ABSTRACT
Qualitative research, the analysis of nonquantitative and nonquantifiable data through methods such as interviews and observation, is integral to the field of biomedical and health informatics. To demonstrate the integrity and quality of their qualitative research, authors should report important elements of their work. This perspective article offers guidance about reporting components of the research, including theory, the research question, sampling, data collection methods, data analysis, results, and discussion. Addressing these points in the paper assists peer reviewers and readers in assessing the rigor of the work and its contribution to the literature. Clearer and more detailed reporting will ensure that qualitative research will continue to be published in informatics, helping researchers disseminate their understanding of people, organizations, context, and sociotechnical relationships as they relate to biomedical and health data.

Key words: qualitative research, qualitative methods, reliability, biases, data collection, data analysis

INTRODUCTION
Many informatics researchers, including those with training and expertise in quantitative or computational methods, come to recognize the value of applying qualitative methods. Qualitative research is the collection and analysis of nonquantitative and nonquantifiable data through methods such as interviews and observation to understand perspectives, beliefs, and experiences. Qualitative research is invaluable for understanding context, explaining phenomena and processes, understanding the rationale underlying behavior and decisions, generating hypotheses, and developing or extending theory about sociotechnical phenomena. In health and biomedical informatics, qualitative research is often used to gain insights into the patients and professionals who use informatics innovations, the contexts in which they live and work, and the life experiences that are the sources of the medical, technology, and digital trace data that informaticists analyze.

As a recent scoping review noted, qualitative research represents a small but important portion of published articles in JAMIA. In this perspective, we provide a brief outline of expectations for qualitative research published in informatics. Our goal is to explain how authors can demonstrate rigor and avoid common pitfalls by reporting how they have sought to reduce bias, improve reliability, and verify their findings. Moreover, we aim to provide guidance that researchers can use at the early stages of studies to ensure that they are able to rigorously report methodological details in a subsequent manuscript.

Theoretical rigor
All research, whether quantitative or qualitative, is strengthened with a firm foundation in relevant theory or an explanation of why new theory is needed. Authors should use their literature review to cite the relevant theory or theories that have grounded their work. For example, a study...
of the adoption of a novel technology would typically begin with a discussion of existing theories of technology diffusion and adoption. In some cases, existing theories are not adequate to describe the phenomenon being studied. For example, theories of technology adoption developed in advantaged populations may not be entirely relevant for understudied groups. In these cases, qualitative work may be needed to develop new concepts or theories or extend existing ones. Grounded theory research, in particular, is designed to develop new theory. Authors using grounded theory should consider discussing current theories and explaining where they fall short, which will strengthen the rationale for theory-building work.

Rigor in stating the research question and clarifying the study design and methodological perspective
Like other research papers, qualitative research papers should contain a clear statement of the research question. Authors should describe the gap in knowledge, discuss why the question needs to be answered, and explain why qualitative or mixed-methods research is appropriate to answer it. Qualitative research questions should not be phrased as hypotheses.

Describing the study design and stating the methodological approach that guided the work is extremely helpful, especially in an inherently multidisciplinary field such as informatics. The methodological approach should be aligned with the research question. For example, phenomenological research is appropriate for developing a nuanced, sensitive understanding of the lived experience of a phenomenon and the meaning attributed to it by those who experience it. Grounded theory researchers seek to develop novel social theories emerging from data analysis, especially around social processes. Recently, grounded theory-based mixed-methods research has also been proposed. Ethnography is useful for describing groups and interpreting their cultures, contexts, and shared meanings; such studies may use a combination of qualitative and quantitative data. A case study approach can prove particularly useful for evaluating interventions or programs or investigating critical events; these studies may also use mixed methods. Participatory design, usability research, and user-centered design are examples of qualitative or mixed-methods approaches intended to produce new technologies adapted to the needs and capabilities of their users and stakeholders.

When using mixed methods, researchers should state their selected study design. Examples include sequential exploratory (qualitative then quantitative) to discover a new phenomenon and then determine its generalizability; sequential explanatory (quantitative then qualitative) to explain unexpected findings from quantitative work; measure development (quantitative, qualitative, then quantitative) in which qualitative methods are used to develop instruments for quantitative surveys; and parallel mixed designs such as evaluating the implementation of a technology intervention through qualitative research while simultaneously collecting quantitative usage data.

