A Comprehensive Literature Review of Definitions and Methods for Assessing Saturation of Qualitative Data

Miranda Lauher-Charest[1], Kristi Jackson[1], Cory Saucier[1]

[1]IQVIA

Background

- An early definition of saturation (theoretical saturation) was described by Glaser and Strauss (1967) as part of grounded theory; they defined this as the point where *"no additional data are being found whereby the sociologist can develop properties of the category"* [1]
- This has since evolved into the concept of saturation, broadly defined as the point where additional data are not likely to produce any new relevant themes. This method is used as an indicator for stopping data collection in qualitative research
- The objective of this review was to identify and evaluate existing definitions and methods for assessing saturation of qualitative data

Methods

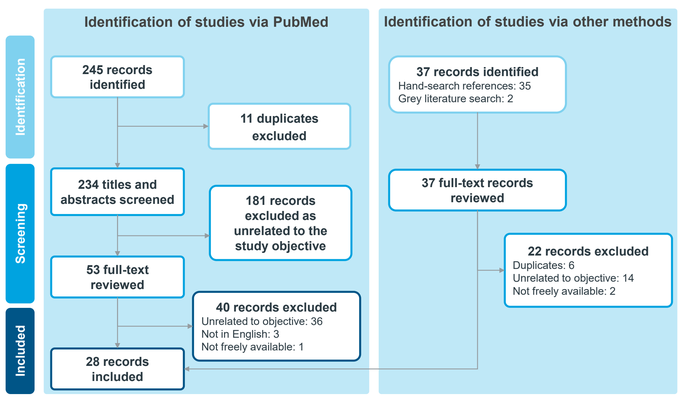
- Two search strings were developed: one targeting the topic of interest (saturation) in qualitative research, and another capturing articles more broadly exploring qualitative research methodology – which may include saturation (Table 1)
- This was supplemented by hand-searching article references and reviewing grey literature, which included select FDA, ISPOR, and ISOQOL guidance

Table 1. Search Strings

Topic	PubMed search string	Hits
Saturation	((“saturation”[Title/Abstract]) AND ((“qualitative” [Title/Abstract]) AND (“qualitative research” [MeSH]) AND (“method*”[MeSH])) AND ((interview [All Fields]) OR (focus group [All Fields])) AND ((“sample” [Title/Abstract]) OR (“sample size”[Title/Abstract])))	15
Qualitative Methodology	((“phenomenology”[Title/Abstract] OR “thematic analysis”[Title/Abstract] OR “content analysis”[Title/Abstract] OR “grounded theory”[Title/Abstract]) OR “saturation”[Title/Abstract] AND (“qualitative”[Title] AND “qualitative research” [MeSH] AND “method*”[MeSH]))	230

- The following selection criteria were used for records identified from all sources: 1) the article was in English, 2) the article was published between 2013 and 2023, and 3) the primary focus of the article was qualitative methodology, including but not necessarily limited to, saturation
- Full texts were retrieved if the record: 1) met selection criteria, 2) presented insufficient information to decide, or 3) was a review article. Articles considered relevant were advanced to full-text review (Figure 1)

Figure 1. Screening process PRISMA diagram



Results

- A total of 28 articles were included in this review. A few articles (Bowen et al. 2008 [2]; Francis et al. 2010[5]; Guest et al. 2006[7]; and Patrick et al. 2011[29]) were reviewed and included from hand-searching references despite falling outside of the 10-year window (2013 – 2023)
- Twenty-five articles defined “saturation,” 5 with original definitions and 20 citing at least 1 source, commonly Glaser and Strauss (1967; n=7)[1], Guest et al. (2006; n=3)[7], and Morse (1995; 2015; n=3)[18]
- Two articles, LaDonna et al. (2021)[14] and Leese et al. (2021)[15], described saturation alternatives: information power (the smaller the required sample size to reach data adequacy) and theoretical sufficiency (data collection can stop once the study team has reached an adequate depth of understanding)

