological content reported in our study. Another possible reason for such low engagement may be participant concern about the privacy and security of personal data shared online [55,56]. Taking necessary measures to ensure the confidentiality of patient information over online platforms would also be important when attempting to boost the utilization of social media sites.

In addition, the sociodemographic characteristics of our participants were associated with their level of concern regarding dermatological information on social media and social media engagement. Women, married respondents, and Facebook users were more likely to be interested in dermatological issues on social media. These findings suggest that Facebook, in particular, could serve as the primary social media platform for disseminating dermatological information to targeted groups of online users. For example, this information could focus on dermatological issues that are most relevant to women and married individuals. Previous studies have argued the potential of Facebook for knowledge distribution, owing to its large user base, versatility in providing information in various forms, and capacity to allow for interactions and connections between users [57,58]. The positive association between engagement and the credibility score given to family members and dermatology companies as sources of social media information may also be partially explained by the ability to facilitate communications between sources and receivers of information: the more credible the sources, the more those sources would be used.

One of the implications that can be drawn from our findings is that the presence of dermatologists on social media platforms to provide official, scientific, evidence-based dermatological information should be increased. To facilitate this, the marketing capabilities of dermatology health facilities should be enhanced, so they can reach a wider population. And, patient-physician communication should be encouraged through social networking sites, notably Facebook, with careful consideration of privacy protection measures. The adoption of the Law on Cybersecurity in Vietnam provided stricter regulations on the content and transfer of data online and can be expected to impact the extent of online channel use by both patients and physicians. This should be taken into consideration when designing communication and education campaigns and programs. Regarding content, dermatological information should be created and distributed in an individually customized manner, targeting specific groups with relevant information.

### Limitations

This study has certain limitations. First, the self-reported nature of the questionnaire may have introduced recall bias. Second, although effort was made to enhance the diversity of participants by conducting research at a central dermatological hospital, the adoption of a convenient sampling technique and the fact that the study was conducted at a single hospital could affect the generalizability of our study. In addition, although the instrument we used to measure social media engagement has a theoretical background and provides valuable insights on the topic, it has not been officially validated. There are also several possible research considerations and directions not yet covered in our study that may be recommended for further studies. Further research may benefit from studying other online media platforms such as Google and YouTube that have considerable influence on dermatology patients in developed countries [59,60] or exploring the relationship between social media and cutaneous concerns associated with stigmatized medical conditions (eg, substance abuse). Moreover, further research may consider examining the effect of dermatological conditions on family quality of life, such as atopic dermatitis [61], as well as the impact of social media on caregivers of dermatology patients [62].

## Conclusion

In conclusion, this study found high levels of social media use among dermatology patients, but only a moderate level of utilization and credibility regarding the use of social media as a source of dermatological information. Sociodemographic characteristics were associated with dermatology-related social media use and engagement. The results of this study recommend enhancing the involvement of dermatologists on social media platforms, in terms of knowledge contribution through social media for both patients and the general public. In addition, more efforts should be given in developing individualized information that targets specific groups of dermatology patients.

## Acknowledgments

Our thanks to Vietnam National Hospital of Dermatology and Venereology that gave us permission and support to conduct this study. In addition, we are grateful to all the patients who gave us their time and willingness to participate in this study.

## Conflicts of Interest

None declared.

## References

1. Karimkhani C, Dellavalle RP, Coffeng LE, Flohr C, Hay RJ, Langan SM, et al. Global Skin Disease Morbidity and Mortality: An Update From the Global Burden of Disease Study 2013. JAMA Dermatol 2017 May 01;153(5):406-412 [FREE Full text] [doi: 10.1001/jamadermatol.2016.5538] [Medline: 28249066]
2. Schofield J, Fleming D, Grindlay D, Williams H. Skin conditions are the commonest new reason people present to general practitioners in England and Wales. Br J Dermatol 2011 Nov;165(5):1044-1050. [doi: 10.1111/j.1365-2133.2011.10464.x] [Medline: 21692764]
3. Awadalla F, Rosenbaum DA, Camacho F, Fleischer AB, Feldman SR. Dermatologic disease in family medicine. Fam Med 2008;40(7):507-511. [Medline: 18928078]
4. Kerr O, Tidman M, Walker J, Aldridge R, Benton E. The profile of dermatological problems in primary care. Clin Exp Dermatol 2010 Jun;35(4):380-383. [doi: 10.1111/j.1365-2230.2009.03586.x] [Medline: 19874334]
5. Hong J, Koo B, Koo J. The psychosocial and occupational impact of chronic skin disease. Dermatol Ther 2008;21(1):54-59. [doi: 10.1111/j.1529-8019.2008.00170.x] [Medline: 18318886]
6. Fortune DG, Richards HL, Main CJ, Griffiths CE. What patients with psoriasis believe about their condition. Journal of the American Academy of Dermatology 1998 Aug;39(2):196-201. [doi: 10.1016/s0190-9622(98)70074-x]
7. Zhang M, Cheow E, Ho CS, Ng BY, Ho R, Cheok CCS. Application of low-cost methodologies for mobile phone app development. JMIR Mhealth Uhealth 2014;2(4):e55 [FREE Full text] [doi: 10.2196/mhealth.3549] [Medline: 25491323]
8. Zhang MW, Ho RC, Loh A, Wing T, Wynne O, Chan SWC, et al. Current status of postnatal depression smartphone applications available on application stores: an information quality analysis. BMJ Open 2017 Nov 14;7(11):e015655 [FREE Full text] [doi: 10.1136/bmjopen-2016-015655] [Medline: 29138195]
9. Rajagopalan A, Shah P, Zhang MW, Ho RC. Digital Platforms in the Assessment and Monitoring of Patients with Bipolar Disorder. Brain Sci 2017 Nov 12;7(11) [FREE Full text] [doi: 10.3390/brainsci7110150] [Medline: 29137156]
10. Zhang MWB, Ho RCM, Cassin SE, Hawa R, Sockalingam S. Online and smartphone based cognitive behavioral therapy for bariatric surgery patients: Initial pilot study. Technol Health Care 2015;23(6):737-744. [doi: 10.3233/THC-151026] [Medline: 26409514]
11. Zhang MWB, Fang P, Ho RCM. Global outreach and user preferences of a smartphone application developed for drinkers. Technol Health Care 2016 Jul 27;24(4):495-501 [FREE Full text] [doi: 10.3233/THC-161143] [Medline: 26890229]
12. Tran BX, Le XTT, Nguyen PN, Le QNH, Mai HT, Nguyen HLT, et al. Feasibility of e-Health Interventions on Smoking Cessation among Vietnamese Active Internet Users. Int J Environ Res Public Health 2018 Dec 20;15(1) [FREE Full text] [doi: 10.3390/ijerph15010165] [Medline: 29361694]
13. Zhang MWB, Ho RCM. Harnessing the potential of the Kinect sensor for psychiatric rehabilitation for stroke survivors. Technol Health Care 2016 Dec 04;24(4):599-602 [FREE Full text] [doi: 10.3233/THC-161147] [Medline: 27061386]
14. Zhang MWB, Ho CSH, Fang P, Lu Y, Ho RCM. Methodology of developing a smartphone application for crisis research and its clinical application. Technol Health Care 2014;22(4):547-559. [doi: 10.3233/THC-140819] [Medline: 24898865]
15. Leow MQH, Chan SWC. Evaluation of a video, telephone follow-ups, and an online forum as components of a psychoeducational intervention for caregivers of persons with advanced cancer. Palliat Support Care 2016 Dec;14(5):474-478 [FREE Full text] [doi: 10.1017/S1478951516000225] [Medline: 27071801]

16. Zhang MWB, Chan S, Wynne O, Jeong S, Hunter S, Wilson A, et al. Conceptualization of an evidence-based smartphone innovation for caregivers and persons living with dementia. Technol Health Care 2016 Sep 14;24(5):769-773. [doi: 10.3233/THC-161165] [Medline: 27129032]
17. Ho K, Peter WWP. Harnessing the social web for health and wellness: issues for research and knowledge translation. J Med Internet Res 2014;16(2):e34 [FREE Full text] [doi: 10.2196/jmir.2969] [Medline: 24518432]
18. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]
19. George DR, Rovniak LS, Kraschnewski JL. Dangers and opportunities for social media in medicine. Clin Obstet Gynecol 2013 Sep;56(3):453-462 [FREE Full text] [doi: 10.1097/GRF.0b013e318297dc38] [Medline: 23903375]
20. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. P T 2014 Jul;39(7):491-520 [FREE Full text] [Medline: 25083128]
21. Househ M, Borycki E, Kushniruk A. Empowering patients through social media: the benefits and challenges. Health Informatics J 2014 Mar;20(1):50-58. [doi: 10.1177/1460458213476969] [Medline: 24550564]
22. Grajales FJ, Sheps S, Ho K, Novak-Lauscher H, Eysenbach G. Social media: a review and tutorial of applications in medicine and health care. J Med Internet Res 2014;16(2):e13 [FREE Full text] [doi: 10.2196/jmir.2912] [Medline: 24518354]
23. We Are Social. 2018 Nov 19. Digital in 2018: World's Internet Users Pass the 4 Billion Mark URL: https://wearesocial.com/uk/blog/2018/01/global-digital-report-2018 [accessed 2020-02-02]
24. Travers RL. Social media in dermatology: moving to Web 2.0. Semin Cutan Med Surg 2012 Sep;31(3):168-173 [FREE Full text] [doi: 10.1016/j.sder.2012.06.003] [Medline: 22929354]
25. Rupert DJ, Moultrie RR, Read JG, Amoozegar JB, Bornkessel AS, O'Donoghue AC, et al. Perceived healthcare provider reactions to patient and caregiver use of online health communities. Patient Educ Couns 2014 Sep;96(3):320-326. [doi: 10.1016/j.pec.2014.05.015] [Medline: 24923652]
26. Ho Y, O'Connor BH, Mulvaney SA. Features of online health communities for adolescents with type 1 diabetes. West J Nurs Res 2014 Oct;36(9):1183-1198. [doi: 10.1177/0193945913520414] [Medline: 24473058]
27. Smailhodzic E, Hooijsma W, Boonstra A, Langley DJ. Social media use in healthcare: A systematic review of effects on patients and on their relationship with healthcare professionals. BMC Health Serv Res 2016 Aug 26;16:442 [FREE Full text] [doi: 10.1186/s12913-016-1691-0] [Medline: 27562728]
28. Hawn C. Take two aspirin and tweet me in the morning: how Twitter, Facebook, and other social media are reshaping health care. Health Aff (Millwood) 2009;28(2):361-368 [FREE Full text] [doi: 10.1377/hlthaff.28.2.361] [Medline: 19275991]
29. Antheunis ML, Tates K, Nieboer TE. Patients' and health professionals' use of social media in health care: motives, barriers and expectations. Patient Educ Couns 2013 Sep;92(3):426-431. [doi: 10.1016/j.pec.2013.06.020] [Medline: 23899831]
30. Hardiker NR, Grant MJ. Factors that influence public engagement with eHealth: A literature review. Int J Med Inform 2011 Jan;80(1):1-12. [doi: 10.1016/j.ijmedinf.2010.10.017] [Medline: 21112244]
31. Bosslet GT, Torke AM, Hickman SE, Terry CL, Helft PR. The patient-doctor relationship and online social networks: results of a national survey. J Gen Intern Med 2011 Oct;26(10):1168-1174 [FREE Full text] [doi: 10.1007/s11606-011-1761-2] [Medline: 21706268]
32. D'Souza K, Henningham L, Zou R, Huang J, O'Sullivan E, Last J, et al. Attitudes of Health Professional Educators Toward the Use of Social Media as a Teaching Tool: Global Cross-Sectional Study. JMIR Med Educ 2017 Aug 04;3(2):e13 [FREE Full text] [doi: 10.2196/mededu.6429] [Medline: 28778841]
33. Frisch NC, Borycki EM, Mickelson G, Atherton P, Novak-Lauscher H, Hooker D, et al. Use of social media and web 2.0 technologies to increase knowledge and skills of british columbia nurses. NI 2012 (2012) 2012;2012:117 [FREE Full text] [Medline: 24199067]
34. Wang JV, O'Donnell M, Albornoz CA, Keller M, Saedi N. Resident experiences with social media: modernizing dermatology training. Int J Dermatol 2018 Dec;57(12):e169-e170 [FREE Full text] [doi: 10.1111/ijd.14182] [Medline: 30101440]
35. Zhang MWB, Ho RCM, Hawa R, Sockalingam S. Analysis of the Information Quality of Bariatric Surgery Smartphone Applications Using the Silberg Scale. Obes Surg 2016 Jan;26(1):163-168. [doi: 10.1007/s11695-015-1890-5] [Medline: 26424704]
36. DeBord LC, Patel V, Braun TL, Dao H. Social media in dermatology: clinical relevance, academic value, and trends across platforms. J Dermatolog Treat 2018 Sep 28:1-30 [FREE Full text] [doi: 10.1080/09546634.2018.1530444] [Medline: 30265614]
37. Tran BX, Huong LT, Hinh ND, Nguyen LH, Le BN, Nong VM, et al. A study on the influence of internet addiction and online interpersonal influences on health-related quality of life in young Vietnamese. BMC Public Health 2017 Jan 31;17(1):138 [FREE Full text] [doi: 10.1186/s12889-016-3983-z] [Medline: 28143462]
38. Tran BX, Mai HT, Nguyen LH, Nguyen CT, Latkin CA, Zhang MWB, et al. Vietnamese validation of the short version of Internet Addiction Test. Addict Behav Rep 2017 Dec;6:45-50 [FREE Full text] [doi: 10.1016/j.abrep.2017.07.001] [Medline: 29450235]
39. Karimkhani C, Connett J, Boyers L, Quest T, Dellavalle RP. Dermatology on instagram. Dermatol Online J 2014 Jul 15;20(7) [FREE Full text] [Medline: 25046455]

40. Amir M, Sampson BP, Endly D, Tamai JM, Henley J, Brewer AC, et al. Social networking sites: emerging and essential tools for communication in dermatology. JAMA Dermatol 2014 Jan;150(1):56-60. [doi: 10.1001/jamadermatol.2013.6340] [Medline: 24196212]
41. Benabio J. The value of social media for dermatologists. Cutis 2013 Jun;91(6):269-270. [Medline: 23837146]
42. Patel RR, Yazd NKK, Dellavalle RP. Dermatology on Snapchat. Dermatol Online J 2017 Jul 15;23(7) [FREE Full text] [Medline: 29469694]
43. Falzone AE, Brindis CD, Chren M, Junn A, Pagoto S, Wehner M, et al. Teens, Tweets, and Tanning Beds: Rethinking the Use of Social Media for Skin Cancer Prevention. Am J Prev Med 2017 Sep;53(3S1):S86-S94 [FREE Full text] [doi: 10.1016/j.amepre.2017.04.027] [Medline: 28818251]
44. Balieva F, Kupfer J, Lien L, Gieler U, Finlay AY, Tomás-Aragonés L, et al. The burden of common skin diseases assessed with the EQ5D™: a European multicentre study in 13 countries. Br J Dermatol 2017 May;176(5):1170-1178. [doi: 10.1111/bjd.15280] [Medline: 28032340]
45. Nguyen SH, Dang AK, Vu GT, Nguyen CT, Le THT, Truong NT, et al. Lack of Knowledge about Sexually Transmitted Diseases (STDs): Implications for STDs Prevention and Care among Dermatology Patients in an Urban City in Vietnam. Int J Environ Res Public Health 2019 Mar 26;16(6) [FREE Full text] [doi: 10.3390/ijerph16061080] [Medline: 30917565]
46. Di Gangi PM, Wasko M. Social Media Engagement Theory: Exploring the Influence of User Engagement on Social Media Usage. JOEUC 2016;28(2):53-73. [doi: 10.4018/joeuc.2016040104]
47. Voorveld HAM, van Noort G, Muntinga DG, Bronner F. Engagement with Social Media and Social Media Advertising: The Differentiating Role of Platform Type. Journal of Advertising 2018 Feb 13;47(1):38-54. [doi: 10.1080/00913367.2017.1405754]
48. Ross NA, Todd Q, Saedi N. Patient seeking behaviors and online personas: social media's role in cosmetic dermatology. Dermatol Surg 2015 Feb;41(2):269-276 [FREE Full text] [doi: 10.1097/DSS.0000000000000267] [Medline: 25654198]
49. Mak K, Lai C, Watanabe H, Kim D, Bahar N, Ramos M, et al. Epidemiology of internet behaviors and addiction among adolescents in six Asian countries. Cyberpsychol Behav Soc Netw 2014 Nov;17(11):720-728. [doi: 10.1089/cyber.2014.0139] [Medline: 25405785]
50. Russell J. Hootsuite Social Media Management. 2017 Jan 31. Social Media in Health Care: The Benefits, Challenges, and Opportunities URL: http://www.marketinghub.today/social-media-in-health-care-the-benefits-challenges-and-opportunities/ [accessed 2020-02-02]
51. Hanson AH, Krause LK, Simmons RN, Ellis JI, Gamble RG, Jensen JD, et al. Dermatology education and the Internet: traditional and cutting-edge resources. J Am Acad Dermatol 2011 Oct;65(4):836-842 [FREE Full text] [doi: 10.1016/j.jaad.2010.05.049] [Medline: 21820206]
52. Hay AA, Gamble RG, Huff LS, Dellavalle RP. Internet social networking sites and the future of dermatology journals: promises and perils. J Am Acad Dermatol 2011 Sep;65(3):e81-e83 [FREE Full text] [doi: 10.1016/j.jaad.2011.04.028] [Medline: 21839304]
53. Galadari HI. Social media and modern dermatology. Int J Dermatol 2018 Dec;57(1):110-111 [FREE Full text] [doi: 10.1111/ijd.13821] [Medline: 29090452]
54. George DD, Wainwright BD. Dermatology resources on the internet. Semin Cutan Med Surg 2012 Sep;31(3):183-190. [doi: 10.1016/j.sder.2012.06.002] [Medline: 22929356]
55. Thompson LA, Dawson K, Ferdig R, Black EW, Boyer J, Coutts J, et al. The intersection of online social networking with medical professionalism. J Gen Intern Med 2008 Jul;23(7):954-957 [FREE Full text] [doi: 10.1007/s11606-008-0538-8] [Medline: 18612723]
56. Thompson MA, Younes A, Miller RS. Using social media in oncology for education and patient engagement. Oncology (Williston Park) 2012 Sep;26(9):782, 784-785, 791 [FREE Full text] [Medline: 23061330]
57. Kim WB, Marinas JEC, Vender RB. Public Engagement with Dermatology Contents on Facebook. J Cutan Med Surg 2015;19(3):304-308 [FREE Full text] [doi: 10.2310/7750.2014.14119] [Medline: 25775650]
58. Kim W, Vender R. Use of facebook as a tool for knowledge dissemination in dermatology. J Cutan Med Surg 2014 Oct;18(5):341-344 [FREE Full text] [doi: 10.2310/7750.2014.14022] [Medline: 25186996]
59. Hill MK, Patel RR, Anand P, Dellavalle RP. Dermatology on Google. Dermatol Online J 2018 Apr 15;24(4) [FREE Full text] [Medline: 29906016]
60. Boyers LN, Quest T, Karimkhani C, Connett J, Dellavalle RP. Dermatology on YouTube. Dermatol Online J 2014 Jun 15;20(6) [FREE Full text] [Medline: 24945641]
61. Chernyshov PV, Ho RC, Monti F, Jirakova A, Velitchko SS, Hercogova J, et al. An International Multi-center Study on Self-assessed and Family Quality of Life in Children with Atopic Dermatitis. Acta Dermatovenerol Croat 2015;23(4):247-253. [Medline: 26724875]
62. Ho RCM, Giam YC, Ng TP, Mak A, Goh D, Zhang MWB, et al. The influence of childhood atopic dermatitis on health of mothers, and its impact on Asian families. Pediatr Allergy Immunol 2010 May;21(3):501-507 [FREE Full text] [doi: 10.1111/j.1399-3038.2009.00972.x] [Medline: 20546527]

## Abbreviations

**NHDV:** National Hospital of Dermatology and Venereology.
**OR:** odds ratio.

---

*Edited by G Eysenbach; submitted 18.01.19; peer-reviewed by HP Do, M El Tantawi, J Xu; comments to author 02.10.19; revised version received 25.11.19; accepted 26.01.20; published 23.03.20.*

*Please cite as:*
*Nguyen SH, Vu GT, Nguyen LH, Nguyen CT, Le THT, Tran TH, Tran BX, Latkin CA, Tam WWS, Ho CSH, Ho RCM*
*Understanding Social Media Use and Engagement Among Dermatology Patients to Inform Dermatological Prevention and Care in Vietnam: Cross-sectional Study*
*JMIR Dermatol 2020;3(1):e13424*
*URL: http://derma.jmir.org/2020/1/e13424/*
*doi:10.2196/13424*
*PMID:*

Original Paper

# Assessment of Altmetrics and PlumX Metrics Scoring as Mechanisms to Evaluate the Top 100 Trending Hidradenitis Suppurativa Articles on Social Media: Cross-Sectional Study

Chapman Wei[1], BSc; Aaron Fong[2], MD; Theodore Quan[1], BSc; Puneet Gupta[1], BSc; Adam Friedman[2], MD, FAAD

[1]George Washington University School of Medicine and Health Sciences, Washington, DC, United States

[2]Department of Dermatology, George Washington University School of Medicine and Health Sciences, Washington, DC, United States

**Corresponding Author:**
Adam Friedman, MD, FAAD
Department of Dermatology
George Washington University School of Medicine and Health Sciences
2150 Pennsylvania Avenue, 2B-430
Washington, DC, 20037
United States
Phone: 1 202 741 2625
Email: ajfriedman@mfa.gwu.edu

## Abstract

**Background:** Dermatologists are increasingly utilizing social media platforms to disseminate scientific information. New tools, such as altmetrics and PlumX metrics, have been made available to rapidly capture the level of scientific article dissemination across social media platforms. However, no studies have been performed to assess the level of scientific article dissemination across social media regarding hidradenitis suppurativa, a disease that is still currently not well understood.

**Objective:** The aim of our study was to evaluate the utility of altmetrics and PlumX metrics by characterizing the top 100 “trending” hidradenitis suppurativa articles in the altmetric database by the altmetric attention score and PlumX score.

**Methods:** Altmetric data components of the top 100 hidradenitis suppurativa articles were extracted from the altmetric database. Article citation count was found using Web of Science. PlumX field-weighted impact scores for each article were collected from the Scopus database. Journal title, open-access status, article type, and study design of original articles were assessed. Additionally, the altmetric attention score, PlumX score, and citation count were log transformed and adjusted by +1 for linear regression, and Spearman correlation coefficients were utilized to determine correlations.

**Results:** Most of the top 100 “trending” hidradenitis suppurativa articles were published in *JAMA Dermatology* (n=27, 27%). The median altmetric attention score, PlumX score, and citation count were 25.5, 3.7, and 10.5, respectively. The most mentions regarding social media platforms came from Twitter. Although no correlation was observed between the citation count and altmetric attention score ($r^2$=0.019, $P$=.17), positive correlation was observed between the citation count and PlumX score ($r^2$=0.469, $P$<.001).

**Conclusions:** Our research demonstrated that citation count is not correlated with the altmetric attention score, but is strongly correlated with the PlumX score regarding hidradenitis suppurativa articles at this point in time. With the continual increase of social media usage by medical professionals and researchers, this study can help investigators understand the best way to captivate their audience.

***(JMIR Dermatol 2020;3(1):e23724)*** doi:10.2196/23724

**KEYWORDS**

altmetric; PlumX; social media; impact factor; hidradenitis suppurativa

## Introduction

Dissemination of science-based peer-reviewed information is essential for increasing awareness of hidradenitis suppurativa, an inflammatory skin disease that is often underrecognized, leading to delayed diagnosis and higher severity at initial presentation [1]. Dermatologists are increasingly utilizing social media platforms to disseminate scientific information [2]. In contrast to medical journals, social media serves as a useful platform to inform the wider general public, both medical professionals and laypeople alike. New tools, including alternative metrics (altmetrics) and PlumX metrics, have been made available to evaluate the extent of an article's dissemination across social media platforms [3,4]. Provided by altmetrics, the altmetric attention score is a weighted score of the amount of "online attention" a research article has received across social media platforms. These platforms include Facebook, Twitter, Google+, blogs, and others [5]. Similarly, PlumX has a PlumX field-weighted citation impact score (PlumX score) that is also a weighted score of the level of article dissemination across similar social media metrics and includes the number of citations, linkouts, and abstract views [6]. While the citation count or "impact factor" reflects the number of citations by other articles or journals, altmetrics and PlumX metrics reflect an article's instantaneous attentiveness among news outlets, blogs, Twitter, Facebook, and other media platforms [7]. While these tools allow one to assess what kind of articles are garnering social media attention or "trending" on social media platforms, the assessment of the utility of altmetrics regarding hidradenitis suppurativa research has not been evaluated yet. Owing to the ability of these metrics to rapidly capture the level of scientific article dissemination, they have the potential of being used complementary with citation count and identifying high-impact hidradenitis suppurativa articles, since the traditional tool takes years to measure an article's impact on field advancement. To evaluate the utility of these tools, we characterized the top 100 "trending" hidradenitis suppurativa articles across social media platforms captured in the altmetric database by altmetric attention scores complemented with PlumX scores and more established markers of manuscript value. Our study aimed to assess the social media platforms that most contribute to the dissemination of literature-based hidradenitis suppurativa information. We also sought to examine how alternative metrics compare with traditional metrics, such as citation count. We hypothesized that alternative metrics will be able to better complement more traditional metrics when evaluating an article's quality.

