



OPEN ACCESS

EDITED BY

Ben Singh,
University of South Australia, Australia

REVIEWED BY

Annemiek J. Linn,
University of Amsterdam, Netherlands
Lisa Gualtieri,
Tufts University, United States

*CORRESPONDENCE

Prathamesh Karmalkar
✉ prathamesh.karmalkar@merckgroup.com

RECEIVED 08 July 2024

ACCEPTED 25 October 2024

PUBLISHED 20 November 2024

CITATION

Spies E, Flynn JA, Oliveira NG, Karmalkar P and Gurulingappa H (2024) Artificial intelligence-enabled social media listening to inform early patient-focused drug development: perspectives on approaches and strategies. *Front. Digit. Health* 6:1459201. doi: 10.3389/fdgth.2024.1459201

COPYRIGHT

© 2024 Spies, Flynn, Oliveira, Karmalkar and Gurulingappa. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Artificial intelligence-enabled social media listening to inform early patient-focused drug development: perspectives on approaches and strategies

Erica Spies¹, Jennifer A. Flynn¹, Nuno Guitian Oliveira², Prathamesh Karmalkar^{3*} and Harsha Gurulingappa³

¹Work Completed While Employees of EMD Serono Research & Development Institute, Inc., Billerica, MA, United States, ²Healthcare Business of Merck KGaA, Darmstadt, Germany, ³Merck IT Centre, Merck Data & AI Organization, Merck Group, Bangalore, India

This article examines the opportunities and benefits of artificial intelligence (AI)-enabled social media listening (SML) in assisting successful patient-focused drug development (PFDD). PFDD aims to incorporate the patient perspective to improve the quality, relevance, safety, and efficiency of drug development and evaluation. Gathering patient perspectives to support PFDD is aided by the participation of patient groups in communicating their treatment experiences, needs, preferences, and priorities through online platforms. SML is a method of gathering feedback directly from patients; however, distilling the quantity of data into actionable insights is challenging. AI-enabled methods, such as natural language processing (NLP), can facilitate data processing from SML studies. Herein, we describe a novel, trainable, AI-enabled, SML workflow that classifies posts made by patients or caregivers and uses NLP to provide data on their experiences. Our approach is an iterative process that balances human expert-led milestones and AI-enabled processes to support data preprocessing, patient and caregiver classification, and NLP methods to produce qualitative data. We explored the applicability of this workflow in 2 studies: 1 in patients with head and neck cancers and another in patients with esophageal cancer. Continuous refinement of AI-enabled algorithms was essential for collecting accurate and valuable results. This approach and workflow contribute to the establishment of well-defined standards of SML studies and advance the methodologic quality and rigor of researchers contributing to, conducting, and evaluating SML studies in a PFDD context.

KEYWORDS

artificial intelligence (AI), social media listening (SML), patient focused drug development (PFDD), unmet needs, machine learning (ML), deep learning (DL)

Abbreviations

AE, adverse event; AI, artificial intelligence; BERT, bidirectional encoder representations from transformers; FDA, US Food and Drug Administration; ML, machine learning; NLP, natural language processing; PFDD, patient-focused drug development; QoL, quality of life; SML, social media listening.

1 Introduction

There is increased recognition that patients' perspectives have not always been factored into therapy research and development, clinical trial development and execution, labeling, and access to emerging therapies. Given the importance of patients' varied health and therapy experience, this lack of inclusion has spurred a new effort to incorporate patient perspectives throughout the drug development life cycle (1). This focused effort on patient perspectives and the incorporation of patient-experience data is facilitated by traditional approaches (e.g., patient interviews, patient engagement meetings), as well as the increased participation of patient groups in communicating their treatment experiences, needs, and preferences through online platforms. Regulatory agencies have recognized the importance of including patient-focused drug development (PFDD) in the overall drug development process and have begun providing guidance on incorporating patient and caregiver perspectives to improve the quality, relevance, safety, and efficiency of drug development, and to inform drug evaluation and regulatory decision making (2–4). PFDD approaches can give patients and caregivers a voice in the drug development process by elucidating aspects of a target product value profile, identifying patient needs and preferences, and selecting clinical study endpoints (5). Various methods of gathering patient input have been endorsed by the US Food and Drug Administration (FDA), including social media listening (SML) (3, 6).

SML involves collecting and analyzing online statements and conversations from social media platforms (e.g., X[®], Facebook[®]), patient-centric social media platforms, blogs, forums, and other means of online communication (7–9). A growing number of patients are taking to online communities and platforms to describe their individual health care journeys, acting as an invaluable source of qualitative data on unique patient perspectives that may otherwise go unheard by health care providers and drug developers. SML approaches in PFDD have been reported in numerous use cases, including identifying unmet medical needs, characterizing target patient populations, repurposing drugs, recruiting patients, detecting adverse events (AEs), and detailing patient experiences with a disease and/or treatment (10–13). SML offers the potential to gather a range of viewpoints directly from patients without compromising personally identifiable information to supplement/complement more traditional PFDD methods. See Figure 1 for an overview of how novel approaches can support traditional PFDD approaches throughout the drug development life cycle.

