Moving beyond "until saturation was reached": Critically examining how saturation is used and reported in qualitative research

Jenny Leese, MA[1,2], **Linda C. Li**, PT, PhD[1,2], **Laura Nimmon**, PhD[3,4], **Anne F. Townsend**, PhD[2,5], **Catherine L. Backman**, OT, PhD[2,3]

[1] Department of Physical Therapy, University of British Columbia, Vancouver, Canada

[2] Arthritis Research Canada, Vancouver, Canada

[3] Department of Occupational Science and Occupational Therapy, University of British Columbia, Vancouver, Canada

[4] Centre for Health Education Scholarship, University of British Columbia, Vancouver, Canada

[5] Division of Health Research, Health Innovation One, Lancaster University, Lancaster, UK

Current word count: 1286

Saturation is well known in qualitative research as a key methodological concept and criterion for discontinuing data collection and/or analysis. Often, it is used by qualitative researchers in making decisions related to the adequacy of their sample size (1). It is becoming increasingly common for decisions to take place at the planning stage in order to propose sample size to potential funders or ethical review boards. A concise statement, such as the phrase "until saturation was reached", is frequently included in the reporting of qualitative research to justify the discontinuation of data collection and/or analysis. There is a trend, however, to not offer explanation of what saturation means or how it occurred in the context of a study. These statements are therefore at risk of being vacuous (included perhaps to tick a checkbox that will drive away reviewer and reader criticism) or overly simplistic (perhaps a response to word limit constraints preventing an accurate portrayal of authors' efforts to richly engage in considerations around saturation). Formal quality guidelines (2,3) and sometimes a journal's own quality criteria prefer saturation is discussed and, if authors do not discuss saturation in line with these assessment criteria, reviewers may request revision (4). In this editorial, we argue that uncritical and/or overly simplistic statements about saturation are not sufficient to indicate quality in qualitative research. We contribute to increasing calls to clarify the concept of saturation in research reports by inviting authors and reviewers to critically re-examine and elaborate on how they justify and report saturation in qualitative research (1,4–8).

The notion of saturation has multiple distinct meanings. While its origins lie within grounded theory, at least four different forms of saturation are used (sometimes in isolation or in hybrid combinations) across a broad range of qualitative approaches, with each form having its own theoretical and methodological assumptions that underpin them (9). Rooted within grounded theory, *theoretical saturation* relates to the degree of development of theoretical categories (10,11). This aligns with the underlying epistemological position and general goals of grounded theory, which aim to propose explanatory theoretical frameworks (10,11). The notion of saturation has since evolved in the literature. Saunders et al. identify three variants of saturation including *inductive thematic saturation*, *a priori thematic saturation* and *data saturation* (9). Briefly, *inductive thematic saturation* is similar to theoretical saturation but relates to identifying new codes or new themes, rather than the degree of development of theoretical categories already identified (9). *A priori thematic saturation* takes a largely deductive approach in which data are collected to exemplify pre-determined theoretical categories (9). Lastly, *data saturation* relates to the degree to which new data repeat what was expressed in previous data, with no reference necessary to theoretical categories (9). Other more recent concepts such as *information power* and *theoretical sufficiency* have also been proposed as alternative tools to make decisions about discontinuing sampling and analysis in qualitative research (12,13).

For decades, literature has drawn attention to the concept of saturation as nebulous, a source of confusion even, or a process that lacks systematization, all while it has also come to be described as the gold standard by which purposive sample sizes are determined in health science research (5–7,14). While the notion of theoretical saturation is well-developed with clear guidance on its practical application in grounded theory (10,11), there tends to be more vagueness surrounding other variants of saturation, in relation to whether and how they may apply to other qualitative methodologies (1,4,6). There is no singular formula for determining saturation and, as identified in a review by Walker et al., little consistency in how saturation is *reported* across various qualitative methods, including grounded theory, qualitative description, phenomenology, ethnography, and narrative analysis, (not to imply there should be necessarily) (6,8,14,15). If saturation is deemed relevant by authors, it is thus typically determined through their own interpretation of the term and judgment informed by the literature. It is therefore incumbent upon authors to make thoughtful decisions around determining saturation (when relevant) and also in how it is reported. Of 15 articles that reported reaching saturation identified in a review by Francis et al., none were descriptive or transparent in their reporting about what saturation meant in practice or how it was achieved (1). This is not isolated to the one journal that was the source of Francis et al.'s review. Authors who claim to have achieved saturation do not always explain what this means in the context of their studies, and often mention saturation only sparingly, an issue that has been highlighted by several qualitative scholars (1,7,12,14).