## Methods

Utilizing Altmetric Explorer, we identified articles using the PubMed search criterion "hidradenitis suppurativa." Altmetric data components of the top 100 altmetric attention score hidradenitis suppurativa articles were extracted in July 2020. The number of mentions from the following altmetric data components was extracted and examined: news mentions, blog mentions, policy mentions, Twitter mentions, Facebook mentions, Wikipedia mentions, Reddit mentions, Mendeley readers, and number of Dimension citations [6]. Mentions represent a measure of how many times a specific term, such as "hidradenitis suppurative," has been referenced on social media channels. Dimension citations include grants, publications, citations, alternative metrics, clinical trials, and patents. Data were also collected about the journal impact factor, journal title, open-access status, where the articles were published, article type, article study design, and article citation count [4,5]. PlumX field-weighted citation impact scores for each article were collected from the Scopus database. Statistical analysis was performed using Prism 8 (GraphPad software), and statistical significance was defined as a *P* value <.05. Spearman correlation coefficients were utilized to determine correlations [2]. Altmetric attention scores, PlumX field-weighted citation impact scores, and citation count were log transformed and adjusted by +1 for linear regression.

## Results

Of the top 100 "trending" hidradenitis suppurativa articles by highest altmetric attention score, most were published in *JAMA Dermatology* (n=27, 27%), followed by *Journal of the American Academy of Dermatology* (n=16, 16%) (Table 1). The median altmetric attention score, PlumX score, and citation count were 25.5, 3.7, and 10.5, respectively. Majority of the "trending" hidradenitis suppurativa research articles were published by authors from Europe (n=47). The median journal impact factor was 6.9. The highest altmetric attention score article was a review article (altmetric attention score=352) [8]. This review article discussed the diagnosis, epidemiology, and treatment of hidradenitis suppurativa, with specific emphasis on advances in the past 5 years [8]. A total of 58 articles had an altmetric attention score above 20, a marker for the top 5% of all scientific output [4]. Of the social media platforms, the most mentions came from Twitter (2830 mentions), which was mostly used to share European-affiliated articles. The second most mentions came from news (573 mentions), which was mostly used to share North American–affiliated articles. The number of Mendeley readers and Dimension citations were 5130 and 4152, respectively. Most of the articles were original articles (53 articles). A total of 49 articles were open access. The most common study design utilized was the clinical observational design (34 articles). A correlation was observed between the altmetric attention score and journal impact factor ($r^2$=0.17, *P*<.001). Although no correlation was observed between citation count and the altmetric attention score ($r^2$=0.019, *P*=.17; Figure 1), a positive correlation was observed between citation count and the PlumX score ($r^2$=0.469, *P*<.001). The altmetric attention score was further found not to be correlated with citation count for particular article types in a chi-square analysis (*P*=.95). Open-access status did not affect the altmetric attention score (*P*=.71). Out of the 100 articles, most of the articles were published in 2017 (19 articles), followed by 18 articles published in 2019 (Figure 2).

**Table 1.** Characteristics of the top 100 “trending” hidradenitis suppurativa articles by the altmetric score.

| Characteristic | Value (N=100) |
|---|---|
| **Journal, n** | |
| *JAMA Dermatology* | 27 |
| *Journal of the American Academy of Dermatology* | 16 |
| *British Journal of Dermatology* | 5 |
| *Journal of Investigative Dermatology* | 4 |
| Other journals | 48 |
| Altmetric score, median (range) | 25.5 (13-352) |
| PlumX field score, median (range) | 3.7 (0.0-31.8) |
| Journal impact factor, median (range) | 6.9 (0.2-70.7) |
| Traditional citation, median (range) | 10.5 (0-389) |
| News mentions, total (range) | 573 (0-109) |
| Policy mentions, total (range) | 17 (0-3) |
| Twitter mentions, total (range) | 2830 (0-706) |
| Facebook mentions, total (range) | 190 (0-24) |
| Number of Mendeley readers, total (range) | 5130 (0-245) |
| Number of Dimension citations, total (range) | 4152 (0-418) |
| **Article type, n** | |
| Original investigation | 53 |
| Research letter/comments to the Editor/brief report | 6 |
| Review | 30 |
| Editorial | 4 |
| Viewpoint/clinical pearls | 2 |
| Guidelines/specific statement | 2 |
| Other | 3 |
| Open access | 49 |
| **Study design, n** | |
| Case report/series | 6 |
| Clinical observation | 34 |
| Clinical trial | 5 |
| Basic science | 7 |
| **Region of the article’s authors, n** | |
| Europe | 47 |
| North America | 37 |
| Other | 16 |

**Figure 1.** Citation count correlation with the altmetric attention score (AAS) and PlumX score.

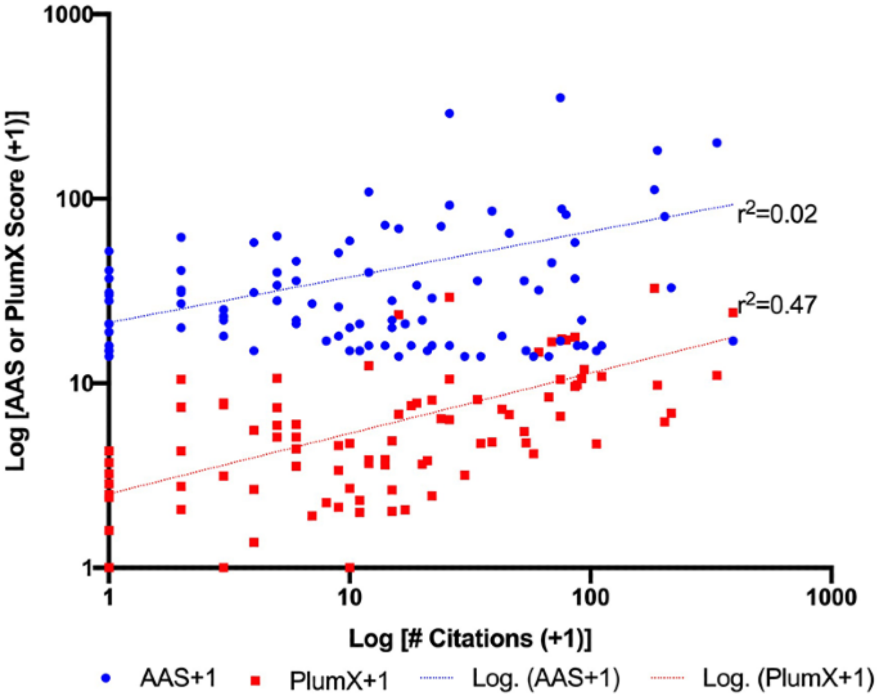
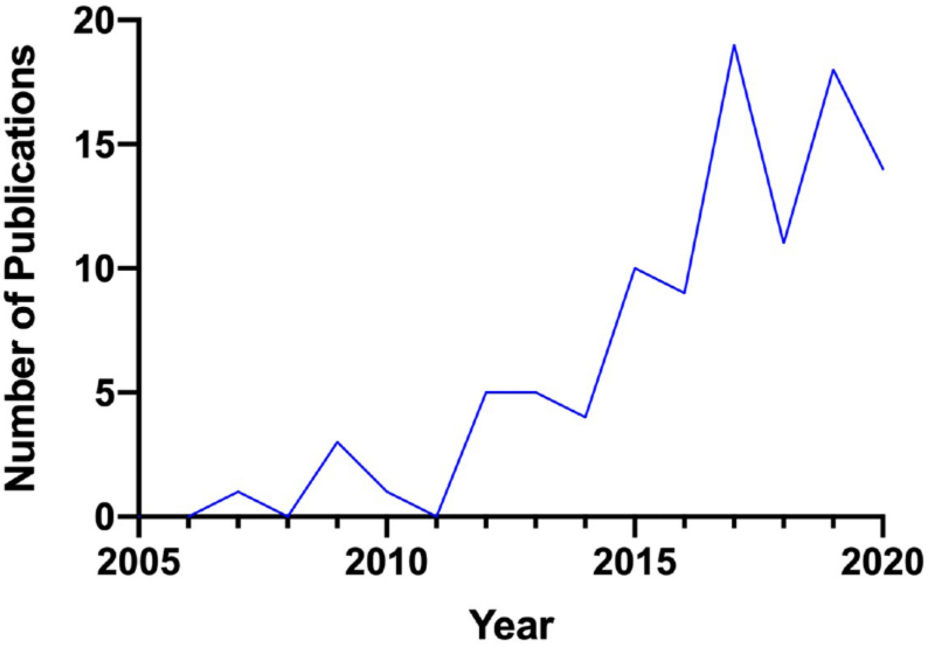


**Figure 2.** Year of publication for the included articles.



## Discussion

### Principal Findings

This study determined the characteristics of the top 100 "trending" hidradenitis suppurativa articles that received the greatest amount of social media attention captured in the altmetric database. Twitter was the most widely used social media platform when mentioning hidradenitis suppurativa. The article with the highest altmetric attention score, a review paper with European affiliation, achieved this high score likely because this article not only summarized the current knowledge about hidradenitis suppurativa, but also went on to provide new insights into the advances in hidradenitis suppurativa research with clear recommendations [8]. Open access, which is often an added cost to authors, did not contribute to a higher altmetric attention score. This likely occurred owing to the decline of the open-access advantage from the ease of article redistribution [9].

### Comparison With Prior Work

Currently, there are no published studies evaluating the utilization of altmetrics with regard to the hidradenitis suppurativa literature. There are however multiple past altmetric studies that investigated top “trending” articles in the altmetric database regarding the correlation between traditional bibliometrics, such as citation count, and online attention, specifically the altmetric attention score. These studies included research articles about specific conditions, such as stroke and oral cancer, and also about specific medical fields, such as urology, emergency medicine, and plastic surgery [10-14]. In our study, altmetrics did not correlate with citation count in hidradenitis suppurativa literature despite its ability to rapidly assess how widely an article is disseminated, while the PlumX score did show correlation. Variable results have been reported regarding the utility of altmetrics [2,4,10-14]. While some articles reported findings similar to those in our study [10,12,13], others only showed a modest correlation at best between citation count and the altmetric attention score [2,4,11,14]. This is likely because the altmetric attention score is more dynamic than the PlumX score since the altmetric attention score focuses on tracking of real-time public interest in a particular topic. Thus, the altmetric attention score does not factor in citation count. In contrast, the PlumX score is likely a better marker because PlumX metrics factor in citation count, suggesting that PlumX metrics may be more useful in identifying high-impact hidradenitis suppurativa articles.

### Limitations

There are several limitations with this study. This study was a cross-sectional study performed in July 2020, so altmetric and PlumX scores may change in the future. Altmetric and PlumX scores should be cautiously interpreted as these alternative metrics do not reflect article quality. Thus, a research article that receives a wide amount of social media attention should not be interpreted as having important results and should be examined in conjunction with in-depth article analysis to determine research quality. Additionally, since the article sample utilized in this study represented a minority of published articles on hidradenitis suppurativa, the generalizability of the findings may be limited regarding hidradenitis suppurativa literature. However, our analysis assessed all hidradenitis suppurativa articles that were in the top 5% of scientific output in the altmetrics database. More recently published hidradenitis suppurativa articles can experience a delay in their citation count compared with older articles. Lastly, our analysis did not assess the presence of journal Twitter accounts or whether the journals were active on Twitter. It is possible that journals having a larger social media presence may bolster their hidradenitis suppurativa articles with higher altmetric scores.

### Future Directions

The correlation of the PlumX score and the altmetric attention score with citation count for other dermatology topics remains to be explored. Further research into identifying specific traits of hidradenitis suppurativa articles with a higher PlumX score or altmetric attention score is warranted. Given the increasing usage of social media by medical professionals and researchers, such research can be useful to investigators by helping them understand the best way to maximize their reach, including the general public. However, as is suggested by our study, it must be understood that the level of research dissemination across social media does not necessarily translate to an impact in the scientific community.

### Conclusion

Since altmetrics and PlumX metrics are dynamic reflections of the general social media interest, there is still some discordance between scientific articles and social media. Although these metrics can identify the impact and dissemination to the public, they do not measure the stringent review process that articles undergo for publication in scientific journals. Despite these limitations, both metrics can be used complementary with traditional citation analysis to assess article quality. Altmetrics with PlumX may be used to rapidly capture what the general public is interested in regarding hidradenitis suppurativa, while traditional metrics can be used to assess an article’s impact.

## Conflicts of Interest

None declared.

## References

1. Dufour DN, Emtestam L, Jemec GB. Hidradenitis suppurativa: a common and burdensome, yet under-recognised, inflammatory skin disease. Postgrad Med J 2014 Apr;90(1062):216-21; quiz 220 [FREE Full text] [doi: 10.1136/postgradmedj-2013-131994] [Medline: 24567417]
2. Jia J, Nguyen B, Mills D, Polin DJ, Sarin KY. Comparing online engagement and academic impact of dermatology research: An Altmetric Attention Score and PlumX Metrics analysis. J Am Acad Dermatol 2020 Aug;83(2):648-650 [FREE Full text] [doi: 10.1016/j.jaad.2019.12.003] [Medline: 31836559]
3. Peters I, Kraker P, Lex E, Gumpenberger C, Gorraiz J. Research data explored: an extended analysis of citations and altmetrics. Scientometrics 2016;107:723-744 [FREE Full text] [doi: 10.1007/s11192-016-1887-4] [Medline: 27122647]
4. Patel RB, Vaduganathan M, Bhatt DL, Bonow RO. Characterizing High-Performing Articles by Altmetric Score in Major Cardiovascular Journals. JAMA Cardiol 2018 Dec 01;3(12):1249-1251 [FREE Full text] [doi: 10.1001/jamacardio.2018.3823] [Medline: 30476952]
5. Elmore SA. The Altmetric Attention Score: What Does It Mean and Why Should I Care? Toxicol Pathol 2018 Apr;46(3):252-255 [FREE Full text] [doi: 10.1177/0192623318758294] [Medline: 29448902]

6. Ortega J. Disciplinary differences of the impact of altmetric. FEMS Microbiol Lett 2018 Apr 01;365(7). [doi: 10.1093/femsle/fny049] [Medline: 29518193]
7. Patthi B, Prasad M, Gupta R, Singla A, Kumar JK, Dhama K, et al. Altmetrics - A Collated Adjunct Beyond Citations for Scholarly Impact: A Systematic Review. J Clin Diagn Res 2017 Jun;11(6):ZE16-ZE20 [FREE Full text] [doi: 10.7860/JCDR/2017/26153.10078] [Medline: 28764311]
8. Saunte DM, Jemec GB. Hidradenitis Suppurativa: Advances in Diagnosis and Treatment. JAMA 2017 Nov 28;318(20):2019-2032. [doi: 10.1001/jama.2017.16691] [Medline: 29183082]
9. Davis PM. Author-choice open-access publishing in the biological and medical literature: A citation analysis. J. Am. Soc. Inf. Sci 2009 Jan;60(1):3-8. [doi: 10.1002/asi.20965]
10. Asaad M, Howell S, Rajesh A, Meaike J, Tran NV. Altmetrics in Plastic Surgery Journals: Does It Correlate With Citation Count? Aesthet Surg J 2020 Oct 24;40(11):NP628-NP635. [doi: 10.1093/asj/sjaa158] [Medline: 32506129]
11. Nocera AP, Boyd CJ, Boudreau H, Hakim O, Rais-Bahrami S. Examining the Correlation Between Altmetric Score and Citations in the Urology Literature. Urology 2019 Dec;134:45-50. [doi: 10.1016/j.urology.2019.09.014] [Medline: 31560915]
12. Kim Y, Kim J, Kim YH, Yoon DY, Kim YJ, Bae JS. Social attention and scientific articles on stroke: Altmetric analysis of top-50 articles. Clin Neurol Neurosurg 2019 Aug;183:105386. [doi: 10.1016/j.clineuro.2019.105386] [Medline: 31207458]
13. Hassona Y, Qutachi T, Dardas L, Alrashdan MS, Sawair F. The online attention to oral cancer research: An Altmetric analysis. Oral Dis 2019 Sep;25(6):1502-1510. [doi: 10.1111/odi.13111] [Medline: 31033085]
14. Barbic D, Tubman M, Lam H, Barbic S. An Analysis of Altmetrics in Emergency Medicine. Acad Emerg Med 2016 Mar;23(3):251-268 [FREE Full text] [doi: 10.1111/acem.12898] [Medline: 26743680]

---

*Edited by G Eysenbach; submitted 20.08.20; peer-reviewed by A Gu, V Kotha, S Oska; comments to author 11.09.20; revised version received 06.10.20; accepted 28.10.20; published 19.11.20.*

*Please cite as:*
*Wei C, Fong A, Quan T, Gupta P, Friedman A*
*Assessment of Altmetrics and PlumX Metrics Scoring as Mechanisms to Evaluate the Top 100 Trending Hidradenitis Suppurativa Articles on Social Media: Cross-Sectional Study*
*JMIR Dermatol 2020;3(1):e23724*
*URL: http://derma.jmir.org/2020/1/e23724/*
*doi:10.2196/23724*
*PMID:*

Viewpoint

# LesionMap: A Method and Tool for the Semantic Annotation of Dermatological Lesions for Documentation and Machine Learning

Bell Raj Eapen[1], MD, MSc; Norm Archer[1], PhD; Kamran Sartipi[2], PhD

[1]Information Systems, McMaster University, Hamilton, ON, Canada

[2]Department of Computer Science, East Carolina University, Greenville, NC, United States

**Corresponding Author:**
Bell Raj Eapen, MD, MSc
Information Systems
McMaster University
DSB - 211
1280 Main Street West
Hamilton, ON, L8S 4L8
Canada
Phone: 1 905 525 9140 ext 26392
Email: eapenbp@mcmaster.ca

## Abstract

Diagnosis and follow-up of patients in dermatology rely on visual cues. Documentation of skin lesions in dermatology is time-consuming and inaccurate. Digital photography is resource-intensive, difficult to standardize, and has privacy concerns. We propose a simple method—LesionMap—and an electronic health software tool—LesionMapper—for semantically annotating dermatological lesions on a body wireframe. We discuss how the type, distribution, and progression of lesions can be represented in a standardized way. The tool is an open-source JavaScript package that can be integrated into web-based electronic medical records. We believe that LesionMapper will facilitate documentation in dermatology that can be used for machine learning in a privacy-preserving manner.

***(JMIR Dermatol 2020;3(1):e18149)*** doi:10.2196/18149

**KEYWORDS**

LesionMap; LesionMapper; digital imaging; machine learning; dermatology

## Introduction

Documenting the origin, distribution, and nature of dermatological lesions in a textual form is inefficient and imprecise. Dermatologists often document the images of the patient or draw the lesions on a body wireframe for later reference. Digital photography for clinical documentation is time-consuming and resource intensive to capture, organize, and maintain [1]. Additionally, there is a growing privacy-related concern over the use of these images [2].

Capturing a detailed account of dermatological lesions in a privacy-preserving way is becoming increasingly important in the era of machine learning and artificial intelligence (AI). Documentation in electronic medical records (EMRs) requires a simple and efficient tool that fits into the clinical workflow. There is a growing need for a standardized methodology and an annotation schema to facilitate the capture of rich data related to dermatological conditions for machine learning. For this, we propose a simple method—LesionMap (LM)—and an electronic health (eHealth) software tool—LesionMapper (LMR)—that fits into the clinical workflow.

The sharing of clinical images between dermatologists for learning purposes is common, and most images are published with the consent of the patient [2]. However, increasingly, social media platforms are used for the easy sharing of such resources, with the associated implications on privacy [3]. Machine learning and AI applications need access to a large volume of data to build machine learning models for clinical decision support. Emerging techniques in machine learning and AI such as convolutional neural networks (CNN) and transfer learning [4] have several applications in dermatology [5]. Interestingly, some computer-vision methods can be applied to machine-generated images in addition to digital images [6].

In this paper, we describe common skin lesions, the semantic annotation methodology (LM), and a software tool (LMR) that can be used for semantic annotations. The tool is designed as an extensible software library (JavaScript) that can be incorporated into web-based EMRs. We briefly describe two such integrations with open-source EMRs—OpenMRS and OSCAR EMR.

## Classification of Skin Lesions

Dermatologists use numerous descriptive terms to identify and describe skin lesions [7]. Flat skin lesions that are small are called macules, and when they exceed 1 cm in size, they are called patches. An elevated dome-shaped lesion is called a nodule, whereas a flat elevated lesion is called a plaque. Small fluid-filled lesions are called vesicles, and if they exceed 1 cm, they are called bullae. If vesicles are filled with pus instead of clear fluid, they are called pustules.

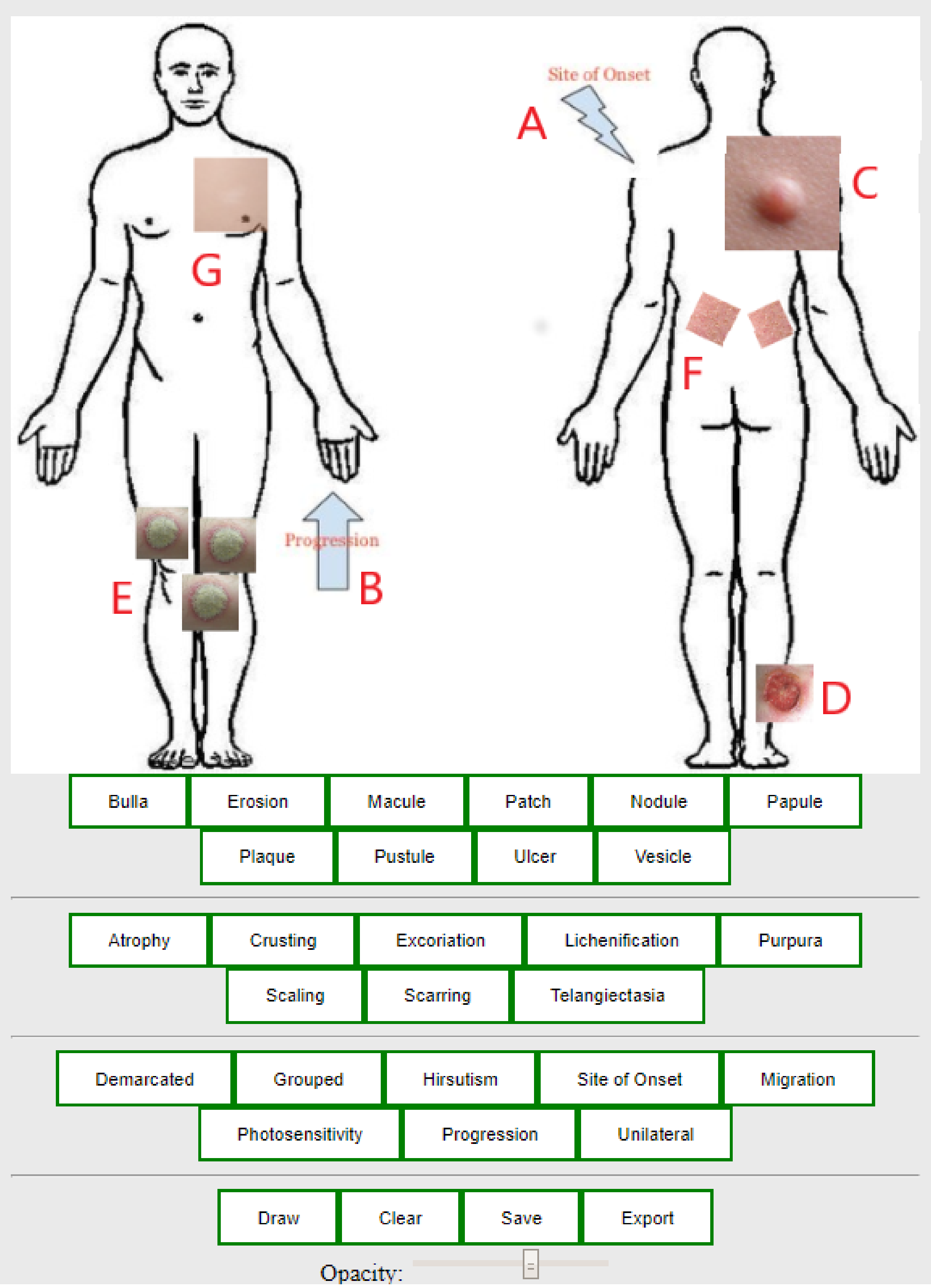
Scales refer to a thickened outer layer of skin while the crust is a liquid debris. An ulcer is an irregularly shaped, deep loss of skin, and if it is superficial it is called an erosion. Atrophy is a thinning of skin, and a fissure is a linear cleft. Necrosis is dead skin tissue, and the scar is the replacement of lost skin by connective tissue. Localized hemorrhage into the skin is called purpura, and petechiae, when the hemorrhagic lesions are small.