*Elements with an asterisk may need to be elaborated in an appendix to avoid lengthening the manuscript.*
The informatics literature includes many excellent examples of studies conducted from the perspective of grounded theory, phenomenology, user-centered design, case study, ethnography, and other qualitative and mixed-methods frameworks.

Rigor in sampling and justification of sample size

Whether researchers work directly with participants or conduct secondary analyses of existing data, they should demonstrate rigor in their sampling approach. What was the population, community, culture, or phenomenon of interest to the research question, and how did the researchers obtain a sample from it? Researchers should explain their sampling approach. Purposeful (or purposeful) sampling describes a group of methods for recruiting nonprobability samples of individuals likely to have perspectives or experiences of interest. With purposive sampling, researchers should report how they identified the groups or perspectives for targeted recruiting (eg, whether it was on the basis of theory or empirical observation). If purposive sampling strategies such as quota sampling and typical case sampling were used to increase representativeness or oversample subgroups of interest, these methods should be described. In grounded theory, in which the goal is theory development, the sampling should be shown to be theoretically justified. For mixed-methods studies, the relationship between quantitative and qualitative samples and analyses should also be specified. Throughout, researchers should explain why their sampling method is appropriate to the research question, discuss whether sampling may be subject to biases, and discuss how they sought to address such biases.

Informatics journals have shown an increased interest in health equity, and authors are encouraged to describe any sampling approaches intended to maximize inclusion of historically marginalized and underserved populations. An informatics-relevant example is this study with men who have sex with men.

Researchers must also provide a convincing rationale for their sample size. A few papers in the qualitative literature have suggested that small sample sizes are sufficient for interviews and focus groups with human subjects. However, these citations are more suited to the planning stage of a qualitative project (as in a funding proposal to justify sample sizes) rather than at the execution and reporting stages of a study. A citation to previous research alone may not provide sufficient assurance that the sample size was adequate for a specific study. Instead, it is the responsibility of the researcher to demonstrate that the sample size is adequate to answer the research question.

One well-accepted sample size criterion is saturation. Researchers invoking saturation should explain which of the several definitions of saturation they used and how they determined that saturation was achieved. Researchers must also show how their criteria for sampling and sample size determination are harmonized with the research question and purpose. (For example, inductive thematic saturation can be achieved with small sample sizes when the population of interest is homogeneous or the research question is simple, but larger samples are usually needed when the population of interest has multiple strata of interest.)

Usability researchers are encouraged to note potential limitations of the Nielsen and the “10 ± 2” sample size heuristics and to recognize that larger samples will tend to help find more usability problems and allow input by more diverse users.

If the researchers sampled units of analysis other than the individual (eg, events, communities, organizations, or social media posts), this should also be justified according to the goals of the study. For example, in a case study, bellwether or ideal case sampling of hospitals might be justified for examining the effects of a clinical informatics intervention. Authors should draw upon any published methodological literature relevant to their sampling approach.

Rigor in data collection and the relationship between researcher and participants

Researchers must describe their data collection methods (eg, interviews, focus groups, observation, etc.), and justify how their choice of data collection method was appropriate for the research question. For mixed-methods studies, the approach to integrating qualitative and quantitative data collection should be explained. The selection or development process for any instruments used in the study should be described. For example, for semistructured interviews, how and why were topic areas selected and how were interview questions developed? For observation, if a template was used to guide data collection, how and why were template components selected? Any pilot testing of the instruments should be discussed, including any use of mixed methods to design the instrument. Instruments developed for the study should be included as appendices.

Researchers should describe how data collection was carried out. It should be clear who conducted interviews, observations, and focus groups, whether they were conducted in person or by telephone or video conference, and how observers and interviewers were trained. When multiple persons collected data, as in team-based ethnographic research, researchers should describe how they ensured reliability across data collectors. Researchers who created field notes should describe their creation (eg, free text notes, observation template).

The interpretivist nature of much qualitative research means that the relationship between the research participants and the researchers is of critical importance. In the research report, it can be helpful to provide a brief summary of the characteristics, training, and perspectives of the qualitative researchers to help readers assess the credibility of their work. Researchers should also be prepared to discuss how they addressed potential data collection biases. These include power differentials that may reduce patient candor when interviewed by physician-researchers or concerns about employment security or professional repercussions that may cause healthcare workers to be concerned about answering questions about their work. These also include the known tendencies for people to provide overly positive assessments of an innovation when interviewed by its developer and to change their behavior when they know they are being observed. For focus groups, researchers should report how they addressed common limitations such as group composition issues (eg, role hierarchies that influence participant discussion), or impacts of dominant personalities.