To help address the issue, authors and reviewers can critically consider the relevance of saturation to the research question, methodological orientation and context of each study. Varpio et al. suggest that saturation may be a concept suited to positivist and post-positivist qualitative methodologies such as Glaser and Strauss's Grounded Theory (12). Framed in this way, they emphasize that saturation has helped qualitative research to gain increased legitimacy (12). They also emphasize, however, the unintended consequence that qualitative researchers work within expectations for research rigour that are generally rooted in positivist and post-positivist traditions, even when working within non-positivist orientations (12). It is important that authors do not skim over the concept of saturation or overlook it if it is meaningful to their study. It is also problematic, however, to assume that saturation is a generic requirement or appropriate criterion for establishing quality across all qualitative approaches, given the wide variety of methodologies within qualitative research (4,8).

We posit that it is important for researchers to decide which form of saturation (if any) can meaningfully address the specific threats to quality in the context of a particular study. Moreover, we emphasize it is important authors demonstrate how they have reached these decisions around saturation (if relevant) in their circumstance. The relevance of saturation should be assessed against expectations for research rigour that align with the theoretical and methodological assumptions of an author's particular qualitative approach. Quantity of data (or number of interviews), for example, is not necessarily theoretically important to achieve saturation across all qualitative approaches. What is determined as an appropriate sample size for one qualitative study is also not necessarily an appropriate sample size for another qualitative study (7). In some circumstances, a detailed description of the appropriateness or richness of data derived may be more important in justifying saturation when reporting a study. There are also practical constraints, such as limited resources (e.g., time, money, energy), that could prevent saturation from being reached in well-intentioned studies for which it is deemed relevant (8,11). Prolonged engagement in the field and in-depth analysis is not always feasible. For a study in which saturation is meaningful, we argue that *not* reaching saturation does *not* necessarily invalidate qualitative findings and may simply mean that a phenomenon has not yet been explored fully or sufficiently.

When authors deem saturation relevant and feasible, it is crucial to demonstrate transparency by reporting what saturation means in the context of their study, how it is part of their methodological orientation, how it occurred, and any limitations encountered. While we acknowledge the importance of concise reporting, we argue too that accurate and transparent reporting of qualitative methodologies is essential to ensure quality. Tables, figures, or appendices are alternative ways to provide evidence of how saturation (or other analytical decisions) was addressed when faced with limited word counts. Moving beyond universal applications and generic phrasing in reporting saturation can help to dispel the myth that saturation is the vital criterion for discontinuining data collection and/or analysis. It can enable the full range of qualitative research approaches to justify and explain discontinuing data collection and/or analysis in ways that are meaningful to their own rationales and contexts.

REFERENCES

1. Francis JJ, Johnston M, Robertson C, Glidewell L, Entwistle V, Eccles MP, et al. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. Psychology & health. 2010;25(10):1229–45.

2. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. International journal for quality in health care. 2007;19(6):349–57.

3. Letts L, Wilkins S, Law M, Stewart D, Bosch J, Westmorland M. Guidelines for Critical Review Form: Qualitative Studies (Version 2.0) [Internet]. 2007. Available from: https://srs-mcmaster.ca/wp-content/uploads/2015/05/Guidelines-for-Critical-Review-Form-Qualitative-Studies.pdf

4. O'Reilly M, Parker N. 'Unsatisfactory Saturation': a critical exploration of the notion of saturated sample sizes in qualitative research. Qualitative research : QR. 2013;13(2):190–7.

5. Morse JM. The Significance of Saturation. Qualitative health research. 1995;5(2):147–9.

6. Bowen GA. Naturalistic inquiry and the saturation concept: a research note. Qualitative research : QR. 2008;8(1):137–52.

7. 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.

8. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qualitative research in sport, exercise and health. 2019;(Journal Article):1–16.

9. Saunders B, Saunders B, Sim J, Sim J, Kingstone T, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Quality & quantity. 2018;52(4):1893–907.

10. Charmaz, K. Constructing grounded theory. London: SAGE; 2014.

11. Strauss AL, Corbin JM. Basics of qualitative research: techniques and procedures for developing grounded theory [Internet]. 2nd ed. Thousand Oaks: Sage Publications; 1998.

12. Varpio L, Ajjawi R, Monrouxe LV, O'Brien BC, Rees CE. Shedding the cobra effect: problematising thematic emergence, triangulation, saturation and member checking. Medical education. 2017;51(1):40–50.

13. Malterud K, Siersma V, Guassora A. Sample Size in Qualitative Interview Studies: Guided by Information Power. Qualitative Health Research. 2016;26(13):1753–60.

14. Morse JM. Data Were Saturated . . . Qualitative health research. 2015;25(5):587–8.

15. Walker JL. The use of saturation in qualitative research. Canadian journal of cardiovascular nursing. 2012;22(2):37–46.