The color of the lesion can provide diagnostic cues, along with the shape, arrangement, and distribution. Discoid and annular are terms used to describe the shape. The distribution can be grouped, discrete, linear, serpiginous, reticular, generalized, symmetrical, or photodistributed. The size, location, and severity are also important. Although this is not an exhaustive list of dermatological descriptions, the most common descriptions are included here. Discrepancies in the terminology of dermatological lesions exist in the literature [8]. LM does not attempt to formalize the ontology, but proposes a pragmatic standard using the iconographic method.

## Iconographic Representations of Skin Lesions

Most descriptive terms used in dermatology can be represented by iconographic images representative of the lesion or feature. The use of iconography in clinical documentation has been demonstrated in the context of pain [9]. The type of lesion can be easily represented by icons due to their visual similarity. The list of icons can be supplemented with custom icons for representing descriptive characteristics, such as the site of onset. LMR provides a set of icons for representing visual and nonvisual characteristics of common skin lesions and additional icons for descriptive characteristics (see Figure 1A and B).

**Figure 1.** The LesionMapper interface: (A,B) icons for descriptive characteristics; (C) large nodule; (D) ulcer on lateral side; (E) multiple discrete plaques; (F) Christmas tree pattern in pityriasis rosea; (G) vitiligo patch showing variation in opacity.



In addition to the type of lesions, there are five other characteristics of each icon that can be changed: size, position, number, orientation, and opacity. Additional information pertaining to the lesion can be encoded using the following characteristics:

- The size of the icon can be used to indicate the average size of the lesion in conditions where lesion size points toward a diagnosis or a particular subtype of the primary diagnosis. For example, the size of the plaques can be a differentiating feature for small-plaque and large-plaque parapsoriasis. The original size of the icon, when placed on the LM, can be used for comparison (see the large nodule in Figure 1C).
- The position of icon placement indicates the distribution of the lesions. The front and back of the body are depicted in the LM. The lateral view is not included to simplify the interface. To represent lateral distribution, the icons can be

placed in the corresponding edge of the wireframe with an overlap of 50% (see the ulcer on the legs in Figure 1D).
- Multiple icons of the same type can be used to represent discrete lesions, and a single large icon can be used to represent confluent distribution (see discrete plaques in Figure 1E).
- The orientation can be used to indicate a pathognomonic distribution, such as the Christmas tree pattern in pityriasis rosea (see Figure 1F).
- The opacity of the lesion can be used to indicate the severity of the presentation. For example, it can be used to represent the degree of depigmentation in a vitiligo patch or the severity of contact dermatitis (see Figure 1G).

Mapping lesions consistently and accurately requires a tool that supports the various functions described above. In addition, from a design perspective, the tool should have the capability to integrate with other health information systems and EMRs.

## LesionMapper

LMR is a prototype implementation of the LM method described above. We adopted the design science principles of Hevner et al [10] for information systems to design LMR. We searched the literature for similar approaches and available tools to address the problem of lesional documentation. Based on the success of similar approaches (Pain-QuILT for annotating pain [9]), we chose iconography as the method and standardized it based on our domain expertise in dermatology. Thereafter, we distinguished some of the easily identifiable characteristics of icons that can be programmatically controlled, such as size, orientation, and transparency. Subsequently, we converged on a popular framework (VueJS JavaScript framework [11]) for implementation. We designed the artifact adopting a modular pattern—as a JavaScript package shared as open source (see the GitHub repository [12])—that can be incorporated into web-based EMRs. LMR provides buttons to add various icons to the canvas. These icons can be independently moved and resized. The opacity and orientation can also be independently modified. The LM can be exported as an image or as a JavaScript Object Notation (JSON) string. LMR supports freehand drawing in the canvas to represent features that are not represented by icons though machine interpretation of the freehand drawing is challenging.

Next, we describe the integration of LMR into two open-source EMRs—OpenMRS and OSCAR.

## Integration With Electronic Medical Records

The modular design helps in the integration of LMR into web-based EMRs. The prototype is created using the VueJS JavaScript framework following the Universal Module Definition (UMD) pattern [13] that can be imported by different module loaders into other browser-based applications. The icons are converted into Base64 strings and included in the JavaScript files.

Open Medical Records System (OpenMRS) is an open-source, Java-based EMR for developing countries with a modular and extensible architecture [14]. OpenMRS supports the Open Web Apps (OWA) specifications that make it possible to design external applications that extend the core functions. The OWA communicates with OpenMRS using REST APIs (representational state transfer application programming interfaces), a software architectural style used for creating Web services, and is embedded in the same server instance. OpenMRS has a custom concept dictionary that helps map data points to a uniform terminology. Nontextual data such as images are stored as ”complex concepts” outside the relational database. LMR can easily fit into an OWA design pattern, and the exported LM images can be stored as complex observations in the patient record. We have a prototype integration that can be used as an example [15].

OSCAR EMR is a web-based EMR system initially developed for primary care practitioners in Canada. OSCAR EMR has a complex data model, and additional data points are supported by an electronic form (eForm) module that stores data as key-value pairs [16]. eForms do not support images or other nontextual data. The ability of LMR to save LMs as a JSON string makes the integration of the LMR module into eForms possible.

## Machine Learning Applications

Dermatological diseases have diverse presentations, with skin type and skin color adding to this variation. Some of these diseases involve hair, nails, and mucous membranes in addition to the skin. Traditional computer-vision algorithms such as convolutional neural networks (CNNs) and other variants of neural networks have limited application when there are many decision alternatives [17]. Hence, AI algorithms have had limited application in dermatology except in problems associated with classification (eg, the presence or absence of cancer) [18]. Such algorithms can classify only a given lesion rather than the patient as a whole (ie, a lesion is cancerous vs patient has cancer). Although few CNN-based image search algorithms have proven to be useful, AI algorithms for diagnostic decision making in clinical dermatology lag far behind areas such as radiology [17].

Text analytics and natural language processing (NLP) can be more useful than image analytics when the decision alternatives are numerous, as in dermatology. Multimodal approaches where an image is combined with metadata have shown promise [19]. Machine learning models built using LMs—especially the models created using the JSON representation—resemble text more than an image. Some relevant metadata such as the position and distribution of the lesions, which are difficult to be captured in text and hard to precisely decipher with NLP, are implicitly captured in LMs. The icons represent ontological concepts from dermatology and can map to any standard terminology system [7]. We posit that LMs are semantically rich enough to be used for machine learning applications. Machine learning models from LMs are likely to be more “explainable” than traditional black box algorithms [20]. The implicit metadata captured by

LMs can supplement regular digital images, leading to better machine learning models.

## Advantages and Limitations

LMs may save time for busy practitioners while capturing the type, distribution, and characteristics of the lesion; these data can be used to assess clinical progress. The LM exported as JSON resembles a markup language amenable to data mining and machine learning methods [21]. LMs are portable and can be easily and safely exchanged without privacy concerns.

LMR can export LMs as images. These images can be used as a proxy for patient images in some computer vision–based applications. Computer vision has been successfully applied to identify metabolic defects from gene expression maps [22]. MNIST (Modified National Institute of Standards and Technology database), a dataset widely used in machine learning, consists of images of handwritten digits [23].

It is widely accepted that machine learning can reinforce some health care disparities in dermatology [24]. Skin color is a significant background noise that needs to be accounted for in any machine learning model. It is possible that some of the existing models are biased toward particular skin types that predominate in the training data set. Such models tend to be less sensitive in making predictions on different skin types [25]. The LMs are not affected by such bias.

LMs, however, do not capture all the features, both explainable and unexplainable, captured by a digital image. Hence, LMs are not useful in scenarios where accurate and sensitive extraction of features from an image is important for prediction. For example, LMs are not appropriate for skin cancer classification [5] and mole mapping [26]. LMs do not support annotating dermatopathology images [27]; they are also not applicable for dermatoscopic images that rely on pixel-level analysis [28]. Examination findings such as fluctuation, consistency, and tenderness are not represented by icons at present to keep the interface simple. More icons can be added if the user community requires them.

LMR and the LM method have not been clinically tested. The integration of LMR into existing EMRs may be difficult. Despite its anticipated ease of use, the actual impact of LMR on physician workflow, if any, needs to be investigated further.

## Discussion

The skin is the largest organ in the human body, and as such, skin conditions are commonly encountered in any health care practice. Although dermatology is a specialization within clinical medicine, 50% of skin conditions are assessed and documented by nondermatologists [29].

There is no universal standard for pictographic documentation of the type, distribution, progression, and severity of lesions in dermatology, as in dentistry [30] and ophthalmology [31]. The LM standardizes visual representation using icons that can be extended to accommodate different use cases in clinical and cosmetic dermatology. The simplicity of the mapping rules facilitates use by nondermatologists in the skincare industry; LMs are also semantically rich enough to capture most relevant information about a skin condition with minimal effort.

Image analytics in dermatology is not as popular, as it is in visually oriented medical specialties such as radiology and pathology; the exception is the field of skin cancer diagnostics. This is because of the privacy concerns associated with dermatological images and the difficulties in standardizing image capture. The LM is not a replacement for a digital image of the lesion. However, some of the diagnostic aspects that are difficult to be captured in images, such as distribution and progression, can be useful for machine learning applications, especially when combined with the textual representation of a patient’s history. Such multimodal approaches mimic the clinical workflow more so than CNN-based algorithms [32]. New computer-vision algorithms are proving to be capable of learning from computer-generated images [22]. We believe that LMs can be similarly used with computer-vision methods. Finally, we urge the open-source community to help us improve LMR and potential users to report issues on the repository [12] so that we can fix them. We will work on a 3D wireframe for better accuracy, and we welcome other feature requests from the user community.

### Conflicts of Interest

None declared.

### References

1. Aspres N, Egerton IB, Lim AC, Shumack SP. Imaging the skin. Australas J Dermatol 2003 Feb;44(1):19-27. [doi: 10.1046/j.1440-0960.2003.00632.x]
2. Kunde L, McMeniman E, Parker M. Clinical photography in dermatology: Ethical and medico-legal considerations in the age of digital and smartphone technology. Australasian Journal of Dermatology 2013 May 29;54(3):192-197. [doi: 10.1111/ajd.12063]
3. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. P & T 2014 Jul;39(7):491-520 [FREE Full text] [Medline: 25083128]
4. Shin H, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Trans. Med. Imaging 2016 May;35(5):1285-1298. [doi: 10.1109/tmi.2016.2528162]

5. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017 Jan 25;542(7639):115-118. [doi: 10.1038/nature21056]
6. Levine MD, Nazif AM. Dynamic Measurement of Computer Generated Image Segmentations. IEEE Trans Pattern Anal Mach Intell 1985 Mar;PAMI-7(2):155-164. [doi: 10.1109/tpami.1985.4767640]
7. Eapen BR. ONTODerm--a domain ontology for dermatology. Dermatol Online J 2008 Jun 15;14(6):16. [Medline: 18713597]
8. Cardili RN, Roselino AM. Elementary lesions in dermatological semiology: literature review. An Bras Dermatol 2016 Oct;91(5):629-633. [doi: 10.1590/abd1806-4841.20164931]
9. Lalloo C, Kumbhare D, Stinson JN, Henry JL. Pain-QuILT: clinical feasibility of a web-based visual pain assessment tool in adults with chronic pain. J Med Internet Res 2014 May 12;16(5):e127 [FREE Full text] [doi: 10.2196/jmir.3292] [Medline: 24819478]
10. Hevner AR, March ST, Park J, Ram S. Design Science in Information Systems Research. MIS Quarterly 2004;28(1):75. [doi: 10.2307/25148625]
11. Street M, Passaglia A, Halliday P. Complete Vue. In: Complete Vue.js 2 Web Development: Practical Guide To Building End-to-end Web Development Solutions With Vue.js 2. Birmingham: Packt Publishing; 2018.
12. Eapen BR. GitHub. 2020. LesionMapper - Vue component for standardized mapping of lesions in dermatology URL: https://github.com/dermatologist/lesion-mapper [accessed 2020-01-16]
13. Rich H. Dismantling the barriers to entry. Queue 2015;13(5):37 [FREE Full text] [doi: 10.1145/2742580.2742813]
14. Mamlin BW, Biondich PG, Wolfe BA, Fraser H, Jazayeri D, Allen C, et al. Cooking up an open source emr for developing countries: OpenMRS, a recipe for successful collaboration. In: American Medical Informatics Association. 2006 Presented at: AMIA Annual Symposium Proceedings; 2006; Washington p. 529.
15. Eapen BR. OpenMRS. 2017. Dermatology LesionMapper URL: https://addons.openmrs.org/show/org.openmrs.module.dermatology-lesionmapper [accessed 2020-01-18]
16. Safadi H, Chan D, Dawes M, Roper M, Faraj S. Open-source health information technology: A case study of electronic medical records. Health Policy and Technology 2015 Mar;4(1):14-28. [doi: 10.1016/j.hlpt.2014.10.011]
17. Li C, Shen C, Xue K, Shen X, Jing Y, Wang Z, et al. Artificial intelligence in dermatology. Chinese Medical Journal 2019;132(17):2017-2020. [doi: 10.1097/cm9.0000000000000372]
18. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. BMC Med Res Methodol 2019 Mar 19;19(1). [doi: 10.1186/s12874-019-0681-4]
19. Wang H, Wang Y, Liang C, Li Y. Assessment of Deep Learning Using Nonimaging Information and Sequential Medical Records to Develop a Prediction Model for Nonmelanoma Skin Cancer. JAMA Dermatol 2019 Nov 01;155(11):1277. [doi: 10.1001/jamadermatol.2019.2335]
20. Andreas H, Chris B, Constantinos SP, Douglas BK. What do we need to build explainable AI systems for the medical domain? arXiv.org 2017 [FREE Full text]
21. Guazzelli A, Lin W, Jena T. PMML in action: unleashing the power of open standards for data mining and predictive analytics. California: Createspace Independent Publishing Platform; 2020.
22. Zhang J, Naik HS, Assefa T, Sarkar S, Reddy RVC, Singh A, et al. Computer vision and machine learning for robust phenotyping in genome-wide studies. Sci Rep 2017 Mar 8;7(1). [doi: 10.1038/srep44048]
23. Li Deng. The MNIST Database of Handwritten Digit Images for Machine Learning Research [Best of the Web]. IEEE Signal Process. Mag 2012 Nov;29(6):141-142. [doi: 10.1109/msp.2012.2211477]
24. Adamson AS, Smith A. Machine Learning and Health Care Disparities in Dermatology. JAMA Dermatol 2018 Nov 01;154(11):1247. [doi: 10.1001/jamadermatol.2018.2348]
25. Kamulegeya LH, Okello M, Bwanika JM, Musinguzi D, Lubega W, Rusoke D, et al. Using artificial intelligence on dermatology conditions in uganda: A case for diversity in training data sets for machine learning. bioRxiv.org 2019 [FREE Full text] [doi: 10.1101/826057]
26. Berk-Krauss J, Polsky D, Stein JA. Mole Mapping for Management of Pigmented Skin Lesions. Dermatologic Clinics 2017 Oct;35(4):439-445. [doi: 10.1016/j.det.2017.06.004]
27. Toro P, Corredor G, Romero E, Arias V. A visualization, navigation, and annotation system for dermatopathology training. In: International Society for Optics and Photonics. 2018 Presented at: 14th International Symposium on Medical Information Processing and Analysis, volume 10975; 2018; Mazatlán, Mexico. [doi: 10.1117/12.2511638]
28. Ferreira B, Barata C, Marques JS. What is the role of annotations in the detection of dermoscopic structures? In: Iberian Conference on Pattern Recognition and Image Analysis. 2019 Presented at: 9th Iberian Conference, IbPRIA 2019; 2019; Madrid, Spain p. 3-11. [doi: 10.1007/978-3-030-31321-0_1]
29. Ahiarah A, Fox C, Servoss T. Brief intervention to improve diagnosis and treatment knowledge of skin disorders by family medicine residents. Fam Med 2007;39(10):720-723 [FREE Full text] [Medline: 17987414]
30. Wu M, Koenig L, Lynch J, Wirtz T. Spatially-oriented EMR for Dental Surgery. In: American Medical Informatics Association. 2006 Presented at: AMIA Annual Symposium Proceedings; 2006; Washington p. 1147.
31. Chiang MF, Boland MV, Brewer A, Epley KD, Horton MB, Lim MC, American Academy of Ophthalmology Medical Information Technology Committee. Special Requirements for Electronic Health Record Systems in Ophthalmology. Ophthalmology 2011 Aug 1;118(8):1681-1687. [doi: 10.1016/j.ophtha.2011.04.015] [Medline: 21680023]

32. Baltrusaitis T, Ahuja C, Morency L. Multimodal Machine Learning: A Survey and Taxonomy. IEEE Trans Pattern Anal Mach Intell 2019 Feb 1;41(2):423-443. [doi: 10.1109/tpami.2018.2798607]

## Abbreviations

**AI:** artificial intelligence
**API:** application programming interface
**CNN:** convolutional neural network
**eForm:** electronic form
**eHealth:** electronic health
**EMR:** electronic medical record
**JSON:** JavaScript Object Notation
**LM:** LesionMap
**LMR:** LesionMapper
**MNIST:** Modified National Institute of Standards and Technology database
**NLP:** natural language processing
**OWA:** Open Web Apps
**REST:** representational state transfer
**UMD:** Universal Module Definition

*Edited by G Eysenbach; submitted 06.02.20; peer-reviewed by F Kaliyadan, A Kt; comments to author 15.02.20; revised version received 26.02.20; accepted 27.02.20; published 20.04.20.*

*Please cite as:*
*Eapen BR, Archer N, Sartipi K*
*LesionMap: A Method and Tool for the Semantic Annotation of Dermatological Lesions for Documentation and Machine Learning*
*JMIR Dermatol 2020;3(1):e18149*
*URL: http://derma.jmir.org/2020/1/e18149/*
*doi:10.2196/18149*
*PMID:*

Original Paper

# Assessment of Patient Satisfaction With Dermatology Clinics According to Clinic Type: Mixed Methods Study

Jennifer Costigan[1*], MSHI; Sue S Feldman[1*], PhD, MEd, RN; Mark Lemak[2*], MHSA

[1]Department of Health Services Research, School of Health Professions, University of Alabama at Birmingham, Birmingham, AL, United States
[2]University of Alabama at Birmingham Health System, Birmingham, AL, United States
[*]all authors contributed equally

**Corresponding Author:**
Sue S Feldman, PhD, MEd, RN
Department of Health Services Research
School of Health Professions
University of Alabama at Birmingham
1716 9th Avenue South
SHPB #590K
Birmingham, AL, 35294
United States
Phone: 1 205 975 0809
Email: suefeldman1009@gmail.com

## Abstract

**Background:** Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey responses are considered significant indicators of the quality of care and patient satisfaction. There is a pressing need to improve patient satisfaction rates as CAHPS survey responses are considered when determining the amount a facility will be reimbursed by the Centers of Medicare and Medicaid each year. Low overall CAHPS scores for an academic medical center’s dermatology clinics were anecdotally attributed to clinic type. However, it was unclear whether clinic type was contributing to the low scores or whether there were other factors.

**Objective:** This study aimed to determine where the efforts of patient satisfaction improvement should be focused for two different types of dermatology clinics (private and rapid access clinics).

**Methods:** This study used a concurrent mixed methods design. Secondary data derived from the University of Alabama at Birmingham Hospital’s Press Ganey website were analyzed for clinic type comparisons and unstructured data were qualitatively analyzed to further enrich the quantitative findings. The University of Alabama at Birmingham Hospital is an academic medical center. The data were analyzed to determine the contributors responsible for each clinic not meeting national benchmarks. Thereafter, a review of these contributing factors was further performed to assess the difference in CAHPS scores between the private and rapid access clinics to determine if clinic type was a contributing factor to the overall scores.

**Results:** The data sample included 821 responses from May 2017 to May 2018. Overall, when both private clinics and rapid access clinics were viewed collectively, majority of the patients reported stewardship of patient resources as the most poorly rated factor (367/549, 66.8%) and physician communication quality as the most positively rated factor (581/638, 91.0%). However, when private clinics and rapid access clinics were viewed individually, rapid access clinics contributed slightly to the overall lower dermatology scores at the academic medical center.

**Conclusions:** This study determined that different factors were responsible for lower CAHPS scores for the two different dermatology clinics. Some of the contributing factors were associated with the mission of the clinic. It was suspected that the mission had not been properly communicated to patients, leading to misaligned expectations of care at each clinic.

***(JMIR Dermatol 2020;3(1):e17171)*** doi:10.2196/17171

**KEYWORDS**

Consumer Assessment of Healthcare Providers and Systems survey scores; patient satisfaction; dermatology; private dermatology clinic; rapid access dermatology clinic; wait time; patient resource stewardship; communication

## Introduction

### Background

Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey responses are considered significant indicators of the quality of care and patient satisfaction. These scores are derived from a series of standardized patient surveys used to assess patients' experiences and satisfaction with health care in the United States. These surveys include quality measures that patients are most qualified to assess, such as aspects related to communication and interaction with medical staff [1,2].

There is a pressing need to improve patient satisfaction rates as CAHPS scores are considered when determining the amount a facility will be reimbursed by the Centers for Medicare and Medicaid each year [2]. Thus, focusing on improving patient satisfaction can, in turn, help to improve the quality of care and patient outcomes, which are both primary initiatives in health care today.

The overall CAHPS scores for the dermatology clinics of a particular academic medical center (AMC; University of Alabama at Birmingham Hospital) have been unsatisfactory for various measures over the past several years. Anecdotally, the private clinic personnel suspected that the rapid access clinics were contributing to the overall low scores. Private clinics are set up such that patients make appointments in advance with the dermatologist of their choosing. In this environment, they can see the same dermatologist for every appointment. On the other hand, rapid access clinics are set up such that patients can get walk-in or "just-in-time" appointments with the dermatologist or dermatology nurse practitioner on duty. There is some crossover in dermatology clinicians; however, there is no guarantee that the patient can see the clinician of choice in the rapid access clinic environment. Additionally, the rapid access clinics involve a rapid approach. Patients are seen for situational care rather than routine care. For example, ideally, an annual dermatology examination would be scheduled in a private clinic where there is more time per appointment, whereas a newly occurring rash would be seen in a rapid access clinic where there is less time per appointment. This scheduling in the rapid access clinic environment is adopted to accommodate more patients. Assessment of the ambulatory CAHPS scores at a granular level was conducted to identify patterns and specific areas of measure between the two different types of dermatology clinics (private and rapid access) that are impacting the overall CAHPS scores. The overarching question being answered was as follows: Where should the efforts of patient satisfaction improvements be focused in the two different types of dermatology clinics?

### CAHPS Survey as a Measure of Patient Experience

CAHPS survey data have been utilized and referenced as direct patient care quality indicators for health care organizations across the United States [1]. The push to move toward a pay-for-performance model and quality-driven health care in the United States has influenced the emergence of such measuring tools and initiatives [3]. According to the Agency for Healthcare Research and Quality, the creator of the CAHPS program, the CAHPS survey is an initiative that was designed to "support investigator-led research to better understand patient experience with health care and develop scientifically valid and feasible strategies and tools to assess patient experience, report survey results, and help organizations use the results to improve the quality of care" [1].

The results of CAHPS measures can effectively bring attention to the shortfalls of health care providers and emphasize the need for improvements in their delivery of quality care. However, there appears to be a gap between understanding the collected data and being able to use and interpret the data to create actionable goals and improvement initiatives [4]. The availability of such data is only valuable if it can be translated and applied to create relevant solutions and to optimize the impact such solutions would have on quality care outcomes for an organization. Thus, determining which care factors and initiatives correlate directly with outcome measures, such as those obtained from CAHPS surveys, remains a challenge in health care today.

### Relationship of Data Analysis and Patient Care

Aside from existing challenges, research studies infer that data collection and analysis efforts of patient care and quality measures are highly useful for identifying shortcomings in the delivery of patient care and indicating the importance of continually measuring care plans [5,6]. According to one source, a key element that strengthens patient satisfaction and confidence in clinical care is viewing quality improvement as a constant, iterative, and transparent process [7]. There is also evidence indicating that the publication of care performance data influences quality improvement initiatives to be adopted by hospitals [8]. As such, the awareness of scores being accessible to the public can encourage organizations to be more conscious of their results [8].

Additionally, satisfaction scores are highly useful to identify opportunities for improvement in aspects of patient care. Furthermore, the literature suggests that patient satisfaction should be consistently evaluated and used to assess specific interventions and improvement areas within care practices [5]. For example, in research associated with a Healthcare Information and Management System Society study (HIMSS), measurement of patient satisfaction revealed that providers' perception of their own performance does not always accurately align with patient reality, as approximately 72% of patients responded being extremely or very satisfied with their overall care experience as compared with only 39.7% of providers believing they were extremely or very effective (N=309 and N=204, respectively) [9], suggesting that anecdotal reports may not accurately reflect actual patient satisfaction. This supports the purpose and relevance of the analysis of the dermatology clinics' CAHPS scores. The information contained in patient feedback and surveys can outline the areas in need of improvement and, based on what those areas are, indicate what type of recommendations need to be pursued. The objective of this study was to analyze the ambulatory CAHPS scores at a granular level in order to identify patterns and specific areas of measure in the two different types of dermatology clinics (private and rapid access clinics), which are impacting the overall CAHPS scores.