It is important to understand that patients' lived experiences vary and often extend beyond treatment efficacy. Prior analyses have found that this is particularly true among patients receiving oncologic therapy, where individual experiences depend on specific cancer type/stage, therapy choice, and the presence/absence of other conditions. During a 1-h, web-based breast cancer support group, patients noted diverse concerns, including communication gaps between patients and caregivers, timeliness of insurance company responses, poor integration of palliative care, and financial challenges (14). An analysis of social media

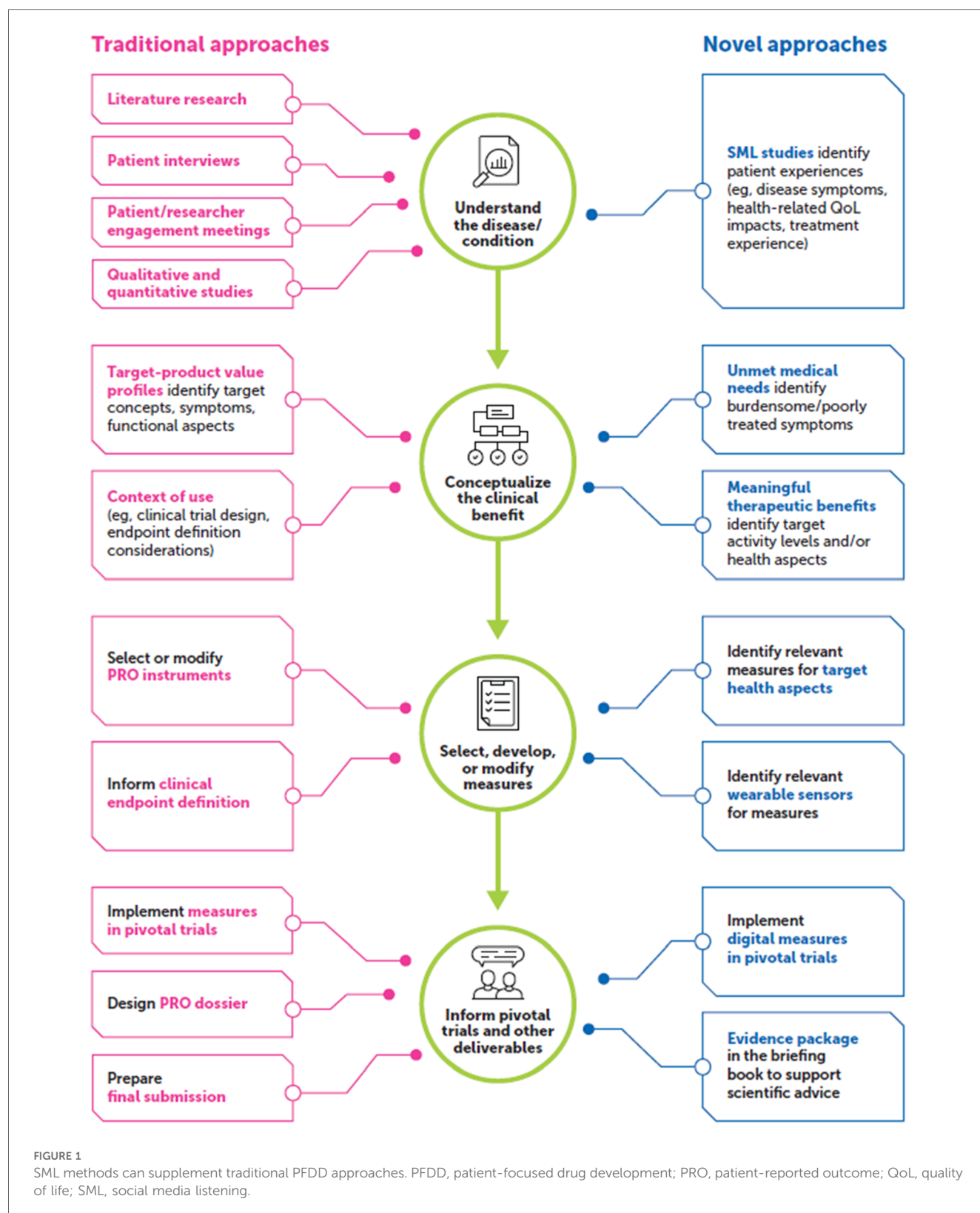
posts by patients with breast cancer found further variation in topics, including treatment patterns, patient journeys, burden of illness, unmet needs, AEs, quality of life (QoL), social impacts from the disease, and perceptions on drug effectiveness (15). Another interview-based study highlighted the importance of advanced care planning and the need for earlier discussion between health care providers and patients and families (16). These reports emphasize the importance of gathering multidimensional perspectives from diverse populations that can collectively provide valuable insights to inform treatment decisions.

