Title

A practical guide for conducting qualitative research in medical education: Part 3—Using software for qualitative analysis

Authors

Clarke, Samuel O
Coates, Wendy C
Jordan, Jaime

Peer reviewed

COMMENTARY - INVITED

A practical guide for conducting qualitative research in medical education: Part 3—Using software for qualitative analysis

Samuel O. Clarke MD, MAS[1] | Wendy C. Coates MD[2,4] | Jaime Jordan MD, MAEd[2,3]

[1]Department of Emergency Medicine, University of California, Davis Health System, Sacramento, California, USA

[2]Department of Emergency Medicine, University of California, Los Angeles, David Geffen School of Medicine at UCLA, Los Angeles, California, USA

[3]Department of Emergency Medicine, Ronald Reagan UCLA Medical Center, Los Angeles, California, USA

[4]Department of Emergency Medicine, Harbor–UCLA Medical Center, Torrance, California, USA

Correspondence
Samuel Clarke, UC Davis Department of Emergency Medicine, 4150 V. St, PSSB 2100, Sacramento, CA 95817, USA.
Email: soclarke@ucdavis.edu

Supervising Editor: Anne Messman, MD, MHPE.

Abstract

The process of performing qualitative analysis can be a daunting task. Technology can be employed to ease the burden of the work; however, the researcher may not fully appreciate how and when computer software can assist in conducting qualitative analysis. In this, the third installment of our "how-to" series on qualitative research methods, we describe basic concepts and approaches to using both simple word processing programs and specific qualitative research software programs to assist in coding and analysis. We hope that the concepts put forth in this paper will help qualitative researchers become more familiar with available technological approaches and that they will, in turn, enhance the efficiency of the research process as well as the depth, clarity and richness of research findings.

KEYWORDS
education research, qualitative analysis, qualitative research methods, software

APPROACHES TO CODING: THE INTERSECTION OF SUBSTRATE AND ANALYTIC APPROACH

Qualitative research is defined by a multiplicity of approaches.[1,2] Data collection, analysis, and interpretation are mutually influenced by the researcher's epistemologic and ontologic stance, the nature of the research question, and the implicit and explicit influence of theoretical frameworks.[3] While the traditions and approaches that guide qualitative research vary widely, a common thread is this: to conduct qualitative analysis, the researcher must first arrive at a system for organizing, displaying, and coding data.[3]

The primary substrate for qualitative research is words organized into *text*, whether in the form of field notes, transcripts of interviews and focus groups, or open-text responses to survey questions. The researcher organizes text into "chunks"—words, phrases, sentences, or paragraphs that refer to a specific idea or theme—and assigns descriptive codes to chunks to signify meaning (please see parts 1 and 2 of our series for further details on qualitative interviewing and coding).[3-5]

To avoid confusion, when we use the term "coding" in this paper, we are referring to the process of organizing qualitative data into meaningful themes that can be analyzed to shed light on the research question, not to the process of generating computer code for software development. The process of coding signifies the initiation of analysis, which Marshall and Rossman[6] describe as "the process of bringing order, structure and meaning" to collected data.[7]

USING SOFTWARE TO ASSIST IN CODING AND ANALYSIS: WHY AND WHICH TYPE?

The process of coding can be done by hand, but the use of software offers a number of advantages. It allows for the fast retrieval and comparison of chunks of text and for assigning and keeping track of codes. Software also allows the researcher to easily cluster codes for the purpose of comparison, enabling the recognition of patterns within data and laying the foundation for higher-order analysis and theory generation.[8]

What type of software is best for qualitative coding? The answer lies in the confluence of the research goals, analytic approach, and the length and complexity of the data to be coded. Text that is brief or that is bounded by specific themes and questions (as in a survey) may be parsed with little need for specific qualitative coding software. Lengthy or less structured interviews, focus groups with multiple participants, or detailed field notes in ethnographic research, on the other hand, will inevitably produce text that requires significant detailed analysis to organize and draw meaningful comparisons and conclusions. Similarly, studies that seek to describe phenomena in broad terms may not require the functionality of qualitative analysis software, whereas studies that rely on granular analysis (e.g., frequency coding) will benefit strongly from this technology. In the following sections we will discuss two software-enabled approaches to coding; their respective indications; and pros and cons based on the study goals, analytic strategy, and data to be analyzed (Table 1).

A SIMPLE APPROACH TO "DIY" CODING

To be clear, coding software is not a requirement for using technology to assist in conducting qualitative research. Here we describe a simple "do it yourself (DIY)" approach to coding using word processing programs such as Microsoft Word, Excel, and Google Docs.