## Methods

### Study Design

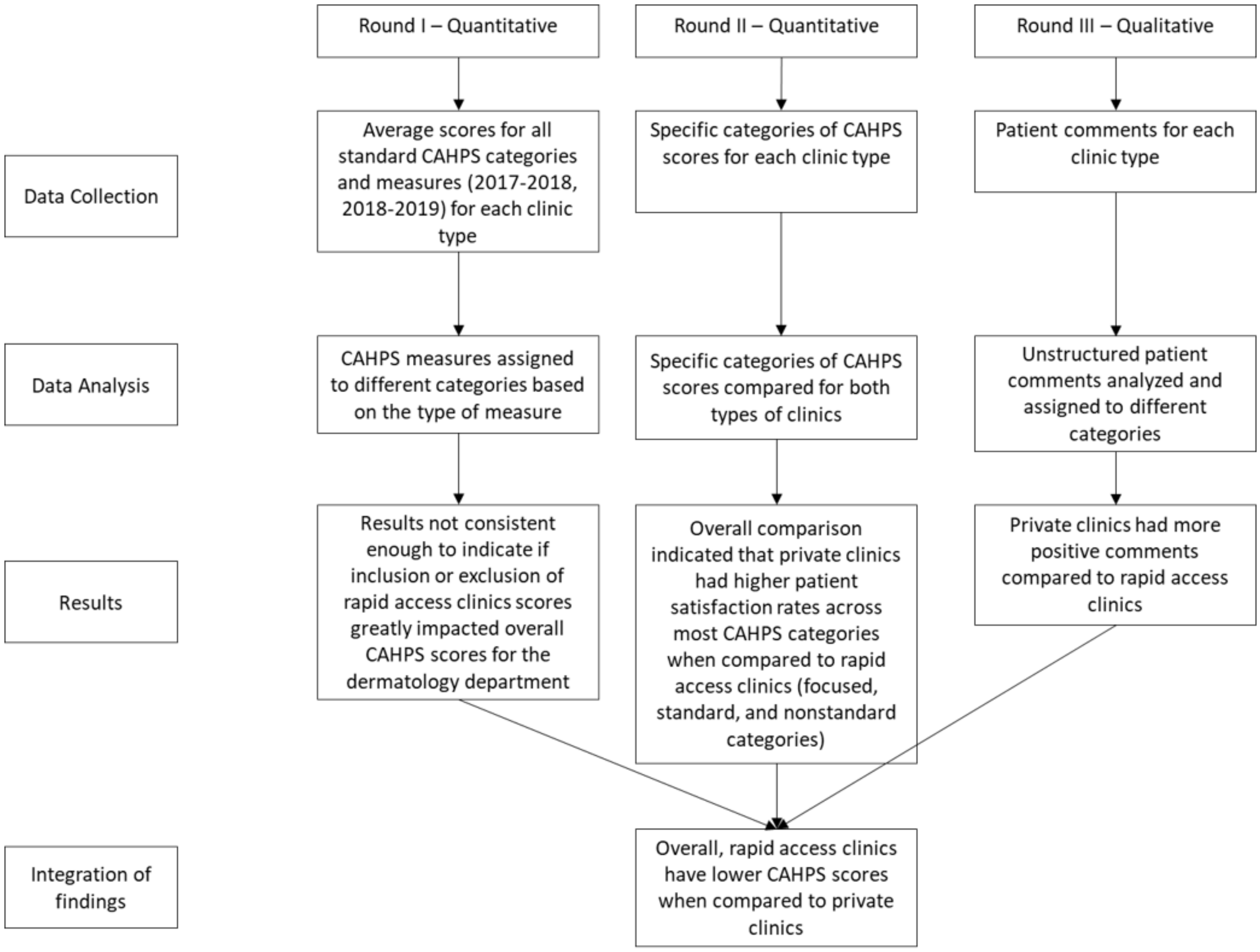
This study had a concurrent mixed methods design, where the findings from a qualitative analysis were used to enrich the findings from a quantitative analysis. Deidentified secondary data derived from the University of Alabama at Birmingham Hospital's Press Ganey website were analyzed for clinic type comparisons. The data were analyzed to determine the contributors responsible for each clinic not meeting national patient satisfaction benchmarks. Thereafter, these contributing factors were further analyzed to determine the difference in CAHPS scores between the private and rapid access clinics to determine if clinic type was a contributing factor to the overall scores.

Based on the most recent survey template provided, these surveys consist of over 60 different standardized CAHPS questions within 10 different domains, 13 additional nonstandard questions, and a section for additional patient comments that assess patient care experience and satisfaction with their visits to the clinics. The scope of this study required filtering of the dataset by dermatology. In order to separate the private clinics and rapid access clinics from the larger dataset, a holistic week-long schedule was assessed, and it outlined the times of operation and staffing for the dermatology clinics and provided the necessary information for comparison of the survey data by clinic type, which was conducted further in the analysis.

The data sample included 821 responses from May 24, 2017 through May 24, 2018. All data were analyzed using Microsoft Excel (Microsoft Corporation, Redmond, Washington, USA) and Tableau (Salesforce, Mountain View, California, USA) for visualization. This study was conducted at the University of Alabama at Birmingham Hospital (approval number: #300003087).

Analysis of the data was performed in three rounds. Round I (quantitative) identified which CAHPS measures and criteria consistently scored the lowest across all clinics and appeared as the greatest contributors to the department's annual CAHPS scores. This round also tried to determine whether the rapid access clinics scores alone were the sole reason why the dermatology scores were low. Round II (quantitative) examined differences between the private and rapid access clinics. Round III (qualitative) took a more granular approach to further investigate the nonstandard quality measures and unstructured patient comments. Figure 1 illustrates the research process from data collection to the integration of findings.

**Figure 1.** Mixed methods research design populated with results and the integration of those results. CAHPS: Consumer Assessment of Healthcare Providers and Systems.



## Quantitative Data Analysis

### *Round I: Contributors to the Annual CAHPS Score*

The process for Round I of the analysis involved extracting the average scores for all standard CAHPS categories and measures for the most recent year, which included survey data of only the private clinics (2018-2019), and the prior year, which included survey data for both the rapid access and private clinics (2017-2018), for comparison.

Comparing CAHPS scores between these two years helped to show the impact of omitting rapid access survey data on dermatology clinics' ratings, which was done in 2018-2019. The measures were then ranked from best to worst based on the average values for both years to determine which categories and individual measures had the most opportunities for improvement. This data review helped to highlight the primary areas in need of attention while also providing insights on specific themes or components of care that may be generally lacking among the clinics.

The format in which the data are presented in Press Ganey-generated reports is not ideal for in-depth analysis and requires substantial restructuring. For example, each measure is separated into different sections in the comma separated value (csv) file, with individual sets of rows and columns for each response.

In order to create standard response categories across measures for a more analysis-friendly format, the different response types were grouped together based on their equivalence to another response. For example, "no" and "never" responses were combined into one response subgroup "no/never," whereas "yes, definitely," "yes," and "always" categories were compiled into the response subgroup "yes, definitely/yes/always." CAHPS questions also included "yes, somewhat/sometimes" and "usually" as responses, but for the purpose of this analysis, only the highest and lowest response categories ("no/never" and "yes, definitely/yes/always") were used for comparison. A sample of the data in restructured format is presented in Table 1.

**Table 1.** Restructured data from Press Ganey survey responses.

| CAHPS[a] measure (2017-2018) | Category | Sample size, n | No/never, n (%) | Yes, definitely/yes/always, n (%) |
|---|---|---|---|---|
| Recommend this provider office | Global | 776 | 16 (2.1) | 706 (91.0) |
| Provider explains in a way you understand | Physician communication quality | 777 | 14 (1.8) | 724 (93.2) |
| Provider listens carefully to you | Physician communication quality | 776 | 18 (2.3) | 723 (93.2) |
| Provider talks with you about a problem/concern | Physician communication quality | 775 | 94 (12.1) | 681 (87.9) |
| Provider gives easy to understand instructions | Physician communication quality | 681 | 8 (1.2) | 628 (92.2) |
| Provider knows important information/medical history | Physician communication quality | 775 | 32 (4.1) | 647 (83.5) |
| Provider shows respect for what you say | Physician communication quality | 775 | 7 (0.9) | 734 (94.7) |
| Provider spends enough time with you | Physician communication quality | 777 | 16 (2.1) | 710 (91.4) |
| Clerks/receptionists are helpful | Office staff quality | 776 | 19 (2.4) | 673 (86.7) |
| Clerks treat you with courtesy/respect | Office staff quality | 775 | 3 (0.4) | 718 (92.6) |

[a] CAHPS: Consumer Assessment of Healthcare Providers and Systems.

### *Round II: CAHPS Score by Clinic Type*

When using Press Ganey data for an organization's specific needs, such as differentiating clinics, a process must be used to merge the data in ways other than the original intent when the survey was developed. For example, our organization's need was to compare two clinics; however, the survey, in its original form, does not allow for this type of comparison. In this section, we detail the process for readers who have interest in duplicating this or a similar study at their organization. Press Ganey does not readily distinguish clinic types. Therefore, the only way to separate data by clinic type was to cross-analyze the Press Ganey data with a physician schedule tracking sheet. The physician schedule tracking sheet facilitated identification of clinics, as well as physicians associated with certain outcomes that could not be determined solely by the CAHPS data. The clinic schedule included days of the week and times of the day (am vs pm) specific physicians work in the private and rapid access clinics, whereas the CAHPS scores in Press Ganey can be viewed by visit date/time. Using the clinic schedules for comparisons with the data available in Press Ganey, we could match the clinic type with the corresponding CAHPS data.

For example, in order to divide the datasets by clinic type, four separate reports had to be pulled from Press Ganey and complied accordingly using schedule and physician criteria included in the schedule. The first report consisted of average patient satisfaction rates for Monday, Wednesday, and Thursday, which were unfiltered to sort out the first section of private clinic data. The second report then consisted of patient satisfaction scores for Tuesday and Friday (am appointments only). To separate the remaining data for only the private clinics, a third report was used to extract scores for Tuesday and Friday (pm appointments), which included only those providers who met patients on these specific days. Finally, the fourth report was extracted for only rapid access data, which included patient satisfaction scores for Tuesday and Friday (pm appointments only), and a filter was applied to exclude the providers of the private clinics (pm) for these specific days. This process had to be completed for each time specific data point (standard CAHPS measures, nonstandard quality measures, and patient comments) and was separated by clinic type.

## Qualitative Data Analysis

### *Round III: Unstructured Patient Comments*

A third round of analysis took a more granular approach to further investigate patient comments. The use of unstructured text comments required an additional data preparation step, in that they had to be assigned to standard groups that could allow for comparison. The preparation for these comments consisted of reviewing all patient comments for both types of dermatology clinics from 2017-2018 and categorizing them into response type themes as follows: positive, negative, indifferent, and not applicable. For example, a comment “I appreciated the receptionist helping me make my next appointment” was categorized as positive. By contrast, a comment “I waited 3 months for my appointment, only to have it cancelled by the doctor the week beforehand” was categorized as negative. The “positive” and “negative” categorized comments were then further categorized into subthemes based on aspects of care to which they were most frequently related. The subthemes included medical needs not met or questions not answered; lack of or delay with results or records; long wait or slow processing; appointment rushed or barely saw the doctor; lack of communication, direction, or follow up; medical staff rude, impersonal, or did not listen; difficult to schedule timely appointment or appointment cancelled by the doctor; check in process frustrating, receptionist unhelpful, or lack of assistance; and general or other. Thereafter, the data were processed through the same organizational approach described in Round I.

# Results

## Results Structure

Results are presented in the following order: demographics, contributors to CAHPS scores (Round I), results of nonstandard quality measures of CAHPS scores by clinic type (Round II), and unstructured patient data (Round III).

## CAHPS Survey Respondents’ Demographic Data

As shown in Table 2, just over half of the respondents were female (464/821, 56.5%), an overwhelming majority were white (650/821, 83.0%), and most were between 50 and 79 years old (589/821, 71.0%). Moreover, a vast majority had some college education (670/821, 87.2%).

**Table 2.** Survey respondents' demographics (N=821).

| Characteristic | Value, n (%) |
|---|---|
| **Gender** | |
| Female | 464 (56.5%) |
| Male | 357 (43.5%) |
| **Ethnicity** | |
| White | 650 (83.0%) |
| African American | 93 (11.9%) |
| Asian | 7 (0.9%) |
| Hispanic/Latino | 5 (0.7%) |
| Hawaiian/Pacific Islander | 2 (0.3%) |
| American Indian/Alaska Native | 3(0.4%) |
| Other | 13 (1.7%) |
| **Age (years)** | |
| 0-17 | 8 (1.0%) |
| 18-34 | 79 (9.6%) |
| 35-49 | 86 (10.5%) |
| 50-64 | 254 (30.9%) |
| 65-79 | 335 (40.8%) |
| 80 or older | 59 (7.2%) |
| **Education** | |
| Eight grade or lower | 7 (0.9%) |
| Some high school | 12 (1.6%) |
| High school graduate | 80 (10.4%) |
| Some college | 195 (25.4%) |
| Four-year college graduate | 183 (23.8%) |
| Greater than 4 years of college | 292 (38.0%) |

## Quantitative Data Analysis Results

### *Round I: Contributors to the Annual CAHPS Score*

As mentioned in the Methods, the annual CAHPS scores were first analyzed with and without the rapid access clinics included. The negative responses (no/never) are shown in Figure 2 and the positive responses (yes, definitely/yes/always) are shown in Figure 3. Both figures represent the comparison of "no/never" and "yes, definitely/yes/always" response rates overall (all weekdays) across CAHPS categories between the year with rapid access clinic scores (2017-2018) and the year without rapid access clinics scores (2018-2019). On examining Figure 2, stewardship of patient resources indicates, among other things, whether clinicians consider the needs of patients when making therapy, medication, or referral recommendations. For this variable, it was found that 66.8% (367/549) of respondents said "no/never" with the inclusion of rapid access clinics as compared with 70.9% (783/1104) of respondents without the inclusion of rapid access clinics, indicating that the data without the rapid access clinics were worse. Another example is how well providers communicate with patients, which is shown in Figure 2 as "physician communication quality," indicating how effective is provider communication with patients. For this variable, it was found that 2.1% (13/638) of respondents said "no/never" with the inclusion of rapid access clinics (meaning communication was poor) as compared with 2.2% (27/1227) of respondents without the inclusion of rapid access clinics, indicating that the data without the rapid access clinics were worse, and thus, the data were worse for private clinics.

Figure 3 shows the positive responses while using the same quality measures. With regard to stewardship of patient resources, involving whether providers consider the needs of patients, 33.2% (182/549) of respondents said "yes, definitely/yes/always" with the inclusion of rapid access clinics as compared with 29.1% (321/1104) of respondents without the inclusion of rapid access clinics, indicating that a greater percentage thought that their needs were considered in data that included the rapid access clinics. Similarly, with regard to physician communication, 91.0% (581/638) of respondents with the inclusion of rapid access clinics answered "yes, definitely/yes/always" as compared with 90.6% (1112/1227) of respondents without the inclusion of rapid access clinics.

Considering Figures 2 and 3, there does not appear to be a wide variance between most of the measures. One exception is "access to specialists." On assessing the data that included the rapid access clinics (Figure 2), 0.0% of respondents reported "no/never" access to specialists, indicating that specialists are accessible. However, on assessing the data that did not include the rapid access clinics (ie, private clinics only; Figure 2), 10.0% (122/1224) of respondents reported "no/never" access to specialists, indicating that specialists were not accessible. The variance is much wider in Figure 3, which presents the positive responses while using the same quality measures. With regard to access to specialists, Figure 3 demonstrates that 66.7% (2/3) of respondents said "yes, definitely/yes/always" with the inclusion of the rapid access clinics as compared with 20.0% (1/5) of respondents without the inclusion of the rapid access clinics, indicating that more respondents felt that there was access to specialists with inclusion of the rapid access clinics.

**Figure 2.** Negative responses with and without the rapid access clinics.

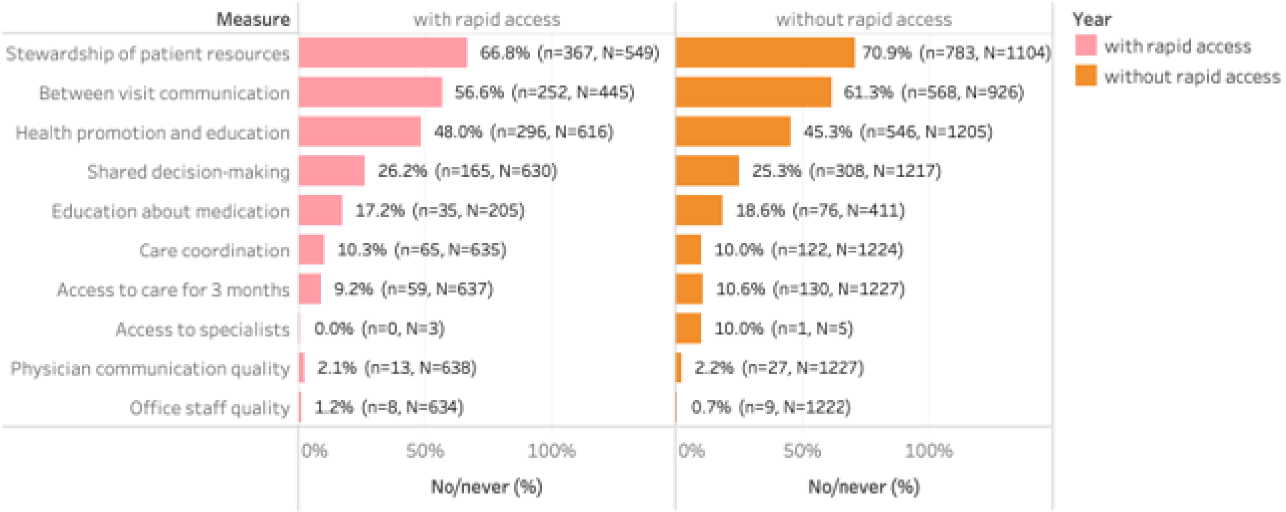
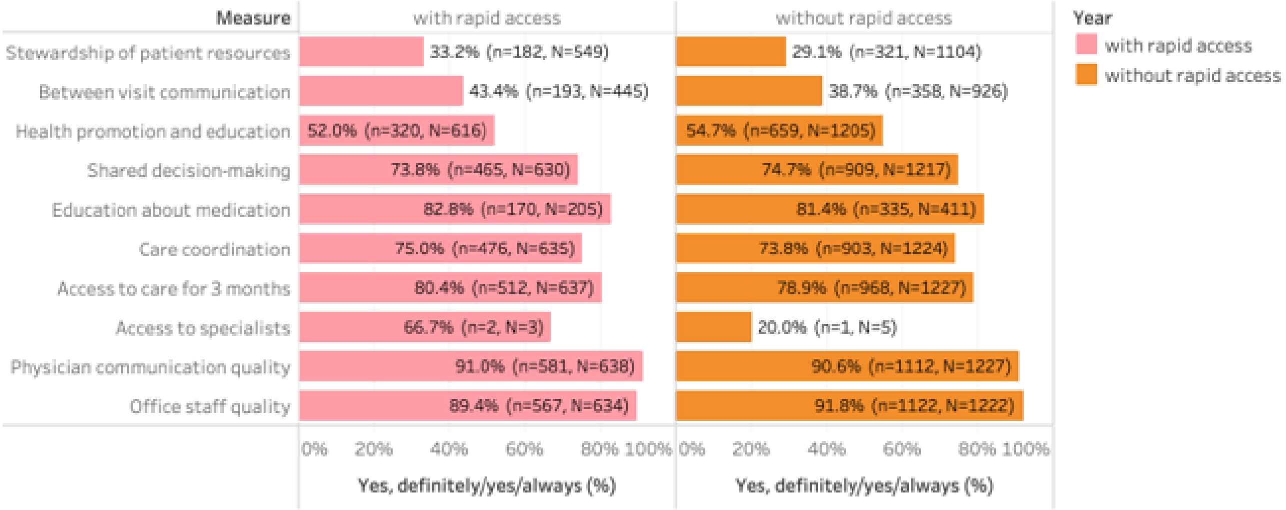


**Figure 3.** Positive responses with and without the rapid access clinics.



Figures 4 and 5 present comparisons of "no/never" and "yes, definitely/yes/always" response rates for Tuesday and Friday (rapid access clinic days) across CAHPS categories for the year with rapid access clinic scores (2017-2018) and the year without rapid access clinic scores (2018-2019). These results suggest the same conclusions as those involving data for all weekdays. It appears that there was also a slight increase in negative response ratings and a slight decrease in positive response ratings for several of the categories. However, in the case of both results, there were also some CAHPS categories that had an increase in the positive response average and a decrease in the negative response average. Essentially, the results are not consistent enough to indicate whether the inclusion or exclusion of rapid access clinic scores greatly impacted the dermatology clinics' CAHPS scores as a whole. However, these results consistently showed that CAHPS scores for the "stewardship of patient resources" category was higher for the negative response rate (Figure 4), indicating that in the rapid access clinics, patients reported better stewardship of patient resources. Unexpectedly, both with and without inclusion of the rapid access clinics did not receive any responses (positive or negative) for the "access to specialists" category (Figures 4 and 5). Figure 4 shows that when the rapid access clinics were included, there was better consideration for patient resources.

It was noted earlier that the rapid access clinics only operated on Tuesday and Friday afternoons. However, the data in Figure 3 considered the entire day because comparisons were performed between two separate years (one including and the other excluding rapid access clinic data). Therefore, if the data were compared between the two years by Tuesday and Friday (pm), there would have been no rapid access clinic data available according to the criteria for 2018-2019. Using the data of the entire day for the weekdays when the rapid access clinics operated allowed us to observe the impact of excluding rapid access data for those particular days and the results associated with the private clinics alone for the most current year (2018-2019).

**Figure 4.** Negative responses with and without the rapid access clinics on Tuesday and Friday only. N/A: not applicable.

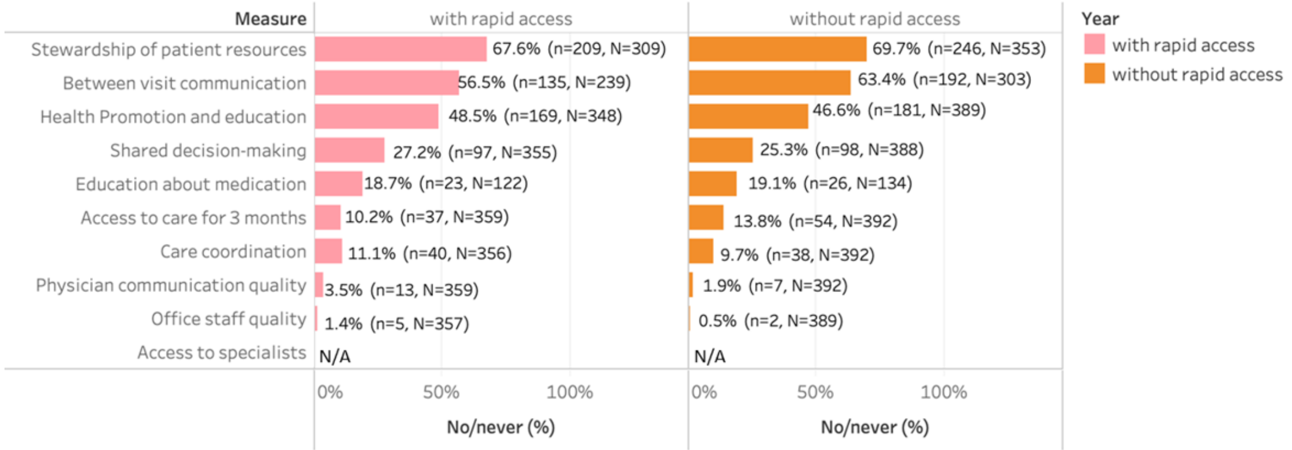
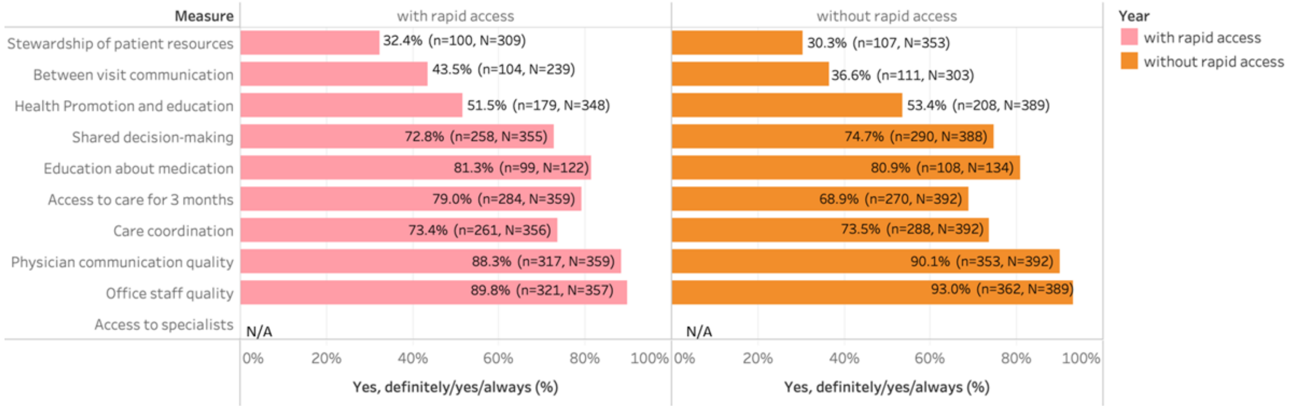


**Figure 5.** Positive responses with and without the rapid access clinics on Tuesday and Friday only. N/A: not applicable.