In this article, we discuss existing evidence of SML application in oncology-based studies and their contributions to PFDD. We highlight the value of a novel, flexible, artificial intelligence (AI)-enabled SML workflow and describe how it can be used to gather patient and caregiver insights that can be used to promote and support PFDD. We report the approach and methodologic learnings from 2 oncology case studies (head/neck and esophageal cancers) to inform future implementations of our workflow and PFDD investigations.

2 Review of SML studies in oncology and their uses for PFDD

SML studies have been employed in multiple therapeutic areas, with the goal of gaining insight into the needs, motivations, behaviors, and considerations of patients and caregivers to inform early drug development strategies (7, 17–19). The oncology space is well suited for SML approaches because cancer entails a high symptom and treatment burden, with patient communities sharing experiences frequently and in detail. In 2005, more than 400,000 internet cancer support groups were found, each having approximately 2,000 members (20). In 2011, 620 breast cancer groups (fundraising, awareness, product/service, patient/caregiver support) were identified on Facebook[®], with more than 1,000,000 members collectively (21). In 2020, 123 apps oriented to patients with cancer were identified across 2 major application marketplaces (22). With increasing smartphone access and a continually growing prevalence and utilization of online platforms, online platforms that address the patient and caregiver experience have become a valuable resource in SML.

Several publications have established the value of SML when seeking to identify patients' unmet medical needs, better characterize the patient population, detect AEs, and examine treatment-switching behaviors. Arun et al. investigated the patient experience among patients with follicular lymphoma using SML and found that the most common topics discussed were the disease and treatment impacts on QoL, curability, and the fear of relapse/progression (18). Chauhan et al. employed SML to elucidate perceptions and needs among patients with melanoma in multiple European countries, further validating SML as a tool to gather evidence from patients and stakeholders across countries. Twitter[®] (now X[®]) was the primary channel across the majority of the countries examined, and the most frequently discussed topics were treatment stage, diagnosis and



tests, surgery and immunotherapy, QoL issues (particularly the emotional burden of disease), and the availability of effective treatments and access to good health care providers (19). Mendelson et al. collected data from Twitter® (now X®), Facebook®, forums, and news/blogs among patients with diffuse

large B-cell lymphoma, revealing that the most discussed symptoms were pain, enlarged lymph nodes, B symptoms, and fatigue. Most posts mentioned switching treatment to stem cell therapy following multiple experiences of relapse and chemotherapy failures (7).

The quantity of available social media data creates challenges to data extraction and analysis. Typically, online aggregators are employed to collect bulk social media data that are refined via natural language processing (NLP) algorithms (10, 11). Common NLP strategies include frequency and co-occurrence analysis, part-of-speech tagging, sentiment analysis, indexing/extracting key terms, term frequency-inverse document frequency, global vectorization word embedding, sequential data processing deep learning, and bidirectional encoder representation (11, 23–26). AI, or machine learning (ML), methods have also been used to process the vast quantities of data produced through social media mining in disease-specific and nonspecific social media communities for a variety of therapeutic areas beyond oncology, including dry eye disease and atopic dermatitis, where approaches such as the skip-gram negative sampling variant of the word2vec neural network and the Biterm Topic Model have been employed (12, 27, 28).

Several studies have described AI with NLP approaches for probing patient-experience data among patients with cancer. Fang et al. explored the application of NLP methods to classify unstructured text from patient interviews and identify patient-reported symptoms and QoL impacts among patients with hepatocellular carcinoma, biliary tract cancer, and gastric cancer. The most promising NLP model identified was a Bidirectional Encoder Representations from Transformers (BERT) model, which outperformed other models in accurately predicting the multiclass classification of unstructured text, demonstrated high predictive performance and could better capture the context of a word given its relative position within a sentence (26). This predictive performance was consistent across interviews from different patient populations and studies, indicating both versatility and generalizability. In summation, this proof-of-concept study showed the potential of NLP-based models to accurately contribute to the automated processing of patient interview transcripts, which could facilitate the incorporation of patient input into the drug development process. Although there are numerous examples demonstrating the value, approaches, and uses of SML, few describe a customizable workflow to conduct SML studies based on an iterative process driven by human experts and enabled by AI methods to facilitate concept identification.

3 Overview of an AI-enabled SML workflow

We have developed an AI-enabled SML model workflow for identifying and processing publicly available social media posts to better understand patient and caregiver experiences. This customizable workflow integrates human milestones and AI-enabled methodologies, with human experts retraining models to iteratively validate and optimize the AI-enabled methods identifying posts from patients and caregivers.