For publication, researchers should describe how they addressed and sought to mitigate these and other potential biases, for example, through reflexivity, prolonged engagement, or persistent observation. If reflexivity practices were used to help researchers understand their own relationship to the research question and the research participants, and their evolving understanding of the data over the course of the project, they should be described. Readers and reviewers value descriptions of reflexivity practices, especially in explaining relationships between researchers and marginalized participants, where the perpetuation of bias is likely. In some research traditions, positionality statements are increasingly used to make re-
flexibility practices in research with marginalized groups more explicit.48

Rigor in data analysis methods
Informatics journals require qualitative researchers to report how they analyzed their qualitative data and provide methodological citations. Many approaches to data analysis are available.3,49,50 What they have in common is that each involves 4 stages: (1) a method for systematically identifying patterns or concepts in the data, (2) a method for reliably labeling these patterns or concepts across different transcripts, fieldnotes, or collected documents/images/artifacts, (3) a method of discovering or identifying relationships between these concepts to synthesize themes or groups of themes comprising theories, and (4) methods to verify and test developing analyses.

These 4 steps are accomplished differently by different analysis approaches. If researchers used deductive approaches such as directed content analysis (which collect and code data on the basis of an existing theoretical framework that predefines the set of applicable concepts and relationships between them), they should report the theory or framework they used.51 By contrast, researchers who use inductive approaches such as thematic analysis49 (analyses conducted in the absence of an existing theoretical framework) should explain how they followed the inductive approach of immersing themselves in the data and allowing concepts and relationships to emerge from reading, review, theorizing, and discussion. Given the large number of first-cycle and second-cycle coding approaches available, researchers should cite which was used.52 Combined inductive-deductive analysis may use components of each of these analysis approaches. For mixed-methods studies, these 4 steps will typically be followed by at least one approach to integrating qualitative and quantitative data (see excellent texts42 on the variety of approaches).

Although inductive data analysis is a component of grounded theory development, not all forms of inductive data analysis meet the definition of grounded theory. Qualitative authors are encouraged to reserve the term “grounded theory” for projects that seek to develop novel theories about social phenomena and conform to one of the main approaches to doing so.50,53–55

Researchers to describe any approaches to improve dependability of coding, for example, whether multiple coders worked on the transcripts, and if so, how they worked together (eg, consensus meetings or establishment of inter-rater reliability). Audit trails (eg, memoing in grounded theory) can also be used for single-authored projects.3,56

Researchers should also describe any methods for improving credibility or verifying their interpretation of the data. If they apply triangulation, they should describe what data sources, researchers, or data types were compared.57 If they applied negative case analysis18,59 to seek out and analyze data that appear to disconfirm a developing concept or theory, they should describe how they identified the negative cases and how the analysis revealed patterns that did or did not hold true. If they exposed their interpretation for critique and reinterpretation by participants in the community being studied, methods for doing so should be described (eg, updating the interview guide to include emerging themes to be discussed with new participants). Alternately, if they conducted formal, terminal member checking by inviting participant discussion and review of findings, they should describe the method, the feedback provided, and how it was addressed in the analysis.60–62

The informatics literature contains many good examples of studies that apply inductive analysis,8 deductive analysis,9 mixed inductive–deductive analysis,64 and mixed qualitative–quantitative methods.55

Rigor in reporting results
The results section should contain a description of the sociodemographic and other relevant characteristics of any human subjects in the sample.

When themes are reported as part of the results, researchers should recognize that a theme cannot be sufficiently described in a sentence or phrase. Themes must be supported with rich examples of actual extracts, quotes, or images gathered during data collection. Quotes not only give life and interest to the research report but also serve the critical function of connecting the source data to the researchers’ interpretation. In addition to the quotes, researchers should provide synthesis by explaining themes and categories, and the depth and range of findings represented by those categories. Labeling quotes and extracts with study-specific labels such as “Participant 1” is one helpful approach to demonstrate that representative quotes are drawn from the entire sample of participants.