### *Round II: Data Review by Clinic Type*

We show the results relative to how well the private clinics and rapid access clinics scored on average across all CAHPS categories in Figure 6. Thereafter, we show more granular results of the focus measures by clinic, with the results of the private clinics presented in Figure 7 and those of the rapid access clinics presented in Figure 8.

The focus measures were the measures that targeted physician communication quality, standard measures as well as additional measures based on relevance to care quality, and the potential for improvement relative to the effort for change management. All CAHPS categories were used in the first comparison, rather than selecting just those related to the focus measures, in order to conduct a general comparison of all CAHPS categories according to clinic type. In other words, Figure 6 shows the higher level CAHPS categories/domains that were measured. Each of these categories/domains had a set of related CAHPS measures. The categories/domains relevant to the focus measures chosen for analysis are indicated in Figure 6. These focus measures are further assessed in Figures 7 and 8.

The results of the overall CAHPS scores by clinic type, as depicted in Figure 6, showed that the private clinics had higher patient satisfaction rates across most CAHPS categories when compared with the rapid access clinics. When looking further at the focus measures by clinic type (Figures 7 and 8), most of the measures were higher for the private clinics as compared with the rapid access clinics, although some measures were only slightly higher.

We present the findings for focus measures in the private clinics and rapid access clinics in Figures 7 and 8, respectively. On comparing the two clinic types, the results showed that the private clinics scored higher in most areas. The rapid access clinics scored higher in “appointment for care right away,” “talk with provider about prescription,” “provider gives written medical instructions,” and “talk with provider about problem/concern.” These findings make sense when considering the nature of the rapid access clinics as a situational environment where a patient shows up with or without an appointment, has a targeted conversation with the provider about typically a singular problem or concern, gets a prescription or some other type of instruction, and leaves.

**Figure 6.** Consumer Assessment of Healthcare Providers and Systems survey score by clinic type. The asterisks indicate the categories/domains relevant to the focus measures chosen for analysis. N/A: not applicable.

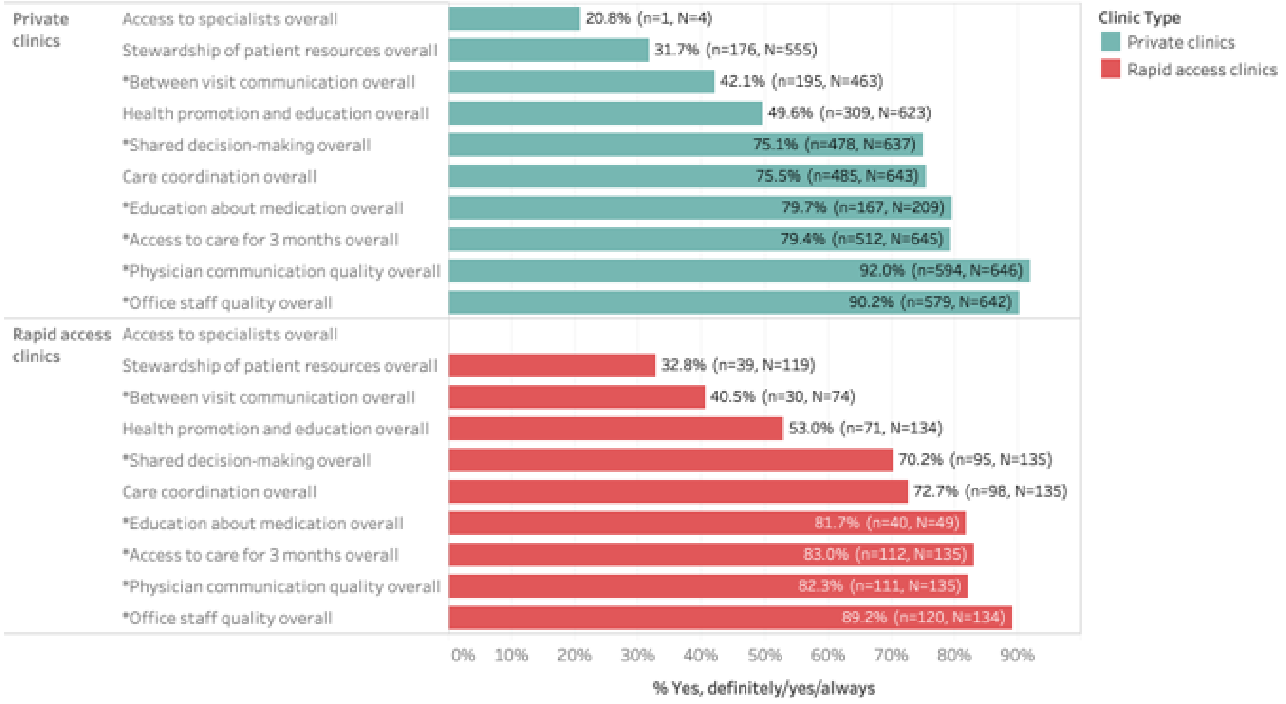


**Figure 7.** Focus measures for the private clinics.

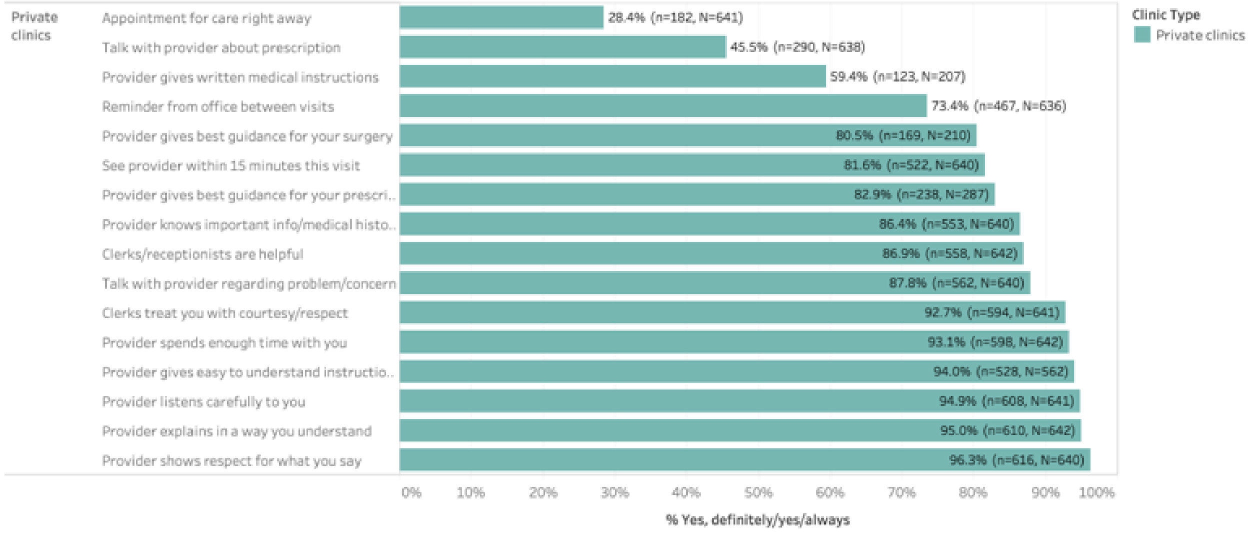
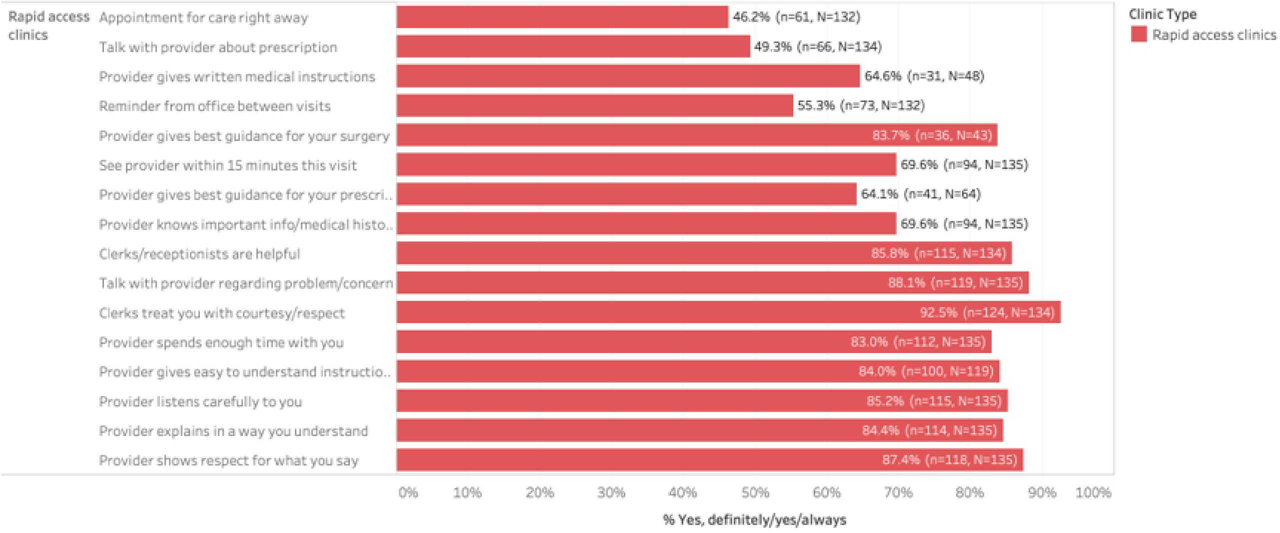
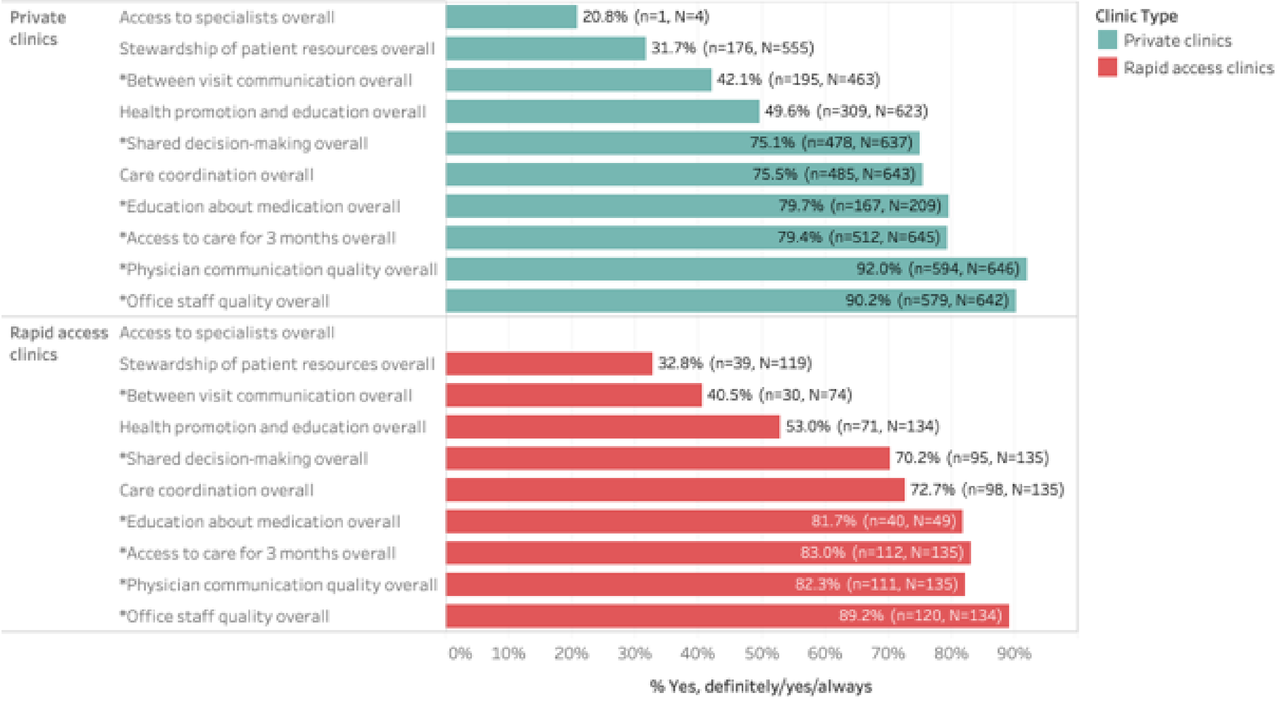


**Figure 8.** Focus measures for the rapid access clinics.



We present the findings for nonstandard focus measures by clinic type in Figure 9. The results showed that the private clinics scored higher than the rapid access clinics. The marginal difference in patient satisfaction between the two clinic types was 12%-15%. Although the primary goal was to find a pattern in patient satisfaction between the clinic types, it can be noted that there is room for improvement in both private and rapid access clinics regarding these measures.

**Figure 9.** Nonstandard focus measures by clinic type. Asterisks indicate the categories/domains relevant to the focus measures chosen for analysis.



## Qualitative Data Analysis Results

### *Round III: Nonstandard Quality Measures and Patient Comments*

We obtained the results of analyzing unstructured patient comments by clinic type. As listed in the Methods, the subthemes included the following: medical needs not met or questions not answered; lack of or delay with results or records; long wait or slow processing; appointment rushed or barely saw the doctor; lack of communication, direction, or follow-up; medical staff rude, impersonal, or did not listen; difficult to schedule timely appointment or appointment cancelled by the doctor; check in process frustrating, receptionist unhelpful, or lack of assistance; and general or other.

Figure 10 shows all patient comments by clinic type. It was found that the private clinics had a much higher rate of positive comments as compared with the rapid access clinics. On breaking down the amount each clinic type contributed to the overall comments for the positive, negative, indifferent, and not applicable categories, the private clinics appeared to contribute the most to each category. However, this finding is most likely influenced by the fact that there are more private clinics that run throughout the week and therefore a greater volume of comments associated with the private clinics. It is worth noting that there were 4454 comments from the private clinics and only 830 comments from the rapid access clinics.

We also assessed the distribution of the comments by clinic type within the sample (ie, the number of negative rapid access comments out of the total number of comments [N] for rapid access only; Figure 10). This analysis showed that the rapid access clinics had a greater makeup of negative comments as compared with the private clinics.

We expanded the negative comments by clinic, and we present the findings for the private clinics and rapid access clinics in Table 3. For the private clinics, the top three negative comments centered around “medical staff rude, impersonal, or did not listen,” “difficult to schedule timely appointment or appointment cancelled by the doctor,” and “check in process frustrating, receptionist unhelpful, or lack of assistance.” The “general or other” category included a variety of comments that did not appropriately fit into any other category (eg, singular words, such as “frustrating” and “disappointing,” without any context).

**Figure 10.** Patient comments by clinic (distributed in positive, negative, indifferent, and not applicable categories).

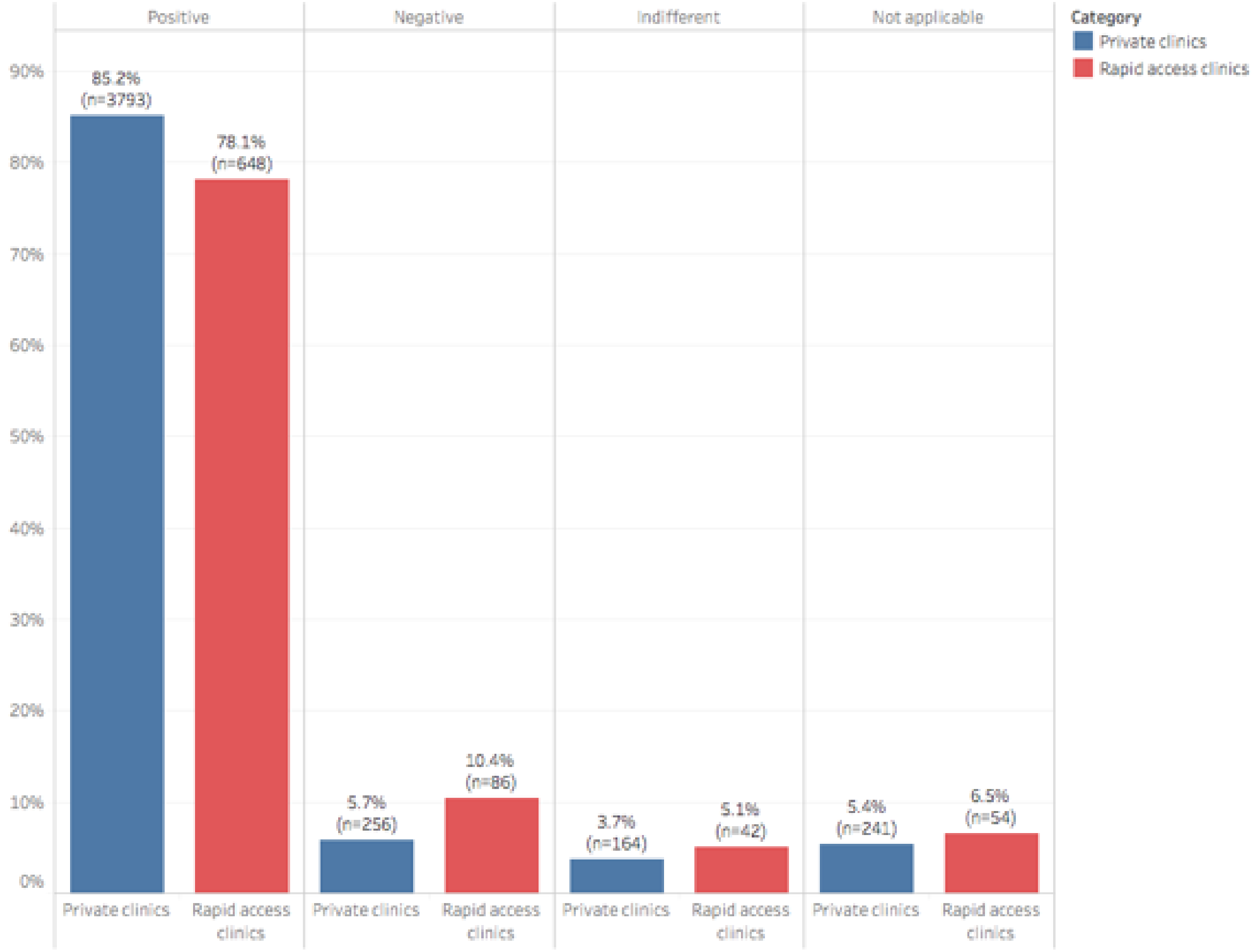


**Table 3.** Patient negative comments in the private clinics and rapid access clinics.

| Comments | Private clinics (N=256), n (%) | Rapid access clinics (N=86), n (%) |
|---|---|---|
| Medical needs not met or questions not answered | 13 (5.1%) | 7 (8.1%) |
| Lack of or delay with results/records | 15 (5.9%) | 2 (2.3%) |
| Long wait or slow processing | 18 (7.0%) | 14 (16.3%) |
| Appointment rushed or barely saw the doctor | 21 (8.2%) | 12 (14.0%) |
| Lack of communication, direction, or follow-up | 25 (9.8%) | 11 (12.8%) |
| Medical staff rude, impersonal, or did not listen | 28 (10.9%) | 15 (17.4%) |
| Difficult to schedule timely appointment or appointment canceled by the doctor | 32 (12.5%) | 5 (5.8%) |
| Check in process frustrating, receptionist unhelpful, or lack of assistance | 37 (14.5%) | 4 (4.7%) |
| General or other | 67 (26.2%) | 16 (18.6%) |

# Discussion

## Principal Findings

This study assessed two distinct dermatology clinics and their individual and collective contributions to CAHPS scores for the Dermatology Department at an AMC to answer the question, “Where should the efforts of patient satisfaction improvements be focused for the two different types of dermatology clinics?” This study utilized a concurrent mixed methods approach (Figure 1) to answer this question.

The clinics consisted of private clinics (patients make appointments in advance with the dermatologist of their choosing and likely see the same dermatologist with whom they have established a relationship) and rapid access clinics (patients can get walk-in or just-in-time appointments with a dermatologist or dermatology nurse practitioner on duty).

The quantitative data analysis showed that when only private clinic data were collected, there was a slight increase in the negative response rate and a slight decrease in the positive response rate across many categories, perhaps indicating that the contribution of rapid access CAHPS scores to the overall scores may be underestimated [2]. When individual clinic data were expanded, the majority of reviewed CAHPS measures had lower satisfaction for the rapid access clinics (9%-17%) as compared with the private clinics. The exception was “talk with provider about problem/concern,” for which the rate was 0.34% higher in the rapid access clinics when compared with the private clinics. Although some of these differences in scores were not substantial, they were consistent enough across the CAHPS categories and measures to give a thought to the individual clinic contribution.

To elaborate further on the qualitative data analysis, the value of patient comments cannot be underestimated as they provide straightforward and insightful feedback from patients. The comments directly express dermatology patients’ perceptions of their care while specifically indicating areas in need of improvement, as well as satisfactory aspects of the clinics. As such, consistent with the literature suggesting continued review of CAHPS scores, a frequent review of these comments could contribute to and increase the richness of aggregated data [5,7].

Both quantitative and qualitative data analyses were performed in this study with the goal of the qualitative data analysis findings enriching the quantitative data analysis findings. Integration of the quantitative and qualitative findings from this study suggests that the rapid access clinics contributed to low overall CAHPS scores, but the contribution was not radical. Qualitative analysis of patient comments elaborated the results of quantitative findings, showing that even when patients know they are in an environment of situational care (ie, a rapid access clinic), they want to spend more time with the clinician, want shorter wait times, and want people to listen. The consideration of these comments provided an additional, valuable, and actionable layer to complement the quantitative findings.

Regardless of the clinic, this study suggested that there is work to be done in regards to how staff or clerks interact with patients, as this aspect was scored low in both clinics. There is also work to be done around communication. Clinicians need to be especially conscientious to avoid speaking to their patients in a frustrated or condescending tone.

The difference between real and perceived wait times presents challenges for even the best clinic environment. A patient’s perception relative to the wait time could be altered by providing frequent updates on the wait time so as to provide patients with a more accurate perception of the time they are actually waiting, which may be a smaller window of time as compared to what they perceive without receiving any updates [10]. Overbooking patients can lead to communication issues. If there are too many patients, the clinician may spend less time with each patient, giving the patient a sense of being rushed and not being heard.

There is also work to be done around effectively managing delays in the clinic and getting to the root cause of those delays. If found that delays are caused by patients, it would be helpful to identify patients who are frequently late or who do not show up to the clinic. These particular patients could then be flagged and scheduled in the afternoon for future appointments. Scheduling such patients later in the day would minimize the interference with the clinic schedule and overall patient flow throughout the day [10].

Our findings show that appointments are much more difficult to get in private clinics and patients need to wait a long time to get an appointment with their established provider. This is consistent with the nationwide shortage of dermatologists [11].

### Limitations

This study had some limitations that need to be acknowledged. First, the data provided for this analysis by Press Ganey were already summarized, and this limited the statistical approaches that could be applied for further assessment. As such, analysis and result reporting were limited to descriptive statistics rather than more traditional statistical analysis to compare one dataset to another.

Another limitation was related to the survey design. The CAHPS survey does not have a not applicable ("N/A") response available as an option for any of the questions. Subsequently, when any of the questions do not apply to patients, which was stated in patient comments, they are forced to answer "no/never" instead, artificially increasing the negative response rate for certain measures as a result.

Additionally, inconsistent response sample sizes (N) for each of the measures likely contributed to skewed reporting. For example, not every patient completed the survey, and those who completed the survey did not always answer every question available (most likely when questions did not apply). This was verified through a review of the original survey copies that were available through Press Ganey. Several of the sample surveys reviewed were only partially filled out.

Lastly, patient comments were fewer for the rapid access clinics as compared with the private clinics. The investigation of methods to increase comments represents an area of future research for researchers.

Nevertheless, this study provides valuable findings related to the degree of the contribution of individual clinics to CAHPS scores.