Table 2. Versions of Saturation

Term	Inductive or deductive	Primary focus	Definition	References
Theoretical Saturation	Inductive: a bottom-up approach that sets out to discover concepts within the data without pre-existing theory	Data Novelty	The point in data collection at which gathering more data reveals no new information, or yields any further insights	Bowen (2008)[2] Braun and Clarke (2021)[3] Constantinou et al. (2017)[4] Guest et al. (2006)[7] Hennink et al. (2017)[10] Hennink et al. (2019)[11] Leese et al. (2021)[15] Moser et al. (2018)[19] Nguyen et al. (2023)[20] O’Reilly and Parker (2013)[21] Saunders et al. (2018)[22] Sebele-Mpofu (2020)[23] Van Rijnsoever et al. (2017)[26]
Thematic Saturation		Data Redundance	The point in data collection when no new relevant codes or themes are identified	
Meaning Saturation		None	The point in data collection when researchers fully understand all emergent concepts and themes, and no further dimensions, nuances, or insights can be found	
A priori Thematic Saturation	Deductive: a top-down approach testing whether data exemplify existing theory,	None	The point in data collection when incoming data repeat what was expressed prior, exemplify existing theory, and no new relevant codes or themes are identified	
Code Saturation		Data Novelty	The point in data collection when no additional concepts are identified, and the codebook begins to stabilize	
Data Saturation		Data Redundance	The point in data collection when incoming data repeat what was expressed in previous data and produce little or no new relevant information	

Table 3. Strategies for assessing saturation

Method	Typology	Description	Example
Statistical Modeling	Sample-Focused	Using a mathematical model to estimate the sample size needed to reach saturation, and at what point within that sample size saturation will likely be achieved. All done prior to starting data collection	Tran et al. (2017)[24] used statistical modelling to estimate the sample size needed to reach saturation
Stopping Criterion		Deciding on a pre-determined point (e.g., X number of consecutive interviews where nothing new emerges) at which to stop data collection. Saturation is reached at this point	Guest et al. (2020)[8] developed a mathematical formula to determine an interview stopping point
Comparative Method	Frequency-Focused	Reviewing data in pre-determined batches and documenting new codes (identified by comparing subsequent batches against the first) in a table. Saturation is reached when a batch (or 2) produces few/no new codes	Turner-Bowker et al. (2018)[25] reviewed data in quartile batches and compared concepts from batch one to subsequent batches
Code Frequency Counts		Reviewing each individual interview transcript and counting the number of new codes present in each successive transcript. Saturation is reached when the frequency of new codes diminishes to few or none	Guest et al. (2006)[7] reviewed each individual transcript and counted the number of new codes identified in each
High Order Grouping	Analysis-Focused	Reviewing data and counting the number of new codes present in each successive transcript/batch. Emphasis is placed on the prevalence and salience of codes. This process continues until the most prevalent and most salient codes have been identified and reached saturation	Hagaman et al. (2017)[9] identified the most common themes/sub-themes in the data and determined the number of interviews needed for these to appear 3 times
Code Meaning		Reviewing an initial interview and documenting all codes, then in subsequent interviews, identifying new codes and any aspects, dimensions, or nuances of already documented codes. This process continues until nothing new is identified and all codes reach saturation	Hennink et al. (2017, 2019)[10,11] compiled a list of significant codes and for each interview, reviewed the data gained about each code

- Fifteen articles identified 6 different versions of saturation with different definitions and focuses. Of these, 3 present bottom-up, or inductive, approaches and 3 present top-down, or deductive, approaches. Furthermore, 2 focus primarily on the novelty of incoming data when determining the point of saturation and 2 focus primarily on the redundancy of incoming data when determining the point of saturation. Neither meaning saturation nor *a priori* thematic saturation prioritize one area of focus over the other (Table 2; Figure 2)
- Sixteen articles described strategies for assessing saturation, each of which aligned to 1 of 6 methods – statistical modeling, stopping criterion, comparative method, code frequency counts, high order grouping, and code meaning – across 1 of 3 typologies – sample-focused, analysis-focused, and frequency-focused (Table 3; Figure 3).

Figure 2. Relationships between versions of saturation

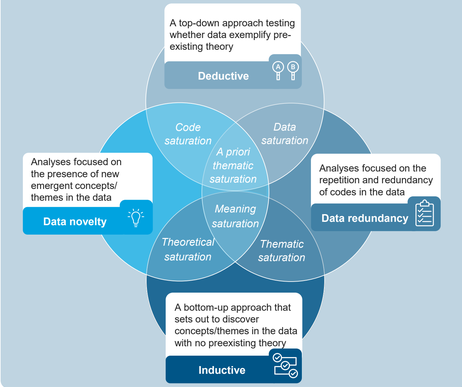
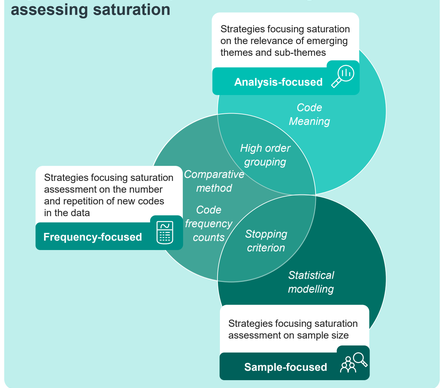


Figure 3. Relationships between strategies for assessing saturation



Conclusion

The concept of saturation is often leveraged as a means of enhancing the strength and credibility of qualitative research; however, it is not always well defined, understood or rigorously assessed. This review offers a central resource for researchers to customize their approach to saturation and transparently communicate that approach to stakeholders (e.g., regulatory, sponsors, researchers). By doing so, saturation can be systematically assessed for fit-for-purpose and compared by stakeholders who rely on this to discuss conceptual coverage, relevance, and salience.