For thematic analysis, a word processing program such as Microsoft Word is often sufficient. Text can be chunked by theme using highlighting, with each theme represented by a color. Chunks corresponding to a given theme can then be moved to a second document where they can be compared. Detailed and collaborative coding can be performed using the "comments" function. The DIY approach also has the advantage of using software that is familiar and readily available, limiting issues of cost and learning a new program. Further

description of general approaches to coding can be found in part 2 of our "how-to" series.[5]

USING QUALITATIVE ANALYSIS SOFTWARE

Computer-assisted qualitative data analysis software (CAQDAS, also called QDAS) has existed since the 1980s and has seen prolific growth over the past two decades.[8,9] Over a dozen software programs exist for the purpose of conducting qualitative and mixed-methods research, all of which share the common functions of searching and linking, coding, mapping, and annotating qualitative data.[3,9] Features of commonly used programs can be found in Table 2.[10-16]

CAQDAS offers a number of purported advantages to the researcher, chief among which is that it improves the efficiency of data organization and coding.[8] The time saved in the clerical work of data preparation frees the researcher to engage in the more meaningful work of analysis.[17] CAQDAS programs allow the researcher to easily create simultaneous codes (multiple codes applied to the same text) and subcodes (codes within sections of text that have already been coded). CAQDAS also increases the auditability of data, and programs that allow for collaborative coding by multiple investigators may serve to strengthen the overall rigor and trustworthiness of research findings.[9]

These strengths, however, must be weighed against the expense of purchasing a software license and the time needed to master a new program. It has also been suggested that CAQDAS programs carry the risk of "disengagement from the data," in which the researcher focuses more on technique rather than meaning.[8,18,19] On a practical level, it offers few advantages over manual coding when dealing with small data sets with brief text.[9] CAQDAS programs may also be a poor fit for analytic strategies

TABLE 1 General indications for choosing a "do-it-yourself" (DIY) approach to coding with general word processing software versus CAQDAS

	DIY coding with word processing programs	CAQDAS software
Goals of study	Description, comparison	Theory generation and/or testing
Analytical approach	Broad, thematic	Granular (e.g., grounded theory)
Data sources	Brief and/or tightly bound to specific research questions (e.g., survey responses)	Lengthy interviews and focus group transcripts, detailed/multiple field notes, open-ended and nuanced text
Advantages	Easy and intuitive to use Inexpensive, widely available Easy to import text files Can search text within documents	Supports multiple data file types Allows for easy organization and searching between multiple documents Provides enhanced coding abilities, annotation and memo generation, visual representations of data
Disadvantages	Can't search between documents More labor intensive to keep multiple documents organized	Steeper learning curve Many require a paid license to use Potential risk of overcoding (losing the forest through the trees)

Abbreviation: CAQDAS, computer-assisted qualitative analysis software.

TABLE 2 Comparison of commonly used CAQDAS programs

Program name	Platform	Cost	Capabilities	Uses	Learning curve
ATLAS.ti	MacOS, Windows, iOS, Android, Cloud	+++	Coding, aggregation, searching and querying data, data visualization, transcription, collaborative analysis (cloud version), multilanguage Supports text, images, audio and video files, social media, geo data, survey data	Grounded theory, ethnography, discourse analysis	More difficult
DeDoose	Cloud	+	Coding, aggregation, searching and querying data, data visualization, collaborative analysis Supports text, audio and video files, images, and spreadsheets	Grounded theory, ethnography, discourse analysis, mixed-methods research	Easier
MAXQDA	MacOS, Windows, iOS, Android	++	Coding, aggregation, searching and querying data, data visualization, transcription, collaborative analysis, statistical analysis, multilanguage Supports text, images, audio and video files, social media, survey data	Grounded theory, ethnography, discourse analysis, mixed-methods research	More difficult
NVivo	MacOS, Windows, Cloud	+++	Coding, aggregation, searching and querying data, data visualization, transcription, collaborative analysis Supports text, audio and video files, webpages, and social media posts	Grounded theory, ethnography, discourse analysis, mixed-methods research	More difficult
QDA Miner Lite	Windows (can run on MacOS with supplemental software)	Free	Coding, aggregation, searching and querying data Supports text, spreadsheets, image files	Grounded theory, ethnography, discourse analysis	Easier
Quirkos	MacOS, Windows, Linux, Cloud	++	Coding, aggregation, searching and querying data, data visualization, collaborative analysis Supports text, social media, survey data	Grounded theory, ethnography, discourse analysis, mixed-methods research	Easier

Note: Multiple software packages are available to assist in analyzing data obtained by qualitative research methods.

Abbreviation: CAQDAS, computer-assisted qualitative analysis software.

that are heavily focused on the situated interpretation of meaning or nuanced structure of language.[9]

The following is an overview of the functionality of NVivo, the CAQDAS program used most commonly by academics and researchers worldwide.[9] Our selection of NVivo as an example is not intended as an endorsement of this particular CAQDAS, but rather a reflection of its prominence within the field. In addition to the worked example using NVivo in this article, a number of instructional manuals, workshops, and online videos are available to the novice.