### Conclusion

This study suggests that clinic type is relevant to patient satisfaction. However, the results also suggest that much of the dissatisfaction could be mitigated with better communication and level-setting expectations. There are gaps in research concerning the relationship between CAHPS scores and specific care factors. The findings and the common patterns between patient satisfaction measures and care components that were discovered through the execution of this analysis provide valuable insights into the drivers of patient satisfaction for not only the AMC's dermatology clinics going forward, but also ambulatory facilities in general.

Future research in this area could investigate physician performance relative to clinic type. Additionally, patient satisfaction by time of the day according to clinic type may provide useful insights into patient satisfaction.

## Acknowledgments

The authors would like to thank everyone in the Department of Radiology Oncology and Dermatology at the University of Alabama at Birmingham Health System, who helped to guide the focus of this study. The authors would like to deeply thank Shikha Modi, MBA, PhD candidate, for her tireless formatting of this manuscript to prepare it for publication. The authors acknowledge the detailed feedback of the reviewers, which helped make this a stronger paper.

## Conflicts of Interest

None declared.

## References

1. Agency for Healthcare Research and Quality. 2018. The CAHPS Program URL: http://www.ahrq.gov/cahps/about-cahps/cahps-program/index.html [accessed 2019-10-08]
2. Centers for Medicare & Medicaid Services. 2019. Consumer Assessment of Healthcare Providers & Systems (CAHPS) URL: https://www.cms.gov/research-statistics-data-and-systems/research/cahps/ [accessed 2019-09-22]
3. Eijkenaar F. Pay for performance in health care: an international overview of initiatives. Med Care Res Rev 2012 Jun;69(3):251-276. [doi: 10.1177/1077558711432891] [Medline: 22311954]
4. Audet A, Davis K, Schoenbaum SC. Adoption of patient-centered care practices by physicians: results from a national survey. Arch Intern Med 2006 Apr 10;166(7):754-759. [doi: 10.1001/archinte.166.7.754] [Medline: 16606812]
5. Merkouris A, Athini E, Hatzimbalasi M, Rovithis M, Papastavrou E. Assessment of patient satisfaction in public hospitals in Cyprus: a descriptive study. Health Science Journal 2013;7(1).
6. Safran DG, Karp M, Coltin K, Chang H, Li A, Ogren J, et al. Measuring patients' experiences with individual primary care physicians. Results of a statewide demonstration project. J Gen Intern Med 2006 Jan;21(1):13-21 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.00311.x] [Medline: 16423118]
7. Renedo A, Marston CA, Spyridonidis D, Barlow J. Patient and Public Involvement in Healthcare Quality Improvement: How organizations can help patients and professionals to collaborate. Public Management Review 2014 Feb 07;17(1):17-34. [doi: 10.1080/14719037.2014.881535]
8. Fung CH, Lim Y, Mattke S, Damberg C, Shekelle PG. Systematic review: the evidence that publishing patient care performance data improves quality of care. Ann Intern Med 2008 Jan 15;148(2):111-123. [doi: 10.7326/0003-4819-148-2-200801150-00006] [Medline: 18195336]

9. HIMSS Analytics. 2019. Patient experience study: shifting consumer demand and health system effort to address patient experience URL: https://enterprise.spectrum.com/insights/resources/reports/patient-experience-study-conducted-by-himss-analytics [accessed 2019-10-09]
10. Press Ganey Associates, Inc. Medical Practice Solutions Starter URL: https://www.premiermedicalhv.com/wp-content/uploads/2013/01/Press-Ganey-Medical-Practice-Solutions-Starter.pdf [accessed 2019-08-28]
11. Glazer AM, Rigel DS. Analysis of Trends in Geographic Distribution of US Dermatology Workforce Density. JAMA Dermatol 2017 May 01;153(5):472-473 [FREE Full text] [doi: 10.1001/jamadermatol.2016.6032] [Medline: 28296988]

## Abbreviations

**AMC:** academic medical center
**CAHPS:** Consumer Assessment of Healthcare Providers and Systems

---

*Edited by G Eysenbach; submitted 23.11.19; peer-reviewed by H Martin, H Gray, F Kaliyadan; comments to author 08.01.20; revised version received 02.03.20; accepted 09.04.20; published 12.05.20.*

*Please cite as:*
*Costigan J, Feldman SS, Lemak M*
*Assessment of Patient Satisfaction With Dermatology Clinics According to Clinic Type: Mixed Methods Study*
*JMIR Dermatol 2020;3(1):e17171*
*URL: http://derma.jmir.org/2020/1/e17171/*
*doi:10.2196/17171*
*PMID:*

Original Paper

# Consumer Preference of Products for the Prevention and Treatment of Stretch Marks: Systematic Product Search

Pengyi Zhu[1*], BS; Andrew Fung[1*], BSN, RN; Benjamin K P Woo[1*], MD

Olive View–University of California Los Angeles Medical Center, Sylmar, CA, United States

[*] all authors contributed equally

**Corresponding Author:**
Pengyi Zhu, BS
Olive View–University of California Los Angeles Medical Center
14445 Olive View Drive, Cottage H1
Sylmar, CA, 91342
United States
Phone: 1 626 817 1089
Email: sharonpzhu2@gmail.com

## Abstract

**Background:** Striae distensae, or stretch marks, are a common and distressing condition affecting females two-and-a-half times more frequently than males. Despite the numerous products available for stretch mark prevention and treatment, there have been few studies that consider consumer product preference.

**Objective:** The aim of this study was to determine which products were preferred by consumers for the prevention and treatment of stretch marks based on product vehicle and product ingredients.

**Methods:** In January 2020, a search was conducted on internet retailer Amazon for products related to stretch marks. The top products were identified as those with 100 reviews or greater and a rating of 4 or higher. The products were classified as either stretch mark–specific or non stretch mark–specific. Price, rating, type of vehicle, and specific ingredients of both product groups were compared. Vehicle-type and ingredients in both product groups were compared with two-tailed two-sample proportion tests to determine if certain vehicles or ingredients were more likely to be found in stretch mark–specific products. *P*<.05 indicated statistical significance.

**Results:** Out of over 10,000 products, 184 were selected as the top products according to the review and rating criteria of which 117 (63.6%) were stretch mark–specific and 67 (36.4%) were non stretch mark–specific. Oil was the most common vehicle (131/184, 71.2%) while vitamin E was the most common ingredient (58/184, 31.5%). Oil, as a vehicle, was more likely to be found in stretch mark–specific products than in non stretch mark–specific products (*P*=.001). Olive oil (*P*=.02) and cocoa butter (*P*=.08), *Centella asiatica* (*P*=.01), and shea butter (*P*=.003) were the ingredients more likely to be found in stretch mark–specific products than in non stretch mark–specific products.

**Conclusions:** This study demonstrated that there are many products available for the prevention and treatment of stretch marks and identified specific ingredients in the products preferred by customers. There are few studies investigating the effectiveness of the major ingredients in the stretch mark products that are preferred by consumers. Future studies can focus on the effectiveness of the ingredients found in the products that are preferred by consumers.

***(JMIR Dermatol 2020;3(1):e18295)*** doi:10.2196/18295

**KEYWORDS**

Amazon; consumer preference; stretch mark; striae distensae; striae gravidarum

## Introduction

Striae distensae, more commonly known as stretch marks, are a common and disfiguring dermatological condition in which the layer of dermis becomes scarred. Striae distensae occur two-and-a-half times more frequently in females than in males [1]. Striae distensae are more likely to develop during pregnancy (striae gravidarum), with obesity, with Cushing syndrome, during adolescence, with steroidal drug use, and after surgery [2].

The pathogenesis of striae distensae is multifactorial. Striae distensae are typically associated with rapid weight gain (such as in pregnancy) and rapid growth (such as in adolescence) as a result of tissue stretching [3]. Striae distensae are more common on the abdomen, breast, thigh, and buttocks [1,4]. Striae distensae caused by mechanical stress have been found to exhibit altered extracellular matrix components, primarily fibronectin, fibrillin, elastin, and collagen [5], through potential mechanisms such as low expression of genes for collagen and fibronectin, reduced fibroblast function, and disorganized fibrils or elastic fiber networks [3-5]. Studies of hormonal changes occurring in pregnancy, during steroid use, and in Cushing syndrome contribute to the understanding of the development of striae distensae; increased steroid levels cause increased protein catabolism and decreased production of collagen by fibroblasts resulting in less elastic fibers in the matrix of the dermis [1,6].

Because of the distress caused by the disfiguring nature of stretch marks, studies have attempted to find methods to prevent and treat stretch marks [7]. Different topical agents have been used to target processes that lead to the development of stretch marks. Trofolastin, a product which contains *Centella asiatica* was found to stimulate fibroblasts and counteract the glucocorticoid effect [8]. Oils and cocoa butter assist with maintaining skin hydration, which influences the expression of cytokines and growth factors via occlusive mechanisms [5,9]. Tretinoin (a derivative of vitamin A) has been found to stimulate fibroblasts to increase collagen I production in tissue while ascorbic acid (vitamin C) has also been found to increases collagen production [5]. Despite the numerous topical therapies available, therapy efficacy has not been well studied.

Because of the lack of evidence regarding the effectiveness of prevention or treatment of stretch marks, individuals look to other sources when determining which products to purchase. In this study, we assess which products are available to consumers, which are popular among consumers, and factors determining which products are used for stretch marks.

## Methods

A search was performed in January 2020 on United States internet retailer Amazon to identify the top products used for stretch marks. The searches performed were (1)“stretch mark” in All Departments: Beauty & Personal Care, (2)“stretch mark cream” in All Departments, and (3) “stretch mark” in Beauty & Personal Care: Skin Care: Maternity. Products were included based on average customer rating (4 stars and higher out of a possible 5 stars) and number of customer reviews (100 reviews and greater). Product titles, product descriptions, customer question and answer sections, and customer review sections were reviewed for the phrase “stretch mark.” Products were excluded if they were nontopical products such as masks, massage balls, massage cups, or socks; product sets with more than one type of item; or if there was no mention of “stretch mark” in the title, product description, customer question and answer section, or customer review section. The product listing with the highest number of reviews was included if the product had multiple listings for different quantities and sizes. Median rating, average rating, average number of reviews, vehicle-type, ingredient list, and median price per ounce or fluid ounce were determined for the overall sample of products. Because of the large variety of ingredients, only ingredients that had been investigated in previous studies for effectiveness in treating stretch marks were included in the analysis [3,5,10,11].

Products were divided into two groups—stretch mark–specific products were those that included “stretch mark” in the product title, product description, or both and non stretch mark–specific products were those that included “stretch mark” in the customer question and answer section or the customer review section. To determine if a specific vehicle-type or specific ingredient was more likely to be found in a stretch mark–specific product, two-tailed two-sample proportion tests were used. $P<.05$ indicated statistical significance.

## Results

Out of over 10,000 products on the United States internet retailer Amazon, 184 products met the required criteria. This group of products had a median consumer rating of 4.5 and a median of 409 reviews per product (range 103-10,573; total 183,366). The median price per ounce or per fluid ounce was $3.19 (range $0.17- $48; products that did not provide volume or weight information were not included: silicone sheet products, Crocodile acne scar removal cream, and Aliver TCM scar and acne mark removal gel ointment).

Products were categorized by vehicle-type. Out of 184 products, 131 (71.2%) were oils, 21 (11.4%) were creams, 9 (4.9%) were butters, 9 (4.9%) were gels, 5 (2.7%) were silicone sheets, 3 (1.6%) were lotions, 2 (1.1%) were scrubs, and the remaining 4 (2.2%) were categorized as other (balm, salve, ointment, and stick).

Of all 184 products, there were 117 (63.6%) stretch mark–specific and 67 (36.4%) non stretch mark–specific. Of the stretch mark-specific products, “stretch mark” was mentioned in the product title (6/117, 5.1%), product description (75/117, 64.1%), or both (36/117, 30.8%). “Stretch mark” was mentioned in the customer question and answer section (84/184, 45.7%) and customer review section (138/184, 75.0%). Customers mentioned “stretch mark” in relation to pregnancy or postpartum use in 70.1% of products (129/184). The top 10 most reviewed products are listed in Table 1. The top 10 most reviewed stretch mark–specific products are listed in Table 2.

Of the 117 stretch mark–specific products, 76 (65.0%) were oils, 17 (14.5%) were creams, 8 (6.8%) were butters, 6 (5.1%) were gels, 3 (2.6%) were silicone sheets, 3 (2.6%) were lotions, 2 (1.7%) were scrubs, 1 (0.9%) was a balm, and 1(0.9%) was a salve. Of the 67 non stretch mark–specific products, 55 (82.0%) were oils, 4 (6%) were creams, 1 (1.5%) was a butter, 3 (4.5%) were gels, 2 (3%) were silicone sheets, 1 (1.5 %) was an ointment, and 1 (1.5%) was a stick. Oil as a vehicle was more likely to be found in the stretch mark–specific product group than in the non stretch mark–specific product group ($P$=.001).

Of all 184 products, 7 (3.8%) contained allantoin, 11 (6%) contained aloe, 10 (5.4%) contained *C. asiatica*, 20 (10.9%) contained cocoa butter, 9 (4.9%) contained collagen, 4 (2.2%) contained hyaluronic acid, 23 (12.5%) contained shea butter, 3 (1.6%) contained vitamin A, 7 (3.8%) contained vitamin C, and

58 (31.5%) contained vitamin E. Out of the 117 stretch mark–specific products, 5 contained allantoin (4.3%), 9 contained aloe (7.7%), 10 (8.5%) contained *C. asiatica*, 18 (15.4%) contained cocoa butter, 7 (6%) contained collagen, 4 (3.4%) contained hyaluronic acid, 21 (17.9%) contained shea butter, 2 (1.6%) contained vitamin A, 6 (5.1%) contained vitamin C, and 40 (34.2%) contained vitamin E. Out of the 67 non stretch mark–specific products, 2 (3%) contained allantoin, 2 (3%) contained aloe, 0 (0%) contained *C. asiatica*, 2 (3%) contained cocoa butter, 0 (0%) contained hyaluronic acid, 2 (3%) contained shea butter, 1 (1.5%) contained vitamin A, 1 (1.5%) contained vitamin C, and 18 (26.9%) contained vitamin E. Three ingredients were more likely to be found in stretch mark–specific products—cocoa butter (*P*=.08), *C. asiatica* (*P*=.01), and shea butter (*P*=.003).

Specific types of oil as ingredients were also investigated [5]. While there were a variety of oils present in the products, almond oil, coconut oil, and olive oil have been found to be present frequently in other literature and were the focus of this analysis. Of all 184 products, 40 (21.7%) contained almond oil, 33 (17.9%) contained coconut oil, and 18 (9.8%) contained olive oil with remaining products containing other oils or no oils. Of the 117 stretch mark–specific products, 22 (18.8%) contained almond oil, 22 (18.8%) contained coconut oil, and 16 (13.7%) contained olive oil. Of the 67 non stretch mark–specific products, 18 (26.9%) contained almond oil, 11 (16.4%) contained coconut oil, and 2 (3%) contained olive oil. Olive oil was more likely to be found in stretch mark–specific products (*P*=.02).

**Table 1.** Top 10 most reviewed products.

| | Manufacturer name | Product name | Number of reviews | Mean rating (out of 5) |
|---|---|---|---|---|
| 1 | Pura D’or | Organic Jojoba Oil | 10,573 | 4.5 |
| 2 | Bio-Oil | Skincare Oil | 9732 | 4.4 |
| 3 | Sky Organics | Organic Castor Oil | 7164 | 4.5 |
| 4 | Now Solutions | Sweet Almond Oil | 6861 | 4.7 |
| 5 | M3 Naturals | Anti Cellulite Massage Oil | 5048 | 4.6 |
| 6 | First Botany Cosmeceuticals | Arabica Coffee Scrub | 4839 | 4.0 |
| 7 | Burt's Bees Mama Bee | Belly Butter | 4660 | 4.4 |
| 8 | Pura D’or | Organic Rosehip Seed Oil | 4332 | 4.6 |
| 9 | Ancient Greek Remedy | Ancient Greek Remedy Oil | 4249 | 4.6 |
| 10 | Honeyskin | Ultimate Face & Body Cream | 3359 | 4.3 |

**Table 2.** Top 10 most reviewed stretch mark–specific products.

| | Manufacturer name | Product name | Number of reviews | Mean rating (out of 5) |
|---|---|---|---|---|
| 1 | Bio-Oil | Skincare Oil | 9732 | 4.4 |
| 2 | Sky Organics | Organic Castor Oil | 7164 | 4.5 |
| 3 | M3 Naturals | Anti Cellulite Massage Oil | 5048 | 4.6 |
| 4 | First Botany Cosmeceuticals | Arabica Coffee Scrub | 4839 | 4.0 |
| 5 | Burt's Bees Mama Bee | Belly Butter | 4660 | 4.4 |
| 6 | Ancient Greek Remedy | Ancient Greek Remedy Oil | 4249 | 4.6 |
| 7 | Honeyskin | Ultimate Face & Body Cream | 3359 | 4.3 |
| 8 | Coco & Co. | The Purest Coconut Virgin Oil | 2858 | 4.7 |
| 9 | Palmer's Cocoa Butter Formula | Massage Lotion For Stretch Marks | 2654 | 4.3 |
| 10 | Majestic Pure Cosmeceuticals | Rosehip Oil | 2647 | 4.4 |

## Discussion

### Principal Findings

Over 10,000 products for stretch marks were found on Amazon, indicating the broad and diverse array of products ranging from topical to physical products (such as masks and socks) that are available for consumers. In this study, only topical products were assessed to limit the discussion of effectiveness to certain ingredients (those previously studied), but even with past studies, there has been no consensus regarding the effectiveness of topical products. As such, consumers rely upon information and suggestions posted by consumers on internet retail websites such as Amazon.

There is a disparity between the characteristics of products that are intended for use on stretch marks and the characteristics of products that are used on stretch marks based on customer reviews. Out of all 184 products, 63.6% (117) were intended for use on stretch marks, while 36.4% (67) were not intended for stretch marks but were used nonetheless based on customer reviews or customer questions and answers. This indicated that customers may be purchasing ineffective products by relying upon the opinions of others; 45.7% (84/184) of products had a question regarding stretch marks in the question and answer section, 75% (138/184) of products mentioned “stretch marks” in the reviews, and 70.1% (129/184) of products mentioned “stretch marks” related to pregnancy. The number of products that had a combination of “stretch marks” being mentioned in either question and answer section, reviews, and in relation to pregnancy was not differentiated. Previous studies have also shown that patients seek advice from friends, product advertisements, and the internet, particularly in the early stages of pregnancy [12].

The majority of the top products that were found when searching “stretch mark” were oils—131 (71.2%) of the top 184 products. Stretch mark–specific products were more likely to be oils than non stretch mark–specific products as demonstrated by the two-sample proportion test. Oils help retain moisture in the skin, but oils rich in linoleic acid and polyunsaturated fatty acids, such as olive oil, have also been found to reduce the formation of scars by reducing pro-inflammatory mediators (tumor necrosis factor alpha and peroxisome proliferator-activated receptors) which assists with wound healing [13]. Other studies, however, have produced conflicting findings. Studies have found that olive oil does not significantly reduce the incidence of striae gravidarum [14,15], while a study on almond oil demonstrated its possible effectiveness through its role in decreasing the progression of striae gravidarum by preventing itchiness when massaged into the skin [16,17]. In our study, specific oils that were found to be ingredients in the top products included, but were not limited to, apricot oil, argan oil, avocado oil, canola oil, castor oil, emu oil, grapefruit oil, jojoba oil, safflower oil, and sunflower oil. Recent studies have focused on almond oil, coconut oil, and olive oil [5]. This raises the question if enough studies on oils have been conducted, particularly because there have been contradictory findings on the effectiveness reported by studies. The amount of available studies on oils does not reflect the popularity of oils among consumers for stretch mark prevention and treatment. Furthermore, several studies suggest the therapeutic effect of massage in relation with oil for the prevention of stretch marks [1,16]. This further indicates that there is not enough understanding concerning oils and their use for stretch marks.

Out of the major ingredients assessed in this study, vitamin E was the most commonly used ingredient in both stretch mark–specific (40/117, 34.2%) and non stretch mark–specific products (18/67, 26.9%). A prior study showed that olive oils were rich in vitamin E, yet the findings did not suggest that olive oils were effective in the prevention of striae gravidarum [14]. Almond oil is also rich in vitamin E and vitamin D, yet a previous study also did not demonstrate the effectiveness of almond oil on its own for striae gravidarum [16]. Thus, future studies may wish to continue to assess which ingredients are effective for stretch mark prevention.

Creams comprised 11.4% (21/184) of the top products; creams comprised 14.5% (17/117) of stretch mark–specific products and 6% (4/67) of non stretch mark–specific products. Creams have been found to be the type of product most used by pregnant patients for prevention of stretch marks [4]. Previous studies stated an incidence of 50% to 90% for the development of striae gravidarum [18], indicating a large population of individuals who may seek out striae gravidarum products. Another cross-sectional study demonstrated high use of Bio-Oil by patients for the prevention of striae gravidarum [12], which is consistent with our finding that Bio-Oil was the most reviewed stretch mark–specific product. Product effectiveness and safety are of concern, especially during pregnancy. Furthermore, cream treatments such as Alphastria (containing hyaluronic acid, allantoin, vitamin A, and vitamin E) and Verum (containing vitamin E and hyaluronic acid) were not found to be in the top 10 product list, nor were their ingredients found in the top 10 products, despite having been studied and shown to have possible benefits in stretch mark prevention [1,5]. This demonstrates a discrepancy between the products that are being studied and the products that consumers are actually using, suggesting the need for consumers to be educated.

We focused our study on ingredients that had been mentioned in previous studies [3,5,10,11]. Cocoa butter, *C. asiatica*, and shea butter were more likely to be found as ingredients in products intended for stretch marks. In a previous study, the findings regarding the efficacy of cocoa butter for striae gravidarum prevention did not support its use despite cocoa butter’s moisturizing effects [6,19]. Trofolastin (containing *C. asiatica*, vitamin E, and collagen-elastin hydrolysates) was also studied for the prevention of striae gravidarum and demonstrated significant prevention of stretch marks during pregnancy likely due to fibroblast stimulation and glucocorticoid antagonism [1,5,8]. Though shea butter is known to reduce fibroblasts [3], and there have been studies on whether shea butter is or is not effective for stretch marks, it is more likely to be found in products intended for stretch marks. Other products may have studies indicating their efficacy but may not be specifically indicated for the prevention of stretch marks. Aloe vera has been studied for its effectiveness in decreasing itchiness, and thus, the progression of stretch marks [17]. Tretinoin (or other vitamin A derivatives) has demonstrated some efficacy in preventing stretch marks, but is not found primarily in stretch mark–specific products [6]. Furthermore, tretinoin as well as ascorbic acid and hyaluronic acid are ingredients known to promote production of collagen [5]. Hyaluronic acid increases the skin’s ability to resist tension, likely due to a mechanism of increased collagen production, yet minimal evidence has been found for its effectiveness [5,6]. These discrepancies show the gaps between what is effective and what is communicated to consumers about a product’s effectiveness. Further studies should focus the ingredients being used the most by consumers in order to understand the ingredient’s mechanism and effectiveness in stretch mark prevention.

## Limitations

One limitation was the size of the overall study. There were over 10,000 products listed for “stretch mark” when searching the online retailer Amazon; however, we only included the top products by screening ratings and number of reviews. As such, we did not assess products with lower ratings and less reviews which would have been able to give us a broader idea of which products were being used by consumers for stretch marks. We also excluded products that did not mention “stretch mark” in the title, description, question and answer section, or review section, thus we did not assess what other types of ingredients were potentially being used for stretch marks. Another limitation of our study was that while we considered products that mentioned “stretch mark” in the product review, we did not assess whether the review was positive or negative. Future studies could potentially compare positive and negative consumer product reviews to understand the subjective effectiveness of products.

## Conclusions

As demonstrated from our search on the internet retailer Amazon for products related to stretch marks, there are a plethora of products and ingredients for stretch marks; however, with few studies on the effectiveness of ingredients and products, consumers may rely upon consumer reviews when selecting products. More studies are needed to determine effectiveness and safety of ingredients, particularly with respect to pregnant women.

## Conflicts of Interest

None declared.