At project initiation, human experts develop key research questions, identify inclusion/exclusion criteria, and establish data sources. This is followed by gathering and processing SML data from online platforms, supported by AI-enabled screening and classification methods that use a retrainable ML model. Outputs

from AI-enabled processes are reviewed throughout the workflow by human experts who provide feedback to iteratively refine the proprietary AI-enabled methodologies to further improve the signal-to-noise ratio. A sufficiently trained ML model can then be employed for each concept of interest based on predefined research question(s), providing relevant posts for analysis via NLP identification of qualitative data for synthesis of data visualization. Further, the insights and analyses produced from these steps can be leveraged to further refine the workflow and address key research questions. The [Supplementary Material](#) provides a detailed description of the consecutive steps and a general overview of this flexible, iterative, and configurable workflow that involves human-led initiatives, AI-enabled steps, and steps of iterative refinement. Given the broad array of online platforms dedicated to oncology and patients with cancer, as described earlier, we propose using this process to identify records from patients with cancer and caregivers of interest.

4 Oncology case studies using the AI-enabled SML workflow

To explore the applicability of this workflow, we conducted 2 studies in oncology: 1 in patients with head and neck cancers and another in patients with esophageal cancer. The complete methods and key findings from these case studies are reported elsewhere (29). In brief, we explored the following research questions: (1) What are patients' expectations relating to treatments, and which treatment attributes are most important to patients? and (2) Which symptoms do patients find most burdensome? We then established social media sources of interest (e.g., forums, blog posts, WordPress posts (30), X[®] (31)) from 7 countries (United States, United Kingdom, Japan, Germany, Spain, France, and Italy), including all languages, with non-English posts translated into English. Identified posts were analyzed by the AI-enabled algorithm and classified as patient, caregiver, or irrelevant. Human experts then reviewed the posts and filtered irrelevant posts with specific tags to refine the AI-enabled algorithm. Relevant posts were then exported to ATLAS.ti for qualitative coding (32).

Through these pilot studies, we found that our customizable AI-enabled workflow, with iterative human expert-led training and ML, effectively identified patient and caregiver posts to reveal valuable concepts related to our research questions. Specifically, the workflow identified relevant posts relating to the most frequent products/treatment categories, potential AEs, symptom severity, and psychosocial health. We note that iterative engagement of experts with the AI-enabled methods was essential to refine algorithms and enhance the utility of the workflow.

5 Limitations and challenges encountered while using the AI-enabled SML workflow

Several limitations impacted the insights gained with this SML approach. Although it was possible to gather relevant data by

indication and treatment type, the specific type and/or name of chemotherapy treatment was rarely detailed in posts. Additionally, when posts included >1 treatment type, it was difficult to determine which descriptions were associated with an individual treatment type, confounding how distinct therapies correlate with specific patient experiences. Further, patients' expectations were based on various points of reference, such as information provided by their care team, online sources or other social media posts, and/or fear of the side effects of cancer treatment. Therefore, interpretation of the patient experience would be impacted and possibly biased by individuals' prior experiences.

Throughout the qualitative analysis processes, we discovered that it was challenging to distinguish between cancer-related symptoms and treatment-related symptomatic AEs. Patients frequently posted about symptomatic AEs and the burden from treatment; however, at times, it was unclear whether those symptoms were also cancer related and, if so, it was challenging to associate specific symptomatic AEs with specific treatments. Despite the utility of our novel and customizable workflow, there remained a high proportion of irrelevant entries, even in posts tagged with high confidence by the AI-enabled algorithm, and it was difficult to determine how many patients experienced symptoms out of a total population given repeat or follow-on posts. [Supplementary Table 1](#) includes sample posts, with ATLAS.ti coding visualizations, and an overview of challenges faced when interpreting the posts.

Previous reports of SML studies have identified additional limitations with SML strategies, including the self-selection bias among those who post on social media, as patients who choose to do so may not be representative of the entire target group (5, 9–11, 33). Thus, some patient groups may be underrepresented in social media-based analyses, such as older patients, low literacy patients, and patients from some geographic regions (10, 13). Additionally, negative perceptions may be posted more commonly than positive experiences, potentially skewing the evaluation of a given issue (11). Determining the volume of data needed is another challenge with SML. For instance, fewer posts may be needed to gather data on breast cancer than on rare diseases, given high occurrence rates and public awareness. A scoping review of 77 documented uses of AI-based SML by Tricco et al. reported a median (range) duration of data collection of 1.13 years (6 months–7 years) and a median (range) number of social media posts of 42,594 (4608–711,562) (34). These ranges highlight the high variability of data volume used in SML efforts.