ORGANIZING AND CODING DATA

NVivo has the ability to quickly import and organize files of almost any type: text, audio, and visual files; emails; and social media posts as well as spreadsheets and data files from statistical analysis packages.[9] The current iteration of NVivo can transcribe audio files into text. The ability to import and organize (as well as transcribe) data

files from multiple sources holds the potential for significant time and cost savings in the early stages of a research project. Further detail on data organization is provided in Figures 1-3.

NVivo allows assignment of codes freely or according to coding schemes constructed by the researcher a priori (Figures 4 and 5). Easily searching and performing both open and axial (crosslinking) coding supports reflexivity and constant comparison in the research process.[9,20]

BUILDING TRANSPARENCY AND TRUSTWORTHINESS

CAQDAS programs such as NVivo hold advantages beyond the organization and initial coding of qualitative data. Data-querying functions, as well as the ability to create memos and annotations, provide depth and clarity to the process of qualitative analysis (Figures 6 and 7). Trustworthiness, the central argument supporting the validity of qualitative research findings, rests on researchers' "demonstrating

their understanding of their context and data (credibility), showing consistency and lack of bias in data analysis (confirmability), providing enough detail for possible replication (dependability), and allowing for assessment of a study's outcomes in relation to other contexts (transferability)."[20] NVivo and similar CAQDAS programs help researchers demonstrate trustworthiness by examining coding specificity, auditability within and between researchers, and illuminating patterns that lead to higher-order coding.[20]
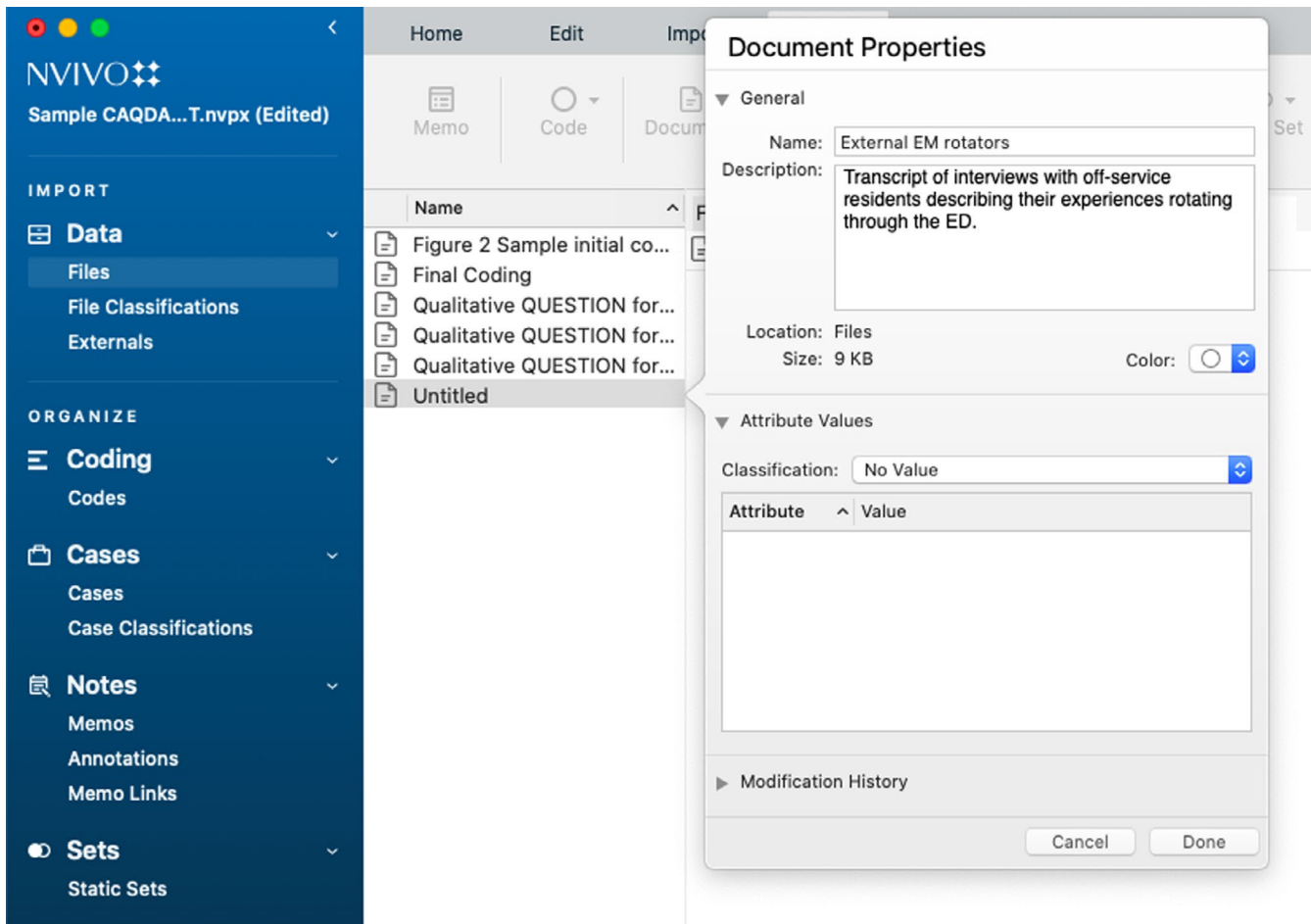
FIGURE 1 Importing files and assigning attributes. The first step in creating a project is uploading data files into the program. NVivo supports many file types (Word documents, pdfs, audio and visual files, Excel spreadsheets), which can simply be dragged and dropped into the program. The investigator can assign attributes (descriptive information, such as the date and location of an interview) to the files to further describe them
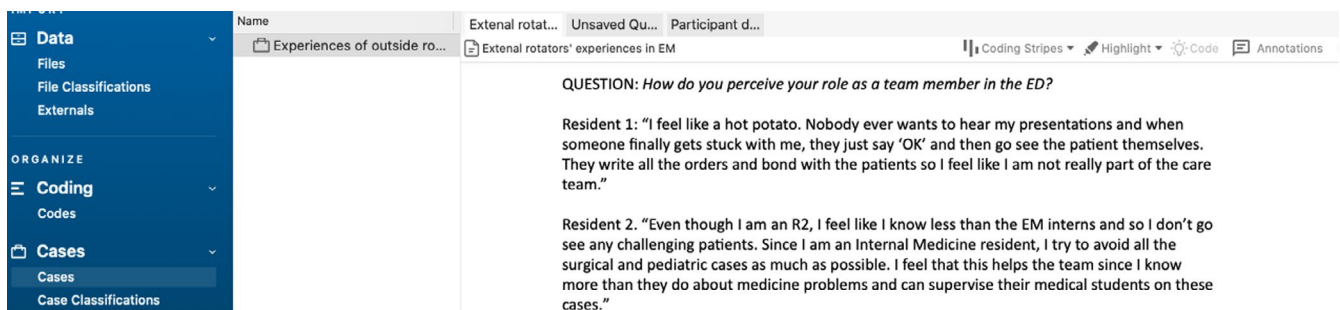
FIGURE 2 Creating cases. NVivo considers the unit of analysis in a qualitative project to be a "case." Cases can refer to individuals (e.g., department chair), sets of interviews or focus groups, or documents that represent a particular facet of a project. In this example, "Experiences of outside rotators in the Emergency Department" (a focus group transcript) will be assigned as a case