## References

1. Elsaie ML, Baumann LS, Elsaaiee LT. Striae distensae (stretch marks) and different modalities of therapy: an update. Dermatol Surg 2009 Apr;35(4):563-573. [doi: 10.1111/j.1524-4725.2009.01094.x] [Medline: 19400881]
2. Tsuji T, Sawabe M. Elastic fibers in striae distensae. J Cutan Pathol 1988 Aug;15(4):215-222. [doi: 10.1111/j.1600-0560.1988.tb00547.x] [Medline: 3183137]
3. Ud-Din S, McGeorge D, Bayat A. Topical management of striae distensae (stretch marks): prevention and therapy of striae rubrae and albae. J Eur Acad Dermatol Venereol 2016 Feb 20;30(2):211-222 [FREE Full text] [doi: 10.1111/jdv.13223] [Medline: 26486318]
4. Al-Himdani S, Ud-Din S, Gilmore S, Bayat A. Striae distensae: a comprehensive review and evidence-based evaluation of prophylaxis and treatment. Br J Dermatol 2014 Mar 12;170(3):527-547. [doi: 10.1111/bjd.12681] [Medline: 24125059]
5. Lokhande A, Mysore V. Striae distensae treatment review and update. Indian Dermatol Online J 2019;10(4):380. [doi: 10.4103/idoj.idoj_336_18]
6. Korgavkar K, Wang F. Stretch marks during pregnancy: a review of topical prevention. Br J Dermatol 2015 Mar 08;172(3):606-615. [doi: 10.1111/bjd.13426] [Medline: 25255817]
7. Yamaguchi K, Suganuma N, Ohashi K. Prevention of striae gravidarum and quality of life among pregnant Japanese women. Midwifery 2014 Jun;30(6):595-599. [doi: 10.1016/j.midw.2013.07.011] [Medline: 23962637]
8. Mallol J, Belda M, Costa D, Noval A, Sola M. Prophylaxis of striae gravidarum with a topical formulation. a double blind trial. Int J Cosmet Sci 1991 Feb;13(1):51-57. [doi: 10.1111/j.1467-2494.1991.tb00547.x]
9. Rawlings AV, Bielfeldt S, Lombard KJ. A review of the effects of moisturizers on the appearance of scars and striae. Int J Cosmet Sci 2012 Dec 21;34(6):519-524. [doi: 10.1111/j.1468-2494.2012.00751.x] [Medline: 22994859]
10. Ross NA, Ho D, Fisher J, Mamalis A, Heilman E, Saedi N, et al. Striae distensae: preventative and therapeutic modalities to improve aesthetic appearance. Dermatol Surg 2017 May;43(5):635-648. [doi: 10.1097/DSS.0000000000001079] [Medline: 28375972]
11. Hague A, Bayat A. Therapeutic targets in the management of striae distensae: A systematic review. J Am Acad Dermatol 2017 Sep;77(3):559-568.e18. [doi: 10.1016/j.jaad.2017.02.048] [Medline: 28551068]
12. Brennan M, Clarke M, Devane D. The use of anti stretch marks' products by women in pregnancy: a descriptive, cross-sectional survey. BMC Pregnancy Childbirth 2016 Sep 21;16(1):276 [FREE Full text] [doi: 10.1186/s12884-016-1075-9] [Medline: 27654661]
13. Bielfeldt S, Blaak J, Staib P, Simon I, Wohlfart R, Manger C, et al. Observer-blind randomized controlled study of a cosmetic blend of safflower, olive and other plant oils in the improvement of scar and striae appearance. Int J Cosmet Sci 2018 Feb 12;40(1):81-86. [doi: 10.1111/ics.12438] [Medline: 29094366]
14. Taavoni S, Soltanipour F, Haghani H, Ansarian H, Kheirkhah M. Effects of olive oil on striae gravidarum in the second trimester of pregnancy. Complement Ther Clin Pract 2011 Aug;17(3):167-169. [doi: 10.1016/j.ctcp.2010.10.003] [Medline: 21742284]
15. Soltanipoor F, Delaram M, Taavoni S, Haghani H. The effect of olive oil on prevention of striae gravidarum: a randomized controlled clinical trial. Complement Ther Med 2012 Oct;20(5):263-266. [doi: 10.1016/j.ctim.2012.05.001] [Medline: 22863639]
16. Timur Taşhan S, Kafkasli A. The effect of bitter almond oil and massaging on striae gravidarum in primiparaous women. J Clin Nurs 2012 Jun 22;21(11-12):1570-1576. [doi: 10.1111/j.1365-2702.2012.04087.x] [Medline: 22594386]

17. Hajhashemi M, Rafieian M, Rouhi Boroujeni HA, Miraj S, Memarian S, Keivani A, et al. The effect of Aloe vera gel and sweet almond oil on striae gravidarum in nulliparous women. J Matern Fetal Neonatal Med 2018 Jul 19;31(13):1703-1708. [doi: 10.1080/14767058.2017.1325865] [Medline: 28521546]
18. Chang ALS, Agredano YZ, Kimball AB. Risk factors associated with striae gravidarum. J Am Acad Dermatol 2004 Dec;51(6):881-885. [doi: 10.1016/j.jaad.2004.05.030] [Medline: 15583577]
19. Osman H, Usta IM, Rubeiz N, Abu-Rustum R, Charara I, Nassar AH. Cocoa butter lotion for prevention of striae gravidarum: a double-blind, randomised and placebo-controlled trial. BJOG 2008 Aug;115(9):1138-1142 [FREE Full text] [doi: 10.1111/j.1471-0528.2008.01796.x] [Medline: 18715434]

---

*Edited by G Eysenbach; submitted 17.02.20; peer-reviewed by S Shu, T Dunnsiri; comments to author 27.03.20; revised version received 01.04.20; accepted 09.04.20; published 04.06.20.*

*Please cite as:*
*Zhu P, Fung A, Woo BKP*
*Consumer Preference of Products for the Prevention and Treatment of Stretch Marks: Systematic Product Search*
*JMIR Dermatol 2020;3(1):e18295*
*URL: http://derma.jmir.org/2020/1/e18295/*
*doi:10.2196/18295*
*PMID:*

Original Paper

# Delivering Clinical Skin Examination Education to Nurse Practitioners Using an Internet-Based, Microlearning Approach: Development and Feasibility of a Video Intervention

Delaney B Stratton[1], PhD, DNP, FNP-BC; Kimberly D Shea[1], RN, PhD, CHPN; Elizabeth P Knight[2], PhD, DNP, FNP-C; Lois J Loescher[1], RN, PhD, FAAN

[1]College of Nursing, The University of Arizona, Tucson, AZ, United States

[2]College of Nursing, Oregon Health & Science University, Portland, AZ, United States

**Corresponding Author:**
Delaney B Stratton, PhD, DNP, FNP-BC
College of Nursing
The University of Arizona
1305 N Martin Ave
Tucson, AZ, 85721
United States
Phone: 1 4802136556
Email: dstratton@email.arizona.edu

## Abstract

**Background:** Skin cancer is the most common cancer; survival of the most serious skin cancers and malignant melanomas depends on early detection. Early detection relies on accessibility to clinical skin examination (CSE). Primary care nurse practitioners (PCNPs) are well-positioned to conduct CSEs; however, they require further education on CSE and have time constraints for continuing education. A digitally delivered intervention grounded in microlearning is a promising approach to deliver new information over a brief period.

**Objective:** Our objective was to develop and explore the feasibility of implementing a 1-week digital video intervention with content on CSE skills, defined as melanoma risk assessment, head-to-toe skin examination, and pigmented lesion assessment, for PCNPs. Specific aims were as follows: (1) Aim 1: to develop three microlearning-based melanoma videos with content on CSE that are suitable for digital delivery to PCNPs in various formats and (2) Aim 2: to assess the feasibility of the video intervention, including enrollment and retention rates, adherence, and acceptability and usability of the video intervention.

**Methods:** For Aim 1, the research team created storyboards for videos that addressed each CSE skill. An expert panel of three dermatologists reviewed the storyboards and videos for relevance, comprehension, and clarity using the content validity index (CVI). The panel evaluated the usability of the video intervention delivery by Research Electronic Data Capture (REDCap) and Vimeo using the System Usability Scale (SUS) and technical video production using Beaudin and Quick's Quality Evaluation of Video (QEV). Aim 2 evaluated enrollment and retention rates of PCNPs, based on metrics from previous studies of CSE in the literature, and video intervention adherence. SUS and the Attitudes toward Web-based Continuing Learning Survey (AWCL) assessed usability and acceptability.

**Results:** CVI scores indicated relevance and clarity for each video: mean scores ranged from 3.79 to 4, where 4 indicated the video was highly relevant and very clear. The integration of REDCap and Vimeo was usable: the SUS score was 96, where 0 was the worst and 100 was the best. The digital delivery of the videos was rated as exceptional on all five technical items: the mean score was 5, where scores ranged from 1 (poor) to 5 (exceptional). Of the 32 PCNPs who were sent emails, 12 enrolled (38%) and, out of these 12, 10 (83%) completed the intervention and the surveys. Video intervention adherence was ≤50%. Participants rated the usability as *better* (mean 85.8, SD 10.6; *better*=70-90) and favorably ranked the acceptability of the AWCL's constructs of *perceived usefulness* (mean 5.26, SD 0.08), *perceived ease of use* (mean 5.40, SD 0.41), *behavior* (mean 5.53, SD 0.12), and *affection* (mean 5.77, SD 0.04), where scores ranged from 1 (strongly disagree) to 7 (strongly agree).

**Conclusions:** The video intervention was feasible to deliver to PCNPs using a digital, microlearning approach. The findings provide support for using the videos as an intervention in a future pilot randomized trial targeting behavioral CSE outcomes among PCNPs and other primary care providers.

***(JMIR Dermatol 2020;3(1):e16714)*** doi:10.2196/16714

**KEYWORDS**
microlearning; clinical skin examination; primary care; nurse practitioners; melanoma; skin cancer

# Introduction

## Background

The incidence of the deadliest skin cancer, melanoma, has doubled in the United States over the past 20 years [1]. An estimated 96,480 new cases of melanoma were diagnosed in 2019, and 7320 deaths resulted from melanoma [2]. Early detection leads to a greater proportion of removal of thin melanomas (<2 mm in thickness), which is associated with better outcomes [3-5]. Early detection is best accomplished by clinical skin examination (CSE), most often performed by a dermatologist. CSE is defined as melanoma risk assessment, head-to-toe skin examination, and pigmented lesion assessment. Currently, there is a shortage of dermatologists—the United States averages just 3.6 dermatologists per 100,000 people [6,7], which has not changed in the last 30 years even as skin cancer incidence continues to rise [8-10]. Patients could benefit from the availability of more practitioners, such as nurse practitioners (NPs), to conduct quality CSEs.

Previous research indicates that NPs may lack confidence and skills to perform CSE or assess skin lesions [11-14]. NPs have demonstrated mixed ability to distinguish lesions that are suspicious for melanoma from nonsuspicious lesions [11,12]; in one study, a majority of NPs stated that they would rather refer any skin lesion to a specialist [15], which is problematic given the shortage of dermatologists. Although NPs' confidence in CSE is low, they believe that primary care providers, in general, help to detect skin cancer early [13]. In a pilot study, NPs showed promise for cultivating good sensitivity (ranging from 50% to 100%, n=4) and excellent specificity (ranging from 99% to 100%, n=4) when asked to identify suspicious lesions [16].

NPs desire more training and resources about skin cancer training [17]. However, a recent systematic review [12] concluded that there are minimal CSE activities for NPs, and the activities that exist are not well explicated. The majority of NP CSE-focused intervention studies had small sample sizes (eg, ranging from 1 to 30) [18-21], were lengthy (eg, 14 weeks to 6 months) [18], or did not describe NPs' dermatological training [18-21]. Fewer interventions had modules lasting under an hour (eg, 15-45 minutes) or were self-directed (eg, reviewing pamphlets, posters, and two presentations) [22-24]. This manuscript describes the development and feasibility testing of a brief CSE educational video intervention for primary care nurse practitioners (PCNPs).

## Purpose Statement

The purpose of this study was to develop and explore the feasibility of implementing a 1-week, digitally delivered video intervention with content on CSE for PCNPs. The videos covered three CSE skills, defined as melanoma risk assessment, head-to-toe skin examination, and pigmented lesion assessment. The specific aims were as follows: (1) Aim 1: to develop three microlearning-based melanoma videos with content on CSE that are suitable for digital delivery to PCNPs in various formats and (2) Aim 2: to assess the feasibility of the video intervention, including enrollment and retention rates, video intervention adherence, and acceptability and usability postcompletion. Table 1 presents the details of the aims, hypotheses, measures, scoring [25-30], and outcomes.

**Table 1.** Aims, measures and tools, scoring, and outcomes of the clinical skin examination (CSE) educational video intervention for primary care nurse practitioners (PCNPs).

| Aim with subaim or hypothesis | Measures or tools | Scoring | Outcomes |
|---|---|---|---|
| **Aim 1: to develop, over 3 months, three theory-based, short skin cancer videos with content on comprehensive CSE skills that are suitable for digital delivery to PCNPs in various formats (eg, mobile phone, tablet, and computer)** | | | |
| Aim 1a: to assess content validity of the video intervention using an established method | Content validity index (CVI) | Relevance: 1 (not relevant) to 4 (highly relevant)<br>Clarity: 1 (not clear) to 4 (very clear) [25-27] | Dermatology experts score content relevance and clarity highly |
| Aim 1b: to assess the integration of the videos and surveys into Research Electronic Data Capture (REDCap) | System Usability Scale (SUS) | 1 (strongly disagree) to 5 (strongly agree)<br>Sum of item scores is calculated and multiplied by 2.5<br>Scores: concerning (<50), passable (50-69), better (70-90), and truly superior (>90) [28] | Dermatology experts score usability as *better* to *truly superior* |
| Aim 1c: to assess the digital delivery of the videos | Beaudin and Quick's Quality Evaluation of Video (QEV) | 1 (poor) to 5 (exceptional)<br>Option for open-ended comments follows each item [29] | Dermatology experts give high scores for technical production and navigation (ie, video design, intended content, visual quality, audio quality, and audio-visual relationship) |
| **Aim 2: to determine enrollment and retention rates, video intervention adherence, and acceptability and usability postcompletion** | | | |
| Hypothesis 2.1: enrollment rates will be equal to or better than 60% | Calculate the number of participants recruited and enrolled compared to those recruited who chose not to consent or enroll | N/A[a] | Enrollment rates |
| Hypothesis 2.2: retention rates will be greater than 50% | Calculate completion of videos and surveys per number of participants enrolled | N/A | Retention rates |
| Hypothesis 2.3: video intervention adherence will be greater than or equal to 50% | Vimeo "finishes": counted from start-to-play to viewing of the very last video frame | N/A | Participant completion rates of the videos |
| Hypothesis 2.4: usability scores will be equal to or higher than 70 | System Usability Scale (SUS) | 1 (strongly disagree) to 5 (strongly agree)<br>Sum of item scores is calculated and multiplied by 2.5<br>Scores: concerning (<50), passable (50-69), better (70-90), and truly superior (>90) [28] | Participants score usability highly |
| Hypothesis 2.5: acceptability scores will be equal to or higher than 5 | Attitudes toward Web-based Continuing Learning Survey (AWCL)<br>Acceptability: *perceived usefulness* (5 items), *perceived ease of use* (4 items), *behavior* (3 items), and *affection* (3 items) | 1 (strongly disagree) to 7 (strongly agree) [30] | Participants score acceptability highly |

[a] N/A: not applicable.

# Methods

The study design was a one-group, posttest, cross-sectional design. This section describes methods for each study aim. The University of Arizona Institutional Review Board approved the study.

## Aim 1: CSE Video Development

Video development was guided by the use of a microlearning conceptual framework and operational transparency. The informational content of the videos was adapted from previous studies [31-33]; adaptation focused on key concepts that could be addressed in a short amount of time [34]. Development also included creation of the video storyboards.

Microlearning allows for the dissemination of short, meaningful knowledge, which benefits practitioners with busy schedules. Microlearning is defined as “special moments or episodes of learning while dealing with specific tasks or content and engaging in small but conscious steps” [35]. Research findings document that the use of short content may increase information retention by 20% [36]. Microlearning is for users who have difficulty creating the time to engage in long stretches of learning activities outside of dedicated study times and institutional programs [34].

For operational transparency, we conducted a systematic review to ascertain the rigor of previous CSE interventions [37] using Sidani and Braden’s clarifying elements (see Multimedia Appendix 1) [38]. The goal of the intervention was to inform the participants about CSE, enhance their CSE skills, and motivate them to perform CSE. Each of the three videos had a specific learning objective. The specific strategies, respective components, and immediate goals for each video are in Multimedia Appendix 2. The media for intervention delivery were written (ie, reading) and verbal (ie, audio). The format consisted of a video of skills instruction along with a PowerPoint presentation with voiceover. The videos were each 5-10 minutes in length (ie, amount). Vimeo was used to house the videos; this is an online platform and community developed to create, upload, and share videos [39]. Participants viewed each video one time and spaced video viewings within a 1-week period (ie, frequency); therefore, the duration of the intervention was 1 week.

The research team created and reviewed storyboards for each video. The videos were produced in collaboration with video technology experts at the institution and were uploaded to Vimeo. The web application Research Electronic Data Capture (REDCap) maintained the surveys and the separate fields for each video link from Vimeo. REDCap is a secure workflow methodology and software application designed for the development and deployment of digital data capture tools to assist with clinical and translational research [40]. REDCap allowed placement of the Vimeo video link into the survey, enabling viewing of the video within the survey, without having to open a new browser window. The expert panel, which consisted of three dermatologists, evaluated the integration of the videos in REDCap.

The components of each module were assessed using Sidani and Braden’s content validity assessment [41] and the content validity index (CVI). The first content validity survey asked the dermatology experts to evaluate the relevance (ie, the degree to which the content has an appropriate sample of activities for the component being measured) and clarity (ie, the extent to which the storyboard is concise, accurate, and direct) of the storyboard content [42]. The dermatology experts could add comments with each item to provide further clarification. Based on the CVI scores and recommendations, video content and activities were refined. The dermatology expert panel reviewed the content validity for a second time while viewing the actual videos 1 month after the first review. They accessed the videos in REDCap and completed the System Usability Scale (SUS).

The five technical production items from Beaudin and Quick’s Quality Evaluation of Video (QEV) evaluated the integration of the videos into Vimeo. The following steps helped to confirm that the REDCap and Vimeo platform was functioning appropriately for the delivery of the surveys and videos:

1. Assemble all surveys and videos into REDCap.
2. Set the timeline for the delivery of each video intervention within REDCap.
3. Conduct an initial test of REDCap to ensure that surveys and videos displayed appropriately and when prompted by the scheduled timeline on both the mobile device and computer.
4. Finalize survey and video delivery schedule.

## Aim 2: Feasibility

Feasibility was assessed via enrollment and retention rates as well as video intervention adherence, usability, and acceptability (see Table 1). Video adherence was monitored by Vimeo, which calculated the number of plays, number of finishes (ie, participant viewed to the very last video frame), and average percentage of the video watched per module [43]. Participants completed a short satisfaction survey at the end of the study, which consisted of free-text and scaled items.

## Sample

A purposive sample of 12 PCNPs was enrolled and 10 PCNPs completed the videos. A sample size of 10 is sufficient for a feasibility study because even a few cases are likely to be very informative with respect to the difficulty of recruitment, the acceptability of the intervention, costs, and logistics [44,45]. Eligibility criteria were as follows:

1. Had a Masters NP Certification or a Doctorate of Nursing Practice with clinical specialty.
2. Had Family, Adult, or Geriatric NP board certification from either the American Nurses Credentialing Center (ANCC) or the American Association of Nurse Practitioners (AANP).
3. Worked in an outpatient setting at least 16 hours/month or 192 hours/year.
4. Had a minimum of one year of experience.
5. Had access to the internet through a computer or a mobile phone.
6. Had English-language proficiency.

Participants were not excluded from the study based on gender, age, or race. Exclusion criteria were any previous skin cancer continuing education in CSE or training in CSE. Participants were recruited during a statewide meeting (ie, Arizona) and a local NP meeting (see Table 2 for demographic information). Interested NPs received an email that summarized the feasibility study along with a link to the consent form, surveys, and first video. Each email contained a unique link to the intervention. REDCap creates closed surveys, where each unique link is assigned a study ID number. The consent document was a disclosure form that listed the intervention length, investigator identity, and the purpose of the study (see Multimedia Appendix 3).

### Data Collection

All data instruments, automated data capture, videos, and contact information were managed using REDCap. The schedule for delivering each video and the posttest survey were automated with REDCap. Human involvement was limited to in-person recruitment and sending the initial email to start the intervention. The videos were incorporated into the questionnaire. Participants completed the posttest survey, along with the usability and acceptability items, after the third video. A total of 1 week after the third video, participants self-reported personal use of CSE in their practice. All outcomes were self-reported though these online surveys and all questions had forced-choice answers. Participants did not have an option to review or change their answers after submission. Prompts were sent daily for up to 2 days if the participant did not view the video within the first 24 hours of enrollment. Participants who failed to submit the survey were not compensated, and the surveys were ineligible for analysis. Each participant who completed the intervention received a US $50 Amazon gift card. Data collection occurred from March to April 2019.

### Data Protection

REDCap was developed specifically around the Health Insurance Portability and Accountability Act (HIPAA) Security Rule guidelines. The REDCap electronic data management system at the University of Arizona is housed on two virtual servers: one supporting database services and the other supporting web services. Hardware is located at the University of Arizona’s Information Technology Services Center (UITS). The space is physically secured within a keyless entry area. Hardware management and support are provided by UITS. The database server is located behind a firewall and the web server is in a Data Management Zone. REDCap software support is provided by the University of Arizona Center for Biomedical Informatics and Biostatistics. All web-based information transmission is password protected and encrypted in transit. Administration of REDCap is managed through virtual servers located at the University of Arizona College of Medicine [46].

### Statistical Analysis

Data from REDCap were exported into SPSS Statistics for Windows, version 26.0 (IBM Corp) [47], for data analysis. Analysis of sample demographic data consisted of descriptive statistics (ie, frequencies and measures of central tendency). For Aim 1a, the CVI was determined by dividing the number of dermatology experts giving the fact or item a score of 3 or 4 and dividing this by the total number of experts (ie, 3) [26]. Scores for the SUS and the QEV were analyzed using descriptive statistics (Aims 1b and 2). For Aim 2, the Attitudes toward Web-based Continuing Learning Survey (AWCL) was used and data analysis consisted of item mean scores, mean construct scores, and correlations between each construct.

## Results

### Overview

Characteristics of the sample are listed in Table 2.

**Table 2.** Demographic and practice characteristics of the sample.

| Characteristic | Value (N=10), n (%) |
|---|---|
| **Gender** | |
| Women | 9 (90) |
| Men | 1 (10) |
| **Age in years** | |
| 30-39 | 3 (30) |
| 40-49 | 2 (20) |
| 50-59 | 4 (40) |
| >60 | 1 (10) |
| **Nurse practitioner (NP) certification** | |
| Family nurse practitioner (FNP) | 8 (80) |
| Adult nurse practitioner (ANP) | 0 (0) |
| Geriatric nurse practitioner (GNP) | 0 (0) |
| FNP + ANP | 1 (10) |
| FNP + GNP | 1 (10) |
| **Type of NP practice** | |
| Group | 9 (90) |
| Individual | 1 (10) |
| **Highest degree obtained** | |
| Masters NP Certification | 8 (80) |
| Doctorate of Nursing Practice | 2 (20) |
| **Years in clinical practice** | |
| 1-5 | 4 (40) |
| 6-10 | 4 (40) |
| 11-20 | 0 (0) |
| 21-30 | 1 (10) |
| 31-40 | 1 (10) |
| **Which electronic device are you using for this intervention?** | |
| Computer | 6 (60) |
| Mobile device | 4 (40) |

## Aim 1a: Content Validity

The dermatology expert panel conducted two reviews to assess the content validity of the intervention. Content validity scores primarily increased or were consistent during the second round of reviews after the storyboards were adjusted to address the reviewers' recommendations from the first round. The scores of the following components decreased during the second round: right arm nevus count (relevance) and discuss systematic approach (relevance) (see Table 3).

**Table 3.** Content validity index (CVI).

| Strategies | First review, mean score (SD) | | Second review, mean score (SD) | |
|---|---|---|---|---|
| | Relevance[a] | Clarity[b] | Relevance | Clarity |
| **Video 1 strategies** | | | | |
| What is skin cancer? | 3.33 (1.52) | 2.33 (1.53) | 4.00 (0) | 3.67 (0.58) |
| Malignant melanoma prevalence | 3.67 (0.58) | 3.67 (0.58) | 3.67 (0.58) | 4.00 (0) |
| Malignant melanoma risk factors | 3.67 (0.58) | 3.00 (1.00) | 4.00 (0) | 4.00 (0) |
| Right arm nevus count | 4.00 (0) | 3.00 (1.73) | 3.67 (0.58) | 3.67 (0.58) |
| Grand mean | 3.67 (0.27) | 3.00 (0.54) | 3.84 (0.19) | 3.83 (0.19) |
| **Video 2 strategies** | | | | |
| Discuss systematic approach | 4.00 (0) | 3.67 (0.58) | 3.67 (0.58) | 4.00 (0) |
| Discuss hard-to-see areas | 3.33 (0.58) | 3.00 (1.73) | 3.67 (0.58) | 4.00 (0) |
| Strategies for incorporating clinical skin examination | 3.00 (1.00) | 2.67 (1.15) | 4.00 (0) | 4.00 (0) |
| Grand mean | 3.44 (0.50) | 3.11 (0.58) | 3.78 (0.19) | 4.00 (0) |
| **Video 3 strategies** | | | | |
| ABCDE rule (Asymmetry, Border, Color, Diameter, Evolution) | 4.00 (0) | 4.00 (0) | 4.00 (0) | 4.00 (0) |
| Ugly duckling sign | 4.00 (0) | 3.33 (1.15) | 4.00 (0) | 4.00 (0) |
| Images of nonsuspicious vs suspicious pigmented lesions | 3.67 (0.58) | 2.00 (1.73) | 3.67 (0.58) | 4.00 (0) |
| Grand mean | 3.89 (0.33) | 4.00 (0.88) | 3.89 (0.33) | 4.00 (0) |

[a] Relevance scale: 1 = not relevant, 2 = somewhat relevant, 3 = quite relevant, and 4 = highly relevant.