6 Discussion

This article describes the value and utility of our newly developed, AI-enabled SML workflow. We describe learnings using 2 oncology case studies that produced descriptions of patient experiences with varying degrees of granularity, including high-level list-type descriptions of treatment regimens, detailed treatment experiences, symptomatic AEs, the impact of treatment on day-to-day life, and psychosocial impacts.

Incorporating evidence and insights from SML can elevate patients' voices in PFDD, adding valuable perspectives that may go unheard when using more traditional methods, such as patient and caregiver advisory boards, evidence-generation plans, and patient-centered outcome endpoints. Although prospective, qualitative patient interviews are the gold standard for understanding health impacts directly from patients, SML studies can provide greater geographic range and more varied input, and can be used synergistically with patient interviews to enhance the confidence of the gathered input (35, 36). Further, as SML methodologies continue to improve, a greater range of patient input can be incorporated into PFDD.

This SML workflow may be particularly useful when investigating therapies with high impacts on QoL or daily functioning, as is the case for many cancer treatments that lead to most patients experiencing side-effect profiles that would be unacceptable for other non-life-threatening diseases. New drugs continue to be developed across cancer indications, and our AI-enabled SML workflow has the potential to generate key patient/caregiver data. This information may be directly implemented to address the unique challenges of cancer treatments (e.g., identifying emerging unmet needs, how disease/treatment impacts daily living, the preferred method of drug administration, time requirements/how long patients are willing to spend at a doctor's office, and what patients may prioritize when choosing treatments).

Ideally, in drug development, researchers collect this type of data on patients' perspectives and integrate this information with data on treatment outcomes, ultimately providing all data to the FDA during regulatory review. When successfully executed, this process results in the approval of new medicines that are both efficacious and reflective of information that is meaningful to patients, caregivers, and providers (37). Furthermore, this workflow could be used to collect real-world data and identify perspectives from minority groups that are often excluded from clinical trials, which can be submitted as evidence to the FDA. Finally, the workflow may have added utility in collecting important postmarked surveillance data, such as patient satisfaction, patient compliance, and concerns that may interfere with future success (patient trust, insurance coverage).

The workflow reported here for AI-enabled SML offers a model for iterative engagement through the initial quality check/classification, continuous review of the AI-enabled outputs, and examination of the qualitative analysis. Furthermore, human-led fine-tuning will allow this workflow to adapt with the evolving needs and expectations of patients, thus priming AI-based SML models to extract data efficiently and effectively in an automated manner without losing that human guidance. Continued use of the dashboard in real-world settings further improved quality, and this approach has potential value in multiple patient populations of interest.

There are multiple future directions for this AI-enabled SML workflow. One is developing qualitative analysis plans that inform the drug development process in a way that prioritizes meaningful outcomes for patients and caregivers. We are also exploring the capture of quantitative metrics that incorporate

patient data or involve the patients themselves in dashboard development and algorithm refinement. The implementation of patient-centric terms, concerns, and conversational tone may allow for a better understanding and interpretation of context within patient or caregiver posts. **Figure 2** shows a conceptual overview on the benefits and uses derived from our AI-enabled

SML workflow, its adoption and use, and its ultimate implementation to inform PFDD strategies.

Utilizing AI-enabled models to gather patient-focused data from the plethora of social media platforms and to categorize large amounts of data is an opportunistic strategy; however, methods require frequent fine-tuning due to the constant