SUPPORTING THEORY GENERATION

CAQDAS programs such as NVivo have been used for inferential coding and theory generation in a number of qualitative approaches, most familiarly grounded theory.[21,22] NVivo can support the generation of theory by allowing the creation of detailed memos and data displays (Figure 7). Memos are an essential part of theory generation in grounded theory in that they allow the researcher to illustrate ideas, add clarification, and draw connections between categories to direct further data gathering and coding.[21,23] NVivo allows the researcher to create memos linked to particular nodes and "sets" of memos that link similar categories.[21] NVivo also allows researchers to create concept maps and visual models that help them explore and describe connections among categories (Figure 8). Taken together, these tools assist the researcher in recognizing and articulating emerging theory.[21]

CONCLUSION

In this article we have addressed the respective pros and cons of two technological adjuncts to coding qualitative data. We encourage both novice and experienced researchers to consider first the nature of their research question, their analytic strategy,
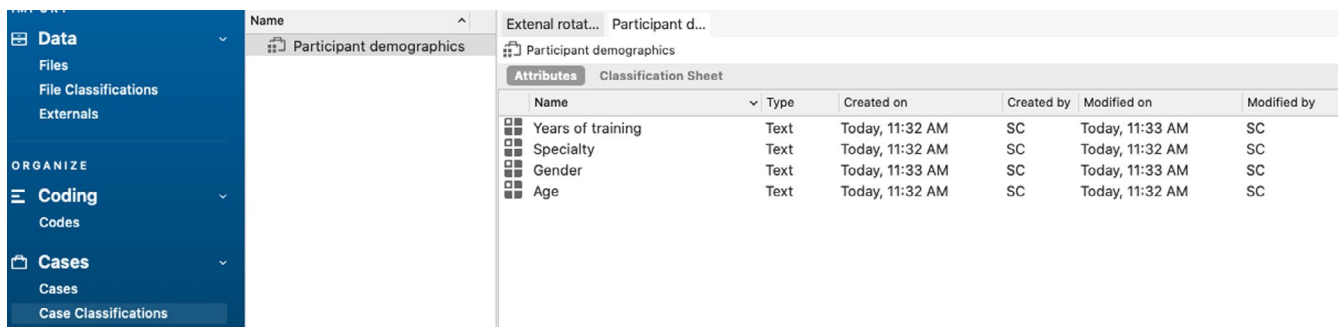
FIGURE 3 Creating case classifications. The investigator can assign case classifications (sets of attributes, such as age, gender, level of training) to facilitate analysis at later stages in the project
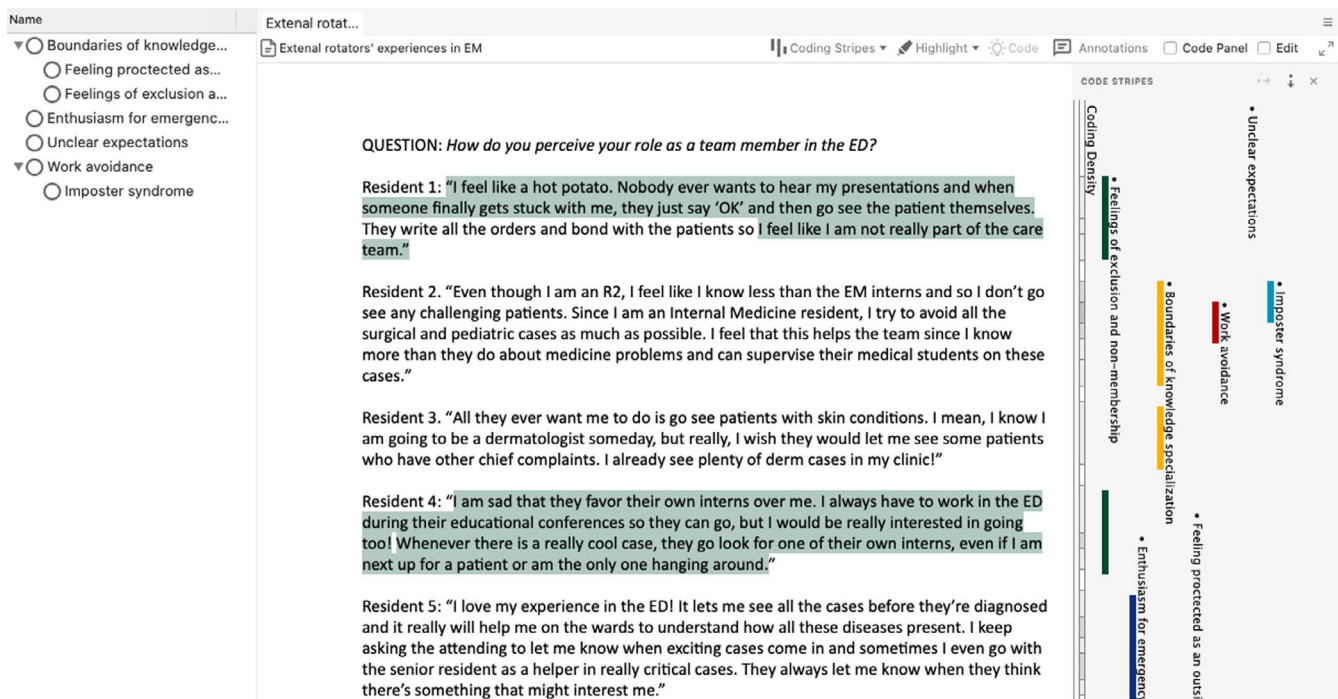
FIGURE 4 Coding. Once a case has been created, the investigator can begin to explore the data by highlighting text and assigning codes to highlighted chunks. Codes can be generated on the fly or according to an a priori coding scheme established by the investigator. The program allows for simultaneous coding (assigning more than one code to a section of text) as well as subcoding (creating new codes within an already coded section). Codes are shown in list form on the left column of the screen and as color-coded "coding stripes" on the right

Files\\Extenal rotators' experiences in EM
6 references coded, 24.09% coverage

Reference 1: 4.62% coverage

"I feel like a hot potato. Nobody ever wants to hear my presentations and when someone finally gets stuck with me, they just say 'OK' and then go see the patient themselves.

Reference 2: 1.30% coverage

I feel like I am not really part of the care team."

Reference 3: 4.81% coverage

I am sad that they favor their own interns over me. I always have to work in the ED during their educational conferences so they can go, but I would be really interested in going too!