[b] Clarity scale: 1 = not clear, 2 = somewhat clear, 3 = quite clear, and 4 = very clear.

## Aims 1b and 1c: Vimeo and REDCap Integration and Digital Delivery

The dermatology expert panel rated the integration of REDCap and Vimeo and the system's usability as *truly superior* (mean 95.8, SD 7.2), with a range of scores from 87.5 (*better*) to 100 (*truly superior*). All expert panel members scored each of the five technical concepts—video design, intended content, visual quality, audio quality, and audio-visual relationship—as *exceptional* (mean 5, SD 0) (see Multimedia Appendix 4).

## Aim 2: Enrollment and Retention Rate and Intervention Adherence

A total of 12 PCNPs consented from a list of 32 emails (38% enrollment rate). Out of 12 participants, 10 completed the intervention and the surveys (83% retention rate). Completion rate of the surveys was 100% (10/10). A total of 50% (5/10) of participants watched the videos in their entirety. Vimeo recorded 6 plays for video 1. Of the 6 participants who played video 1, only 1 (17%) failed to finish the video, ending it 30 seconds before the content was complete. Therefore, the lowest percentage of possible participants who completed all of the videos is 50% (5/10). See Table 4 for the video adherence information for each module.

**Table 4.** Vimeo report regarding video adherence.

| Video | Plays (N=10), n (%) | Finishes (N=10), n (%) | Average amount of the video watched, % |
|---|---|---|---|
| Video 1 | 6 (60) | 4 (40) | 92 |
| Video 2 | 7 (70) | 4 (40) | 97 |
| Video 3 | 8 (80) | 2 (20) | 98 |

## Aim 2: Usability, Acceptability, and Satisfaction

The mean usability was better (mean 85.8, SD 10.6), with a range of scores from 72.5 (*better*) to 100 (*truly superior*). Acceptability of the intervention was assessed using the AWCL. The mean for each of the constructs—*perceived usefulness, perceived ease of use, behavior,* and *affection*—all ranged between *somewhat agree* and *mostly agree* (see Table 5).

**Table 5.** Attitudes toward Web-based Continuing Learning Survey (AWCL) item scores and grand mean scores of the constructs.

| Item | Score[a], mean (SD) |
|---|---|
| Perceived usefulness 1: web-based continuing learning helps my work become more interesting | 5.2 (1.4) |
| Perceived usefulness 2: web-based continuing learning helps to increase my creativity for work | 5.3 (1.3) |
| Perceived usefulness 3: web-based continuing learning facilitates the development of my work | 5.4 (1.2) |
| Perceived usefulness 4: web-based continuing learning effectively enhances my learning | 5.2 (1.0) |
| Perceived usefulness 5: web-based continuing learning helps me attain better learning outcomes | 5.2 (1.4) |
| Perceived ease of use 1: it is convenient to receive training on the job using web-based continuing learning | 5.7 (1.2) |
| Perceived ease of use 2: it is easy to get web-based continuing learning to do what I want it to | 5.1 (1.5) |
| Perceived ease of use 3: it is easy for me to solve problems at work when I participate in web-based continuing learning | 4.9 (1.6) |
| Perceived ease of use 4: the flexibility of web-based continuing learning makes me learn in an easier way | 5.9 (1.1) |
| Behavior 1: I hope to spend more time using web-based continuing learning | 5.7 (1.5) |
| Behavior 2: I hope to use web-based continuing learning more often | 5.5 (1.4) |
| Behavior 3: I want to increase my use of web-based continuing learning in the future | 5.4 (1.4) |
| Affection 1: I think it is interesting to use web-based continuing learning | 5.8 (1.3) |
| Affection 2: web-based continuing learning provides an interesting and attractive environment | 5.7 (1.5) |
| Affection 3: using web-based continuing learning can improve my working ability | 5.8 (1.3) |
| **Constructs, grand mean (SD)** | |
| Perceived usefulness | 5.26 (1.03) |
| Perceived ease of use | 5.40 (0.85) |
| Behavior | 5.53 (1.25) |
| Affection | 5.77 (1.37) |

[a] The scores are based on the following scale: 1 = strongly disagree, 2 = mostly disagree, 3 = somewhat disagree, 4 = neither agree nor disagree, 5 = somewhat agree, 6 = mostly agree, and 7 = strongly agree.

The overall mean satisfaction with the study and the intervention was 99% (SD 1.87, N=10; 100% = *best*). All participants (N=10) would watch the videos even if they did not receive compensation, they believed that the length of the video was “just right,” and they believed that the content was “just right.” Out of 10 participants, 8 (80%) preferred that the videos be accessible for 1 month. When participants were asked what they learned, they cited specific videos (eg, “scaling suspicious vs nonsuspicious lesions”) (5/10, 50%) and increased motivation (eg, “11 plus nevus right arm increased likelihood to have 100 plus. So quick and easy to check!”) (1/10, 10%), and some identified having already completed skin assessments (eg, “was a great refresher”) (2/10, 20%).

## Discussion

A strength of this study is the rigorous, transparent development of the videos and the use of an expert panel of dermatologists to ensure valid content. Each specific strategy recommended by Sidani and Braden for operational transparency of the video intervention was outlined (see Multimedia Appendix 2) and graded by the dermatology expert panel using a CVI. This process is absent from many studies describing intervention development [37]. The CVI indicated that the videos were relevant, except for two index scores. Scores for both *right arm nevus count* and *discuss systemic approach* decreased from 4 (*highly relevant*) to 3.67 (between *highly relevant* and *quite relevant*). According to the scale, both components were ranked as at least *quite relevant* and were kept in the videos. One explanation for this alteration is that there was a different, third expert reviewer during the second round. Otherwise, clarity increased overall during the second round of reviews. The expert panel’s comments primarily focused on promoting clarity, such as adding the definition of skin cancer to the first video and the definition of a pigmented lesion to the third video. Wording was also adjusted (eg, “get melanoma” to “develop melanoma” and “11 nevi tool” to “right arm nevus count”). The expert panel also offered relevant CSE tips, such as ensuring that the patient removes glasses or hearing aids during the head-to-toe skin examination to better visualize the conchal bowl.

Prior to the dissemination of videos to the PCNPs, the dermatology expert panel evaluated the videos’ technical production on Vimeo and their delivery through REDCap. Visual quality, audio quality, and audio-visual relationship were considered exceptional, suggesting that it was feasible to use Vimeo and REDCap to deliver the video intervention.

The hypothesis that the enrollment rate will be equal to or better than 60% was not supported. During the study, 22 NPs were recruited at a conference in November 2018. However, enrollment did not start until February 2019, and just 5 NPs consented after the initial study invitations were sent. An additional 10 participants were recruited at another NP meeting and 8 more participants were enrolled. This highlights challenges in recruiting NPs [48], as well as the importance of a timely follow-up after recruitment.

The hypothesis that the retention rate will be greater than or equal to 50% was supported. Out of 12 recruited participants, 10 (83%) completed the surveys. This is comparable to the retention rate of a prior skin cancer study involving NPs, which reported a retention of 10 out of 14 participants (71%) [13].

Adding strength to this study are the comments from the 2 participants who did not progress past the first video. The first participant stated that she was unable to complete the other videos because of work requirements. She rated the first video favorably, stating that it was very educational. The second participant accessed the first video on her mobile phone without problems but had difficulty accessing the second video, receiving a message that she had already reviewed the video. When she contacted the investigator, she was outside the 1-week time limit for video completion and was withdrawn from the study. Future studies will be formatted to have the REDCap email, which includes the video link, request that participants check their Wi-Fi or cellular connections prior to opening the REDCap link and starting the video.

The hypothesis that video intervention adherence will be greater than or equal to 50% was not supported. Vimeo defines a finish as when the participant views a video to the very last frame. However, the amounts viewed of each video were high: the average amount of video 1 watched was 92% and it was played 6 out of 10 times (60%), the average amount of video 2 watched was 97% and it was played 7 out of 10 times (70%), and the average amount of video 3 watched was 98% and it was played 8 out of 10 times (80%). Participants likely viewed all the content; however, they most likely dropped off at the references and acknowledgements portions of the videos. We were unable to retrospectively connect the times the videos were played according to the Vimeo report with the time the participant finished the REDCap module. However, at least 5 out of 10 PCNPs (50%) fully viewed the content before exiting out of the video and met the “finishing” criteria. Future studies will require active monitoring of the reports to be able to connect the participants’ REDCap and Vimeo information or to obtain a time stamp on REDCap. One way to obtain a time stamp would be to keep a daily record of which participants completed the REDCap survey and correlate this record with the time the Vimeo video was viewed. To increase the number of finishes as defined by Vimeo, the references and the acknowledgements will be moved to the beginning of the video. References can be sent with the email invitation and the acknowledgements can be placed at the start of the video.

The hypothesis that usability scores will be equal to or higher than 70 was supported. The PCNP’s mean SUS scores were 85, or *better*, indicating that the participants viewed the usability of REDCap and Vimeo as positive. Similar SUS scores have been recorded for other interventions, such as a web-based simulation in psychiatry residency training (n=16, score 86.5) [49] and a web-based multimedia application called Electrolyte Workshop, which has e-learning tutorials called WalkThru (n=18, score 87.9) and HandsOn (n=27, score 81.5) [50].

Another strength of this study was the use of the microlearning framework to guide intervention development. The hypothesis that acceptability scores for microlearning will be equal to or higher than 5 was supported. Mean scores for the scales of *perceived usefulness*, *perceived ease of use*, *behavior*, and *affection* all were higher than those reported by Liang et al [30]. These findings suggest that microlearning was an acceptable framework for this intervention. However, further testing is needed to further demonstrate that microlearning is an optimal framework for teaching complex concepts to busy practitioners in a short amount of time.

There were also some limitations to the study. Purposive sampling decreases the generalizability of the results and adds to selection bias [51]. Only the participants who completed the intervention were able to complete the usability, acceptability, and satisfaction surveys. Another limitation to this study was the posttest-only design. This design does not allow for a comparison between the participants before and after the video interventions. However, this design is appropriate for a feasibility study, where the focus is on the practicality of the study [51].

In conclusion, three theory-based, short videos with content on CSE for malignant melanoma were developed that are suitable for internet delivery to PCNPs in various formats. Findings from this feasibility study provide a foundation for the use of microlearning as a guide for delivering brief CSE training to PCNPs. The findings also provide support for using the videos as part of an intervention in future trials targeting behavioral CSE outcomes in PCNPs or other practitioners. This feasibility study provided valuable lessons to inform components of the next research phase, such as the timeliness of enrollment and redesign of the modules to support better measurements of intervention adherence. The long-term goal is to promote the early detection of skin cancer by providing CSE education to PCNPs and ultimately improving skin cancer prognoses.

## Acknowledgments

This study was funded by the University of Arizona Cancer Center, Skin Cancer Institute, Seed Grant Program and the College of Nursing Fall 2018 Frederick Lange Memorial Endowment Award. The authors acknowledge Stephen Purcell for video production assistance.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Definitions of Sidani and Braden’s clarifying elements of an intervention.
[DOCX File , 14 KB - derma_v3i1e16714_app1.docx ]

Multimedia Appendix 2
The components and activities of a clinical skin examination (CSE) intervention.
[DOCX File , 23 KB - derma_v3i1e16714_app2.docx ]

Multimedia Appendix 3
Consent form.
[PDF File (Adobe PDF File), 107 KB - derma_v3i1e16714_app3.pdf ]

Multimedia Appendix 4
Descriptions of the different items from Beaudin and Quick’s Quality Evaluation of Video (QEV).
[DOCX File , 21 KB - derma_v3i1e16714_app4.docx ]

## References

1. Guy GP, Thomas CC, Thompson T, Watson M, Massetti GM, Richardson LC, Centers for Disease Control and Prevention (CDC). Vital signs: Melanoma incidence and mortality trends and projections - United States, 1982-2030. MMWR Morb Mortal Wkly Rep 2015 Jun 05;64(21):591-596 [FREE Full text] [Medline: 26042651]
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019 Jan;69(1):7-34 [FREE Full text] [doi: 10.3322/caac.21551] [Medline: 30620402]
3. Bartlett EK, Karakousis GC. Current staging and prognostic factors in melanoma. Surg Oncol Clin N Am 2015 Apr;24(2):215-227. [doi: 10.1016/j.soc.2014.12.001] [Medline: 25769707]
4. Glazer AM, Rigel DS, Winkelmann RR, Farberg AS. Clinical diagnosis of skin cancer: Enhancing inspection and early recognition. Dermatol Clin 2017 Oct;35(4):409-416. [doi: 10.1016/j.det.2017.06.001] [Medline: 28886797]
5. Gershenwald JE, Scolyer RA, Hess KR, Sondak VK, Long GV, Ross MI, for members of the American Joint Committee on Cancer Melanoma Expert Panel and the International Melanoma Database and Discovery Platform. Melanoma staging: Evidence-based changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA Cancer J Clin 2017 Nov;67(6):472-492 [FREE Full text] [doi: 10.3322/caac.21409] [Medline: 29028110]
6. Uhlenhake E, Brodell R, Mostow E. The dermatology work force: A focus on urban versus rural wait times. J Am Acad Dermatol 2009 Jul;61(1):17-22. [doi: 10.1016/j.jaad.2008.09.008] [Medline: 19427058]
7. Slade K, Lazenby M, Grant-Kels JM. Ethics of utilizing nurse practitioners and physician's assistants in the dermatology setting. Clin Dermatol 2012;30(5):516-521. [doi: 10.1016/j.clindermatol.2011.06.022] [Medline: 22902223]
8. Kimball AB, Resneck JS. The US dermatology workforce: A specialty remains in shortage. J Am Acad Dermatol 2008 Nov;59(5):741-745. [doi: 10.1016/j.jaad.2008.06.037] [Medline: 18723242]
9. Resneck J, Kimball AB. The dermatology workforce shortage. J Am Acad Dermatol 2004 Jan;50(1):50-54. [doi: 10.1016/j.jaad.2003.07.001] [Medline: 14699364]
10. Aquino LL, Wen G, Wu JJ. Factors affecting the pursuit of academic careers among dermatology residents. Cutis 2015 Apr;95(4):231-236. [Medline: 25942025]
11. Loescher LJ, Harris JM, Curiel-Lewandrowski C. A systematic review of advanced practice nurses' skin cancer assessment barriers, skin lesion recognition skills, and skin cancer training activities. J Am Acad Nurse Pract 2011 Dec;23(12):667-673. [doi: 10.1111/j.1745-7599.2011.00659.x] [Medline: 22145657]
12. Loescher LJ, Stratton D, Slebodnik M, Goodman H. Systematic review of advanced practice nurses' skin cancer detection knowledge and attitudes, clinical skin examination, lesion detection, and training. J Am Assoc Nurse Pract 2018 Jan;30(1):43-58. [doi: 10.1097/JXX.0000000000000004] [Medline: 29757921]
13. Blake J, Malone L. Current behaviors, attitudes, and knowledge of nurse practitioners in primary care toward skin cancer screening/prevention. J Dermatol Nurses Assoc 2014 Sep 1;6(2):65-69. [doi: 10.1097/JDN.0000000000000031]
14. Azima S. Expert views, opinions, and recommendations. J Dermatol Nurses Assoc 2016 Sep;8(5):312-317. [doi: 10.1097/JDN.0000000000000250]
15. Shelby DM. Knowledge, Attitudes and Practice of Primary Care Nurse Practitioners Regarding Skin Cancer Assessments: Validity and Reliability of a New Instrument [doctoral dissertation]. Tampa, FL: University of South Florida; 2014 Feb 27. URL: https://scholarcommons.usf.edu/cgi/viewcontent.cgi?article=6320&context=etd [accessed 2020-07-07]
16. Oliveria SA, Nehal KS, Christos PJ, Sharma N, Tromberg JS, Halpern AC. Using nurse practitioners for skin cancer screening: A pilot study. Am J Prev Med 2001 Oct;21(3):214-217. [doi: 10.1016/s0749-3797(01)00354-3] [Medline: 11567843]
17. Roebuck H, Moran K, MacDonald DA, Shumer S, McCune RL. Assessing skin cancer prevention and detection educational needs: An andragogical approach. J Nurse Pract 2015 Apr;11(4):409-416 [FREE Full text] [doi: 10.1016/j.nurpra.2015.01.036]

18. Ali FR, Samarasinghe V, Russell SA, Lear JT. Increasing capacity for skin surveillance in a transplant review clinic. Transplantation 2014 Apr 27;97(8):e48-e50. [doi: 10.1097/TP.0000000000000084] [Medline: 24732901]
19. Armstrong A. Dermoscopy: An Evidence-Based Approach for the Early Detection of Melanoma [doctoral dissertation]. Jacksonville, FL: University of North Florida; 2011. URL: https://digitalcommons.unf.edu/cgi/viewcontent.cgi?article=1324&context=etd [accessed 2020-07-07]
20. Wray EV, Brant B, Hussian F, Muller FM. A new model of teledermoscopy combining service and education. Br J Dermatol 2013;169:138-139.
21. DeKoninck B, Christenbery T. Skin cancer screening in the medically underserved population: A unique opportunity for APNs to make a difference. J Am Assoc Nurse Pract 2015 Sep;27(9):501-506. [doi: 10.1002/2327-6924.12216] [Medline: 25728687]
22. Bradley HB. Implementation of a skin cancer screening tool in a primary care setting: A pilot study. J Am Acad Nurse Pract 2012 Feb;24(2):82-88. [doi: 10.1111/j.1745-7599.2011.00669.x] [Medline: 22324863]
23. Chen C, Woyansky S, Zundell C. The effects of education on compliance with skin cancer risk reduction guidelines. J Dermatol Nurses Assoc 2015;7(2):97-100. [doi: 10.1097/jdn.0000000000000119]
24. Hartnett PD, O'Keefe C. improving skin cancer knowledge among nurse practitioners. J Dermatol Nurses Assoc 2016 Sep;8(2):123-128. [doi: 10.1097/JDN.0000000000000206]
25. Polit DF, Tatano Beck C. Developing and testing self-report scales. In: Nursing Research: Generating and Assessing Evidence for Nursing Practice. 9th edition. Hagerstown, MA: Lippincott Williams & Wilkins; 2012:351-378.
26. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health 2007 Aug;30(4):459-467. [doi: 10.1002/nur.20199] [Medline: 17654487]
27. Yaghmaei F. Content validity and its estimation. J Med Educ 2003;3(1):25-27. [doi: 10.32388/a6w263]
28. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594 [FREE Full text] [doi: 10.1080/10447310802205776]
29. Beaudin BP, Quick D. Instructional video evaluation instrument. J Ext 1996 Jun;34(3) [FREE Full text]
30. Liang J, Wu S, Tsai C. Nurses' internet self-efficacy and attitudes toward web-based continuing learning. Nurse Educ Today 2011 Nov;31(8):768-773. [doi: 10.1016/j.nedt.2010.11.021] [Medline: 21176859]
31. Goldsmith LA, Bolognia JL, Callen JP, Chen SC, Feldman SR, Lim HW, American Academy of Dermatology. American Academy of Dermatology Consensus Conference on the safe and optimal use of isotretinoin: Summary and recommendations. J Am Acad Dermatol 2004 Jun;50(6):900-906. [doi: 10.1016/j.jaad.2004.02.012] [Medline: 15153892]
32. Harris JM, Salasche SJ, Harris RB. Can internet-based continuing medical education improve physicians' skin cancer knowledge and skills? J Gen Intern Med 2001 Jan;16(1):50-56 [FREE Full text] [Medline: 11251750]
33. Learning module: The skin exam. American Academy of Dermatology. 2018. URL: https://www.aad.org/education/basic-derm-curriculum/suggested-order-of-modules/the-skin-exam [accessed 2018-04-02]
34. Bruck PA, Motiwalla L, Foerster F. Mobile learning with micro-content: A framework and evaluation. In: Proceedings of the 25th Bled eConference. eDependability: Reliable and Trustworthy eStructures, eProcesses, eOperations and eServices for the Future. 2012 Presented at: 25th Bled eConference. eDependability: Reliable and Trustworthy eStructures, eProcesses, eOperations and eServices for the Future; June 17-20, 2012; Bled, Slovenia p. 527-543 URL: https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1041&context=bled2012
35. Hug T, Friesen N. Outline of a microlearning agenda. In: Hug T, editor. Didactics of Microlearning: Concepts, Discourses and Examples. Münster, Germany: Waxmann; 2007:15-31.
36. Giurgiu L. Microlearning an evolving elearning trend. Sci Bull 2017;22(1):18-23 [FREE Full text] [doi: 10.1515/bsaft-2017-0003]
37. Stratton DB, Loescher LJ. Educational interventions for primary care providers to improve clinical skin examination for skin cancer. J Am Assoc Nurse Pract 2020 May;32(5):369-379. [doi: 10.1097/JXX.0000000000000235] [Medline: 31369454]
38. Sidani S, Braden CJ. Clarifying elements of the intervention. In: Design, Evaluation, and Translation of Nursing Interventions. Ames, IA: John Wiley & Sons; 2011:43-60.
39. Vimeo. URL: https://vimeo.com/features [accessed 2018-04-02]
40. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap): A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
41. Sidani S, Braden CJ. Selecting, training, and addressing the influence of interventionists. In: Design, Evaluation, and Translation of Nursing Interventions. Ames, IA: John Wiley & Sons; 2011:111-124.
42. Polit DF, Beck CT. The content validity index: Are you sure you know what's being reported? Critique and recommendations. Res Nurs Health 2006 Oct;29(5):489-497. [doi: 10.1002/nur.20147] [Medline: 16977646]
43. What does Vimeo consider a play? Vimeo. 2019. URL: https://vimeo.zendesk.com/hc/en-us/articles/360046057111-User-Level-Analytics-in-Showcases [accessed 2019-05-20]
44. Nieswiadomy RM. Foundations of Nursing Research. 4th edition. Upper Saddle River, NJ: Pearson Education; 2002.
45. Hertzog MA. Considerations in determining sample size for pilot studies. Res Nurs Health 2008 Apr;31(2):180-191. [doi: 10.1002/nur.20247] [Medline: 18183564]

46. REDCap boilerplate text. The University of Arizona Health Sciences Center for Biomedical Informatics and Biostatistics. URL: https://redcap.cats.med.arizona.edu/redcap/redcap_boilerplate_text.doc [accessed 2020-05-20]
47. IBM Corp. IBM SPSS Statistics for Windows 26.0. Armonk, NY: IBM Corp; 2019. URL: https://www.ibm.com/support/pages/node/876320 [accessed 2020-07-06]
48. Schoenthaler A, Albright G, Hibbard J, Goldman R. Simulated conversations with virtual humans to improve patient-provider communication and reduce unnecessary prescriptions for antibiotics: A repeated measure pilot study. JMIR Med Educ 2017 Apr 19;3(1):e7 [FREE Full text] [doi: 10.2196/mededu.6305] [Medline: 28428160]
49. Gorrindo T, Baer L, Sanders KM, Birnbaum RJ, Fromson JA, Sutton-Skinner KM, et al. Web-based simulation in psychiatry residency training: A pilot study. Acad Psychiatry 2011;35(4):232-237. [doi: 10.1176/appi.ap.35.4.232] [Medline: 21804041]
50. Davids MR, Chikte UM, Halperin ML. Effect of improving the usability of an e-learning resource: A randomized trial. Adv Physiol Educ 2014 Jun;38(2):155-160 [FREE Full text] [doi: 10.1152/advan.00119.2013] [Medline: 24913451]
51. Shadish WR, Cook TD, Campbell DT. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston, MA: Houghton Mifflin; 2002.

## Abbreviations

**AANP:** American Association of Nurse Practitioners
**ANCC:** American Nurses Credentialing Center
**AWCL:** Attitudes toward Web-based Continuing Learning Survey
**CSE:** clinical skin examination
**CVI:** content validity index
**HIPAA:** Health Insurance Portability and Accountability Act
**NP:** nurse practitioner
**PCNP:** primary care nurse practitioner
**QEV:** Quality Evaluation of Video
**REDCap:** Research Electronic Data Capture
**SUS:** System Usability Scale
**UITS:** University of Arizona's Information Technology Services Center

*Edited by G Eysenbach; submitted 16.10.19; peer-reviewed by E Hacker, M Janda; comments to author 02.04.20; revised version received 12.06.20; accepted 15.06.20; published 27.07.20.*

*Please cite as:*
*Stratton DB, Shea KD, Knight EP, Loescher LJ*
*Delivering Clinical Skin Examination Education to Nurse Practitioners Using an Internet-Based, Microlearning Approach: Development and Feasibility of a Video Intervention*
*JMIR Dermatol 2020;3(1):e16714*
*URL: http://derma.jmir.org/2020/1/e16714/*
*doi:10.2196/16714*
*PMID:*

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org