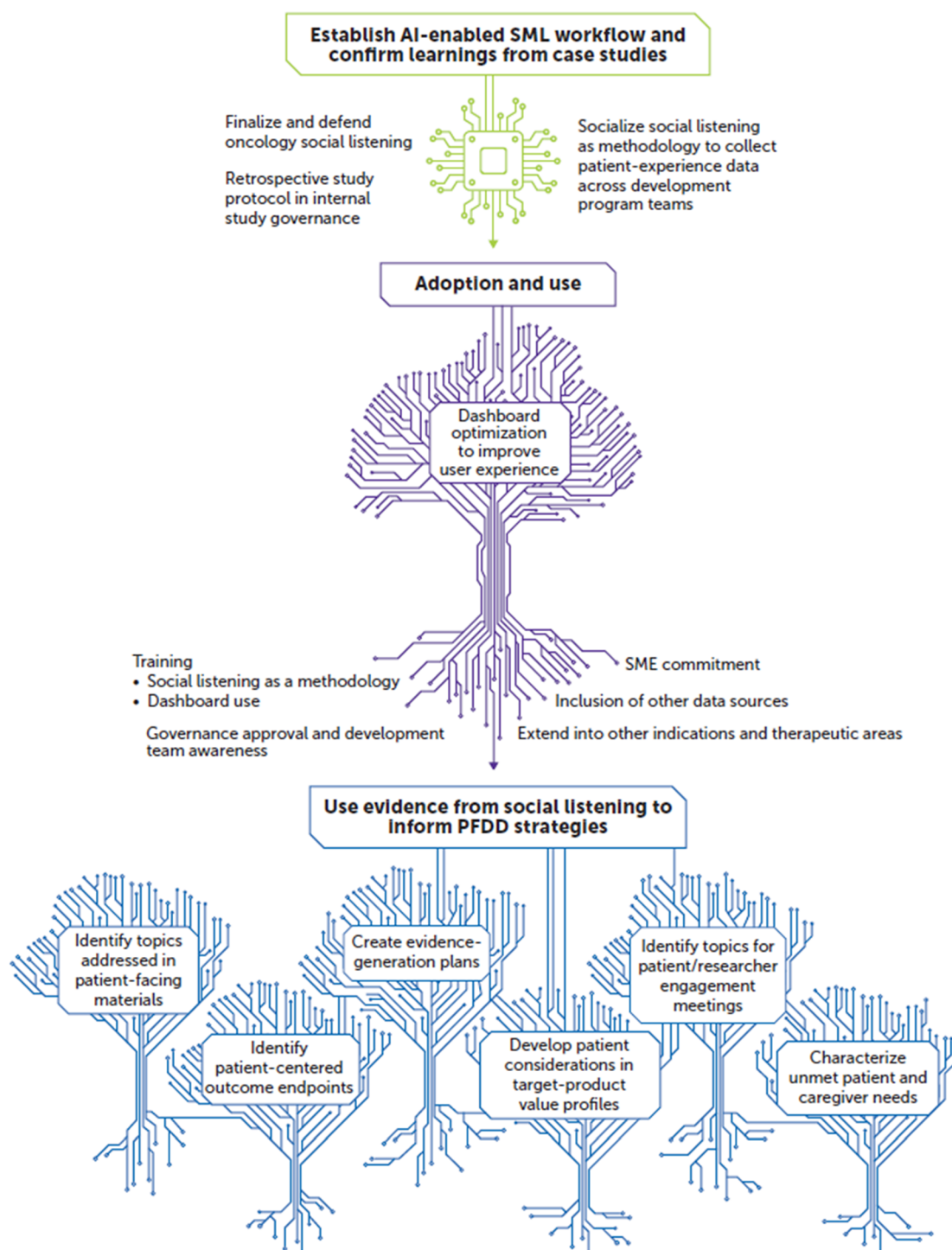


FIGURE 2

Future outlook: working toward a patient-experience knowledge hub. AI, artificial intelligence; PFDD, patient-focused drug development; SME, subject matter expert; SML, social media listening.

evolution of patient needs and expectations and technological advances. Our approach advances the tools available for SML studies and adds to the methodologic quality and rigor from the perspective of researchers contributing to, conducting, and evaluating SML studies in a PFDD context. This perspective manuscript provides further evidence on the utility of SML approaches to gain information on patients' experiences, which align with guidance offered by the FDA and the European Medicines Agency authorities for PFDD. Using SML to inform PFDD complements traditional methods of gathering patient input and enhances many aspects of drug development.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

ES: Conceptualization, Methodology, Writing – review & editing. JF: Conceptualization, Methodology, Writing – review & editing. NO: Conceptualization, Methodology, Writing – review & editing. PK: Conceptualization, Methodology, Writing – review & editing. HG: Conceptualization, Methodology, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This project was sponsored by Merck Healthcare KGaA, Darmstadt, Germany, and EMD Serono Research & Development Institute, Inc., Billerica, MA, United States, a business of Merck KGaA, Darmstadt, Germany. The work was executed in the absence of any external funding.