Reference 4: 3.87% coverage

Whenever there is a really cool case, they go look for one of their own interns, even if I am next up for a patient or am the only one hanging around.
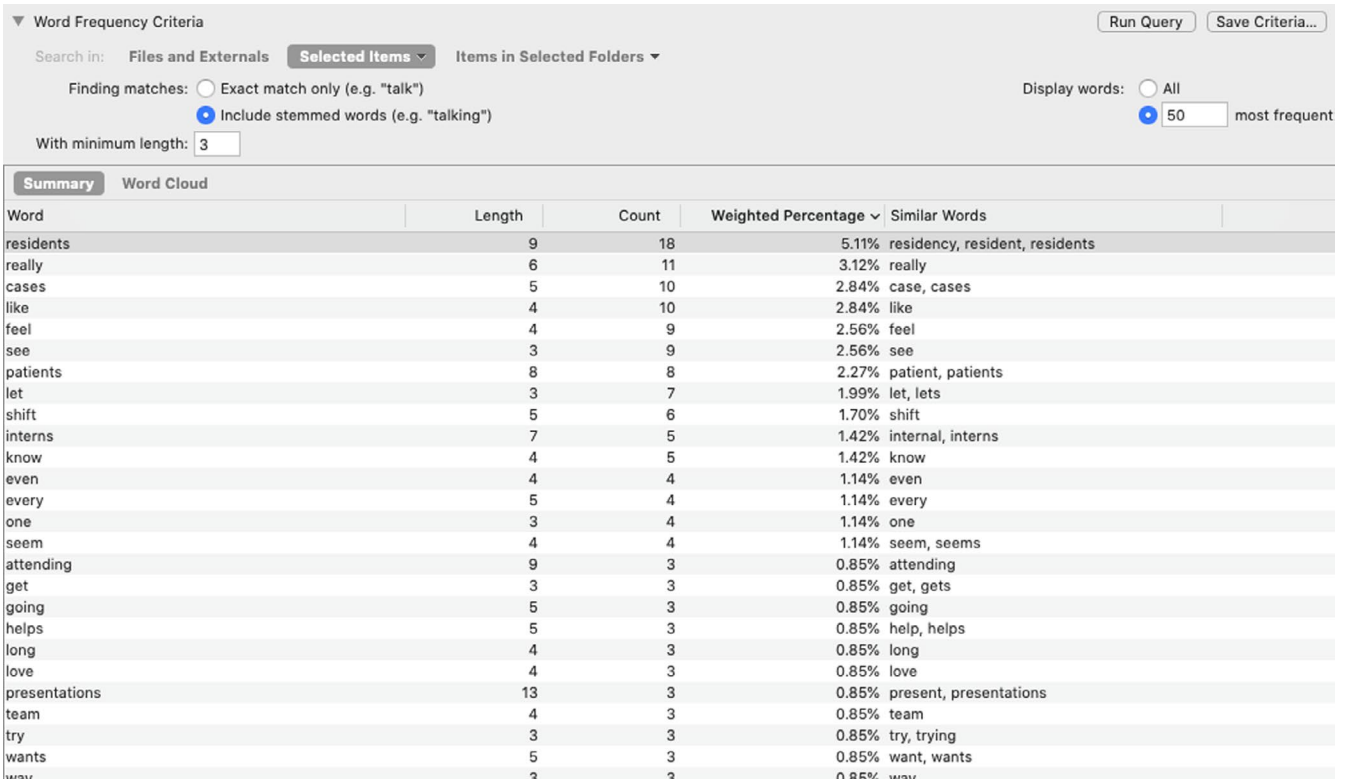
Reference 5: 4.03% coverage

I'd really like to learn more about different presentations, but I have to prove myself every single shift before I can take on this responsibility.

Reference 6: 5.46% coverage

I feel like it's us vs. them. The other rotators at every level get treated one way while the EM residents seem to be treated like gods. I mean, they could hide their favoritism a little bit, couldn't they?

FIGURE 5 Code displays. Text corresponding to particular codes can be viewed separately to facilitate further analysis

▼ Word Frequency Criteria   Run Query   Save Criteria…

Search in: **Files and Externals** | **Selected Items ▼** | Items in Selected Folders ▼