Word limits of biomedical research journals can introduce challenges with integrating rich and descriptive source quotations into the text of a results section. While quotes integrated into the results text are a powerful approach to building a high-quality results description, additional tools such as thoughtfully integrated tables of quotations, boxes for longer-form quotations, and summary visualizations such as timelines and network diagrams can also prove helpful. In mixed-methods studies, joint displays of qualitative and quantitative data can also help show linkages between data.66

Rigor in discussion and conclusions
Qualitative researchers are invited to discuss the potential transferability of their findings to other settings or populations.26 To help readers determine whether the findings might be transferable to other settings and populations, authors should provide detail about the context and setting of the research and about their assumptions about the domain of interest.

As with any publication, qualitative reports should also include a limitations section. Common limitations in qualitative work may include known and unavoidable lack of representation of certain participant perspectives, unavoidable researcher biases, or issues with the transferability of findings. Approaches used to address limitations, such as the use of bracketing to address potential researcher biases, should be discussed in the limitations section.

CONCLUSION
Qualitative research is integral to health and biomedical informatics. High-quality qualitative research is conducted by many informatics researchers whose original backgrounds are in quantitative methods. Examples of excellent qualitative and social science research honored with AMIA’s Diana Forsythe Award are available here (https://www.amia.org/amia-awards/working-group-awards). With this summary of publication expectations, we hope to encourage the submission of high-quality qualitative research that advances the field of informatics. Researchers might find it helpful to refer to some of the formal reporting checklists to learn more about reporting expectations in qualitative research, even when publishing in journals that do not require them.45,67 However, in recognition of restrictive manuscript lengths,
the present guidelines are intended to be more selective than prior formal reporting checklists. They are also crafted to emphasize unique issues arising in health informatics, such as frequent use of usability testing methods.

Addressing all the elements described here is likely to make the qualitative manuscript very long. Authors are encouraged to consider writing an online-only methodological appendix that would allow them to describe their approach in detail without violating length limits. Using tables, boxes, and figures to illustrate methods and results can also help authors stay within page limits.

We hope that this summary will support more complete reporting of research studies, but we recognize that it is not a substitute for training in qualitative research—although this summary can be an instructional tool within broader training programs. We encourage interested researchers to consult the references shared below and the many excellent texts, courses, and mentors available in informatics and beyond. Following the literature to staying abreast of the qualitative research can also alert researchers to innovations in theory and methods that can be applied to continue advancing our understanding of the patients and professionals who use informatics innovations, the contexts in which they live and work, and their beliefs, perspectives, and life experiences.

**FUNDING**

This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**AUTHOR CONTRIBUTIONS**

JSA, NCB, MR, KMU, and TV made substantial contributions to the conception and design of this work; drafted the work and revised it critically for important intellectual content; and gave final approval of the version to be published. All agree to be accountable for all aspects of the work.

**CONFLICT OF INTEREST STATEMENT**

None declared.

**DATA AVAILABILITY**

No data were generated in the course of this study.

**REFERENCES**

1. Hussain MI, Figueiredo MC, Tran BD, et al. A scoping review of qualitative research in JAMIA: past contributions and opportunities for future work. *J Am Med Inform Assoc* 2021; 28 (2): 402–13.

2. Korstjens I, Moser A. Series: practical guidance to qualitative research. Part 2: context, research questions and designs. *Eur J Gen Pract* 2017; 23 (1): 274–9.

3. Moser A, Korstjens I. Series: practical guidance to qualitative research. Part 3: sampling, data collection and analysis. *Eur J Gen Pract* 2018; 24 (1): 9–18.

4. Cresswell JW, Poth CN. *Qualitative Inquiry and Research Design*. 4th ed. Thousand Oaks, CA: Sage Publications, Inc.; 2018.

5. Creamer EG. *Advancing Grounded Theory with Mixed Methods*. United Kingdom: Routledge; 2021.

6. Cresswell JWPC, Vicki L. *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: Sage Publications; 2018.

7. Watkins D, Gioia D. *Mixed Methods Research*. New York, NY: Oxford University Press; 2015.

8. Curry L, Nunez-Smith M. *Mixed Methods in Health Sciences Research: A Practical Primer*. SAGE; 2014.

9. Winkelham WJ, Leonard KJ, Rossos PG. Patient-perceived usefulness of online electronic medical records: employing grounded theory in the development of information and communication technologies for use by patients living with chronic illness. *J Am Med Inform Assoc* 2005; 12 (3): 306–14.