References

1. Borup G, Bach KF, Schmiegelow M, Wallach-Kildemoes H, Bjerrum OJ, Westergaard N. A paradigm shift towards patient involvement in medicines development and regulatory science: workshop proceedings and commentary. *Ther Innov Regul Sci.* (2016) 50(3):304–11. doi: 10.1177/2168479015622668
2. US Food and Drug Administration. *Patient-Focused Drug Development: Collecting Comprehensive and Representative Input.* (2020). Available online at: <https://www.fda.gov/media/139088/download> (accessed September 23, 2024).
3. US Food and Drug Administration. *Patient-Focused Drug Development: Methods to Identify What is Important to Patients.* (2022). Available online at: <https://www.fda.gov/media/131230/download> (accessed September 23, 2024).
4. European Medicines Agency. *ICH Reflection Paper on Proposed ICH Guideline Work to Advance Patient Focused Drug Development.* (2021). Available online at: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-reflection-paper-proposed-ich-guideline-work-advance-patient-focused-drug-development_en.pdf (accessed September 23, 2024).
5. Patalano F, Gutzwiller FS, Shah B, Kumari C, Cook NS. Gathering structured patient insight to drive the PRO strategy in COPD: patient-centric drug development from theory to practice. *Adv Ther.* (2020) 37(1):17–26. doi: 10.1007/s12325-019-01134-x
6. Humphrey L, Willgoss T, Trigg A, Meysner S, Kane M, Dickinson S, et al. A comparison of three methods to generate a conceptual understanding of a disease based on the patients' perspective. *J Patient Rep Outcomes.* (2017) 1(1):9. doi: 10.1186/s41687-017-0013-6
7. Mendelson E, Nast J, D'Alessio D, Aasaitambi S, Chauhan J, Samavedam S. PCN328 understanding of patient experience with diffuse large B-cell lymphoma (DLBCL) through social media listening. *Value Health.* (2020) 23(Suppl 1):S82. doi: 10.1016/j.jval.2020.04.1788
8. Goel K, Parashar N, Aasaitambi S, Verma H, O'Brien P, Sloesen B. P33 patient experiences and insights on chronic ocular pain from a social media listening study. *Value Health.* (2022) 25(Suppl 1):S8. doi: 10.1016/j.jval.2021.11.032
9. Martinez B, Dailey F, Almario CV, Keller MS, Desai M, Dupuy T, et al. Patient understanding of the risks and benefits of biologic therapies in inflammatory bowel disease: insights from a large-scale analysis of social media platforms. *Inflamm Bowel Dis.* (2017) 23(7):1057–64. doi: 10.1097/MIB.0000000000001110
10. Koss J, Rheinlaender A, Truebel H, Bohnet-Joschko S. Social media mining in drug development—fundamentals and use cases. *Drug Discov Today.* (2021) 26(12):2871–80. doi: 10.1016/j.drudis.2021.08.012

Acknowledgments

We would like to acknowledge the contributions of Nancy Schoenherr, MSc, and Paul Kamudoni, MSC, PhD, toward the completion of this work. Medical writing and editorial support were provided by Austin Horton, PhD, of Humanity Communications Inc., and were funded by EMD Serono Research & Development Institute, Inc.

Conflict of interest

PK and HG are current employees of Merck Data & AI organization, Merck IT Centre. ES and JF were employed by EMD Serono Research & Development Institute, Inc. at the time the work was completed. NO is a current employee of Healthcare Business of Merck KGaA.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fdgth.2024.1459201/full#supplementary-material>