Finding matches: ○ Exact match only (e.g. "talk")   Display words: ○ All
                 ● Include stemmed words (e.g. "talking")   ● 50   most frequent
With minimum length: 3

| **Summary**   Word Cloud |

| Word	Length	Count	Weighted Percentage ⌄	Similar Words
residents	9	18	5.11%	residency, resident, residents
really	6	11	3.12%	really
cases	5	10	2.84%	case, cases
like	4	10	2.84%	like
feel	4	9	2.56%	feel
see	3	9	2.56%	see
patients	8	8	2.27%	patient, patients
let	3	7	1.99%	let, lets
shift	5	6	1.70%	shift
interns	7	5	1.42%	internal, interns
know	4	5	1.42%	know
even	4	4	1.14%	even
every	5	4	1.14%	every
one	3	4	1.14%	one
seem	4	4	1.14%	seem, seems
attending	9	3	0.85%	attending
get	3	3	0.85%	get, gets
going	5	3	0.85%	going
helps	5	3	0.85%	help, helps
long	4	3	0.85%	long
love	4	3	0.85%	love
presentations	13	3	0.85%	present, presentations
team	4	3	0.85%	team
try	3	3	0.85%	try, trying
wants	5	3	0.85%	want, wants
way	3	3	0.85%	way

FIGURE 6 Querying data. As analysis progresses, the investigator can query (ask questions of) the data in a number of ways, such word counts and crosstabs. Collectively, these tools assist the researcher in finding linkages between codes and identifying patterns as they analyze the data

FIGURE 7 Creating memos and annotations. As the investigator develops insights into the data, they can create memos and annotations and assign them to individual codes for easy retrieval
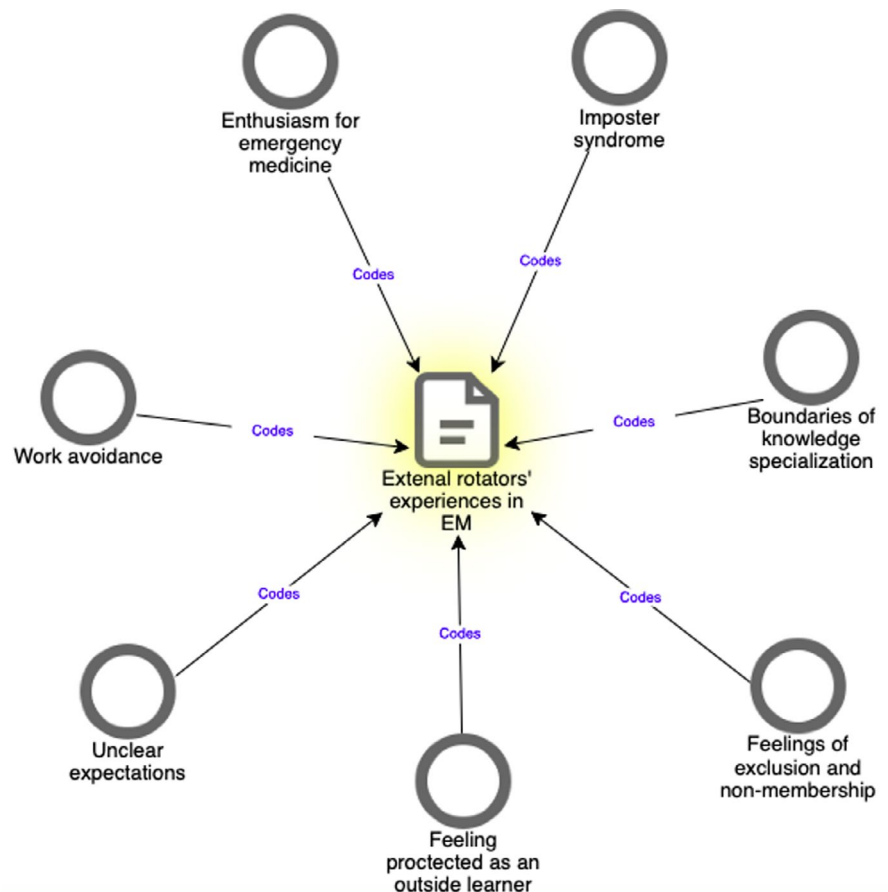
FIGURE 8 Visual representations of data. The culmination of qualitative analysis often involves visual representations of key findings such as theoretical models. NVivo offers a number of tools (e.g., mind maps, word clouds) to help the investigator create these visual representations

and the characteristics of the data to be analyzed prior to committing to a coding approach. While technology has the ability to strengthen and streamline the research process, it cannot replace the insight of the researcher in the critical work of qualitative analysis.

CONFLICT OF INTEREST

The authors have no disclosures to report.