10. Ancker JS, Witteman HO, Hafeez B, Provencer T, Van de Graaf M, Wei E. The invisible work of personal health information management among people with multiple chronic conditions: qualitative interview study among patients and providers. *J Med Internet Res* 2015; 17 (6): e137.

11. Ancker JS, Witteman HO, Hafeez B, Provencer T, Van de Graaf M, Wei E. “You get reminded you’re a sick person”: personal data tracking and patients with multiple chronic conditions. *J Med Internet Res* 2015; 17 (8): e202.

12. Willis MA, Hein LI, Hu Z, et al. Feeling better on hemodialysis: user-centered design requirements for promoting patient involvement in the prevention of treatment complications. *J Am Med Inform Assoc* 2021; 28 (8): 1612–31.

13. Ancker JS, Stabile C, Carter J, et al. Informing, reassuring, or alarming? Balancing patient needs in the development of a postsurgical symptom reporting system in cancer. *AMIA Annu Symp Proc* 2018; 2018: 166–74.

14. Wang X, Kim TC, Hegde S, et al. Design and evaluation of an integrated, patient-focused electronic health record display for emergency medicine. *Appl Clin Inform* 2019; 10 (4): 693–706.

15. Benda NC, Alexopoulos GS, Marino P, Sirey JA, Kiosses D, Ancker JS. The age limit does not exist: a pilot usability assessment of a SMS-messaging and smartwatch-based intervention for older adults with depression. *AMIA Annu Symp Proc* 2021; 2020: 213–22.

16. Cresswell KM, Bates DW, Williams R, et al. Evaluation of medium-term consequences of implementing commercial computerized physician order entry and clinical decision support prescribing systems in two ‘early adopter’ hospitals. *J Am Med Inform Assoc* 2014; 21 (2): e194–202.

17. Abraham J, Reddy MC. Challenges to inter-departmental coordination of patient transfers: a workflow perspective. *Int J Med Inform* 2010; 79 (2): 112–22.

18. Novak L, Brooks J, Gadd C, Anders S, Lorenzi N. Mediating the intersections of organizational routines during the introduction of a health IT system. *Eur J Inf Syst* 2012; 21 (5): doi: 10.1057/ejis.2012.2.

19. Greenhalgh T, Wherton J, Shaw S, Papoutsi C, Vijayaraghavan S, Stones R. Infrastructure revisited: an ethnographic case study of how health information infrastructure shapes and constrains technological innovation. *J Med Internet Res* 2019; 21 (12): e16093.

20. Ungerl KM, Schaeferbauer CL, Campbell TR, et al. Integrating community-based participatory research and informatics approaches to improve the engagement and health of underserved populations. *J Am Med Inform Assoc* 2016; 23 (1): 60–73.

21. Iott BE, Veinot TC, Lovelack J, Kaleh E, Golson L, Benton A. Comparative analysis of recruitment strategies in a study of men who have sex with men (MSM) in Metropolitan Detroit. *AIDS Behav* 2018; 22 (7): 2296–311.

22. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *ADM Policy Ment Health* 2015; 42 (5): 533–44.

23. Draucker CB, Martosol DS, Ross R, Rusk TB. Theoretical sampling and category development in grounded theory. *Qual Health Res* 2007; 17 (8): 1137–48.

24. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. *Field Methods* 2006; 18 (1): 59–82.

25. Guest G, Namey E, McKenna K. How many focus groups are enough? Building an evidence base for nonprobability sample sizes. *Field Methods* 2017; 29 (1): 3–22.

26. Korstjens I, Moser A. Series: practical guidance to qualitative research. Part 4: trustworthiness and publishing. *Eur J Gen Pract* 2018; 24 (1): 120–4.

27. Miller WL, Crabtree BF. Qualitative analysis: how to begin making sense. *Fam Pract Res J* 1994; 14 (3): 289–97.

28. Morse J. The significance of saturation. *Qual Health Res* 1995; 5 (2): 147–9.

29. Hennink MM, Kaiser BN, Weber MB. What influences saturation? Estimating sample sizes in focus group research. *Qual Health Res* 2019; 29 (10): 1483–96.