11. Cook N, Mullins A, Gautam R, Medi S, Prince C, Tyagi N, et al. Evaluating patient experiences in dry eye disease through social media listening research. *Ophthalmol Ther.* (2019) 8(3):407–20. doi: 10.1007/s40123-019-0188-4
12. Voillot P, Riche B, Portafax M, Foulquie P, Gedik A, Barbarot S, et al. Social media platforms listening study on atopic dermatitis: quantitative and qualitative findings. *J Med Internet Res.* (2022) 24(1):e31140. doi: 10.2196/31140
13. Cook NS, Kostikas K, Gruenberger JB, Shah B, Pathak P, Kaur VP, et al. Gutzwiller: patients' perspectives on COPD: findings from a social media listening study. *ERJ Open Res.* (2019) 5(1):00128–2018. doi: 10.1183/23120541.00128-2018
14. Shinkhara R, Attai D, Scheitler AJ, Babey S, Glenn B, Ponce N. Using a twitter chat to rapidly identify barriers and policy solutions for metastatic breast cancer care: qualitative study. *JMIR Public Health Surveill.* (2021) 7(1):e23178. doi: 10.2196/23178
15. Mazza M, Piperis M, Aasaithambi S, Chauhan J, Sagkriotis A, Vieira C. Social media listening to understand the lived experience of individuals in Europe with metastatic breast cancer: a systematic search and content analysis study. *Front Oncol.* (2022) 12:863641. doi: 10.3389/fonc.2022.863641
16. Toguri JT, Grant-Nunn L, Urquhart R. Views of advanced cancer patients, families, and oncologists on initiating and engaging in advance care planning: a qualitative study. *BMC Palliat Care.* (2020) 19(1):150. doi: 10.1186/s12904-020-00655-5
17. Tapi Nzali MD, Bringay S, Lavergne C, Mollevi C, Opitz T. What patients can tell US: topic analysis for social media on breast cancer. *JMIR Med Inform.* (2017) 5(3):e23. doi: 10.2196/medinform.7779
18. Arun P, Ma Q, Bollu V, Aasaithambi S, Verma H, D'Alessio D, et al. PCN206 evaluating patient experiences in follicular lymphoma (FL) through social media listening (SML) research from five countries in North America and Europe. *Value Health.* (2021) 24(Suppl 1):S58. doi: 10.1016/j.jval.2021.04.297
19. Chauhan J, Aasaithambi S, Sagkriotis A, Lau M, Marquez-Rodas I, Formisano L, et al. POSA357 social media listening for melanoma care across European markets. *Value Health.* (2022) 25(Suppl 1):S219. doi: 10.1016/j.jval.2021.11.1068
20. Im EO, Chee W, Tsai HM, Lin LC, Cheng YC. Internet cancer support groups: a feminist analysis. *Cancer Nurs.* (2005) 28(1):1–7. doi: 10.1097/00002820-200501000-00001
21. Bender JL, Jimenez-Marroquin MC, Jadad AR. Seeking support on facebook: a content analysis of breast cancer groups. *J Med Internet Res.* (2011) 13(1):e16. doi: 10.2196/jmir.1560
22. Charbonneau DH, Hightower S, Katz A, Zhang K, Abrams J, Senft N, et al. Thompson: smartphone apps for cancer: a content analysis of the digital health marketplace. *Digit Health.* (2020) 6:2055207620905413. doi: 10.1177/2055207620905413
23. Schmidt AL, Rodriguez-Esteban R, Gottowik J, Leddin M. Applications of quantitative social media listening to patient-centric drug development. *Drug Discov Today.* (2022) 27(5):1523–30. doi: 10.1016/j.drudis.2022.01.015
24. Frank PP, Lu MXE, Sasse CE. Educational and emotional needs of patients with myelodysplastic syndromes: an AI analysis of multi-country social media. *Adv Ther.* (2023) 40(1):159–73. doi: 10.1007/s12325-022-02277-0
25. Gries KS, Fastenau J. Using a digital patient powered research network to identify outcomes of importance to patients with multiple myeloma. *J Patient Rep Outcomes.* (2020) 4(1):74. doi: 10.1186/s41687-020-00242-5
26. Fang C, Markuzon N, Patel N, Rueda DJ. Natural language processing for automated classification of qualitative data from interviews of patients with cancer. *Value Health.* (2022) 25(12):1995–2002. doi: 10.1016/j.jval.2022.06.004
27. Faust G, Booth A, Merinopoulou E, Halhol S, Tosar H, Nawaz A, et al. The experiences of patients with adjuvant and metastatic melanoma using disease-specific social media communities in the advent of novel therapies (excite project): social media listening study. *JMIR Cancer.* (2022) 8(2):e34073. doi: 10.2196/34073
28. Freeman TCB, Rodriguez-Esteban R, Gottowik J, Yang X, Erpenbeck VJ, Leddin M. A neural network approach for understanding patient experiences of chronic obstructive pulmonary disease (COPD): retrospective, cross-sectional study of social media content. *JMIR Med Inform.* (2021) 9(11):e26272. doi: 10.2196/26272
29. Karmalkar P, Gurulingappa H, Spies E, Flynn AJ. Artificial intelligence-driven approach for patient-focused drug development. *Front Artif Intell.* (2023) 6:1237124. doi: 10.3389/frai.2023.1237124
30. WordPress.com. /hp-2024-jul — WordPress.com. (n.d.). Available online at: <https://wordpress.com/> (accessed September 23, 2024).
31. X.com. X (Formerly Twitter). (n.d.). Available online at: <https://x.com/?lang=en> (accessed September 23, 2024).
32. ATLAS.ti Scientific Software Development GmbH. ATLAS.ti (Version 9.0) [Computer software]. Available online at: <https://atlasti.com> (accessed September 23, 2024).
33. Risson V, Saini D, Bonzani I, Huisman A, Olson M. Patterns of treatment switching in multiple sclerosis therapies in us patients active on social media: application of social media content analysis to health outcomes research. *J Med Internet Res.* (2016) 18(3):e62. doi: 10.2196/jmir.5409
34. Tricco AC, Zarin W, Lillie E, Jeeblee S, Warren R, Khan PA, et al. Utility of social media and crowd-intelligence data for pharmacovigilance: a scoping review. *BMC Med Inform Decis Mak.* (2018) 18(1):38. doi: 10.1186/s12911-018-0621-y
35. Findley A, Sharma G, Bentley S, Arbuckle R, Patalano F, Naujoks C, et al. Comparison of qualitative methods to explore key symptom and functional impact concepts of presbyopia: literature review, social media listening, and qualitative interviews. *Qual Life Res.* (2020) 29(suppl 1):s44.
36. Findley A, Sharma G, Bentley S, Arbuckle R, Patalano F, Naujoks C, et al. Comparison of literature review, social media listening, and qualitative interview research methods in generating patient-reported symptom and functional impact concepts of presbyopia. *Ophthalmol Ther.* (2023) 12(1):501–16. doi: 10.1007/s40123-022-00620-w
37. Pharmaceutical Research and Manufacturers of America. *Patient-Focused Drug Development.* (2024). Available online at: <https://www.phrma.org/-/media/Project/PhRMA/PhRMA-Org/PhRMA-Org/PDF/P-R/patient-focused-drug-development.pdf> (accessed September 23, 2024).