ORCID

Samuel O. Clarke https://orcid.org/0000-0003-3762-1727

Wendy C. Coates https://orcid.org/0000-0002-3305-8802
Jaime Jordan https://orcid.org/0000-0002-6573-7041

REFERENCES

1. Merriam SB. *Qualitative Research: A Guide to Design and Implementation*. 2nd ed. City, State: Jossey-Bass; 2009.
2. Schneider NC, Coates WC, Yarris LM. Taking your qualitative research to the next level: a guide for the medical educator. *AEM Educ Train*. 2017;1(4):368-378.
3. Miles MB, Huberman AM. *Qualitative Data Analysis: An Expanded Sourcebook*. 2nd ed. Thousand Oaks, CA: Sage Publications, Inc.; 1994.
4. Jordan J, Clarke SO, Coates WC. A practical guide for conducting qualitative research in medical education: part 1—how to interview. *AEM Educ Train*. 2021;5(X):XXX.
5. Coates WC, Jordan J, Clarke SO. A practical guide for conducting qualitative research in medical education: part 2—coding and thematic analysis. *AEM Educ Train*. 2021;5(X):XXX.
6. Marshall C, Rossman GB. *Designing Qualitative Research*. 6th ed. Thousand Oaks, CA: SAGE Publications; 2014.
7. Este D, Sieppert J, Barsky A. Teaching and learning qualitative research with and without qualitative data analysis software. *J Res Comput Educ*. 2014;31(2):138-154.
8. Cope DG. Computer-assisted qualitative data analysis software. *Oncol Nurs Forum*. 2014;41(3):322-323.
9. Cypress BS. Data analysis software in qualitative research: preconceptions, expectations, and adoption. *Dimensions Critical Care Nurs*. 2019;38(4):213-220.
10. ATLAS.ti 9 website. c2002–2021. Accessed March 24, 2021. https://atlasti.com/
11. DeDoose website. c2018. Accessed March 24, 2021. https://www.dedoose.com/
12. MAXQDA website. c1995–2021. Accessed March 24, 2021. https://www.maxqda.com/
13. NVivo 12. QSR International website. c2021. Accessed March 24, 2021. https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home/
14. QDA Miner Lite. Provalis Research website. c2021. Accessed March 24, 2021. https://provalisresearch.com/products/qualitative-data-analysis-software/freeware/
15. Quirkos 2.3.1 website. c2021. Accessed March 24, 2021. https://www.quirkos.com
16. Choosing an Appropriate CAQDAS Package. Accessed March 24, 2021. https://www.surrey.ac.uk/computer-assisted-qualitative-data-analysis/resources/choosing-appropriate-caqdas-package
17. Silverman D. *Doing Qualitative Research: A Practical Handbook*. 2nd ed. Thousand Oaks, CA: Sage Publications; 2005.
18. Banner DJ, Albarrran JW. Computer-assisted qualitative data analysis software: a review. *Can J Cardiovasc Nurs*. 2009;19(3):24-31.
19. John WS, Johnson P. The pros and cons of data analysis software for qualitative research. *J Nurs Scholarsh*. 2000;32(4):393-397.
20. O'Kane P, Smith A, Lerman MP. Building transparency and trustworthiness in inductive research through computer-aided qualitative data analysis software. *Organ Res Methods*. 2021;24(1):104-139.
21. Bringer JD, Johnston LH, Brackenridge CH. Using computer-assisted qualitative data analysis software to develop a grounded theory project. *Field Method*. 2006;18(3):245-266.
22. Dalkin S, Forster N, Hodgson P, Lhussier M, Carr SM. Using computer assisted qualitative data analysis software (CAQDAS; NVivo) to assist in the complex process of realist theory generation, refinement and testing. *Int J Soc Res Method*. 2020;24(1):1-12.
23. Glaser BG, Strauss AL. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New Brunswick, NJ: AldineTransaction; 2009.

PennState Health
Emergency Medicine
Residency Program Director

Penn State Health Milton S. Hershey Medical Center is seeking an Emergency Medicine Residency Program Director to join our exceptional academic team located in Hershey, PA. This is an excellent opportunity to join an outstanding academic program with a national reputation and inpact the lives of our future Emergency Medicine physicians.

What We're Offering:
- Competitive salary and benefits
- Sign-On Bonus
- Relocation Assistance
- Leadership for Emergency Medicine Residency Program
- Comprehensive benefit and retirement options

What We're Seeking:
- MD, DO, or foreign equivalent
- BC/BE by ABEM or ABOEM
- Leadership experience
- Outstanding patient care qualities
- Ability to work collaboratively within a diverse academic and clinical environment

FOR MORE INFORMATION PLEASE CONTACT:

Heather Peffley, PHR CPRP
Physician Recruiter
Penn State Health

Email: hpeffley@pennstatehealth.psu.edu
Website: careers.pennstatehealth.org

What the Area Offers:

Located in a safe family-friendly setting, Hershey, PA, our local neighborhoods boast a reasonable cost of living whether you prefer a more suburban setting or thriving city rich in theater, arts, and culture. Known as the home of the Hershey chocolate bar, Hershey's community is rich in history and offers an abundant range of outdoor activities, arts, and diverse experiences. We're conveniently located within a short distance to major cities such as Philadelphia, Pittsburgh, NYC, Baltimore, and Washington DC.