References

[1]Glaser BG, Strauss AL. *The discovery of grounded theory: Strategies for qualitative research*. Routledge; 2017.; [2]Bowen GA. Naturalistic inquiry and the saturation concept: a research note. *Qual Res*. 2008;8(1):137-152. doi:10.1177/1468794107085301; [3]Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qual Res Sport Exerc Health*. 2021;13(2):201-216. doi:10.1080/2159676X.2019.1704846; [4]Constantinou CS, Georgiou M, Perdikogianni M. A comparative method for themes saturation (CoMeTS) in qualitative interviews. *Qual Res*. 2017;17(5):571-588. doi:10.1177/1468794116686650; [5]Francis JJ, Johnston M, Robertson C, et al. What is an adequate sample size? Operationalizing data saturation for theory-based interview studies. *Psychol Health*. 2010;25(10):1229-1245. doi:10.1080/08870440903194015; [6]Fusch P, Ness L. Are we there yet? Data saturation in qualitative research. *Qual Rep*. Published online September 8, 2015. doi:10.46743/2160-3715/2015.2281; [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. doi:10.1177/1525822X05279903; [8]Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. Soundy A, ed. *PLOS ONE*. 2020;15(5):e0232076. doi:10.1371/journal.pone.0232076; [9]Hagaman AK, Wutich A. How many interviews are enough to identify metathemes in multisited and cross-cultural research? Another perspective on Guest, Bunce, and Johnson’s (2006) landmark study. *Field Methods*. 2017;29(1):23-41. doi:10.1177/1525822X16640447; [10]Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: How many interviews are enough? *Qual Health Res*. 2017;27(4):591-608. doi:10.1177/1049732316665344; [11]Hennink MM, Kaiser BN, Weber MB. What influences saturation? Estimating sample sizes in focus group research. *Qual Health Res*. 2019;29(10):1483-1496. doi:10.1177/1049732318821692; [12]Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Soc Sci Med*. 2022;292:114523. doi:10.1016/j.socscimed.2021.114523; [13]Kerr C, Nixon A, Wild D. Assessing and demonstrating data saturation in qualitative inquiry supporting patient-reported outcomes research. *Expert Rev Pharmacoecon Outcomes Res*. 2010;10(3):269-281. doi:10.1586/erp.10.30; [14]LaDonna KA, Artino AR, Balmer DF. Beyond the guise of saturation: Rigor and qualitative interview data. *J Grad Med Educ*. 2021;13(5):607-611. doi:10.4300/JGME-D-21-00752.1; [15]Leese J, Li LC, Nimmon L, Townsend AF, Backman CL. Moving beyond “until saturation was reached”: Critically examining how saturation is used and reported in qualitative research. *Arthritis Care Res*. 2021;73(9):1225-1227. doi:10.1002/acr.24600; [16]Lim CT, Tadmor A, Fujisawa D, et al. Qualitative research in palliative care: Applications to clinical trials work. *J Palliat Med*. 2017;20(8):857-861. doi:10.1089/jpm.2017.0061; [17]Marshall B, Cardon P, Poddar A, Fontenot R. Does sample size matter in qualitative research?: A review of qualitative interviews in is research. *J Comput Inf Syst*. 2013;54(1):11-22. doi:10.1080/08874417.2013.11645667; [18]Morse JM. The significance of saturation. *Qual Health Res*. 1995;5(2):147-149. doi:10.1177/104973239500500201; [19]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. doi:10.1080/13814788.2017.1375091; [20]Nguyen KTNH, Stuart JJ, Shah AH, et al. Novel methods for leveraging large cohort studies for qualitative and mixed-methods research. *Am J Epidemiol*. 2023;192(5):821-829. doi:10.1093/aje/kwad030; [21]O’Reilly M, Parker N. ‘Unsatisfactory saturation’: A critical exploration of the notion of saturated sample sizes in qualitative research. *Qual Res*. 2013;13(2):190-197. doi:10.1177/1468794112446106; [22]Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant*. 2018;52(4):1893-1907. doi:10.1007/s11135-017-0574-8; [23]Sebele-Mpofu FY. Saturation controversy in qualitative research: Complexities and underlying assumptions. A literature review. Serpa S, ed. *Cogent Soc Sci*. 2020;6(1):1838706. doi:10.1080/23311886.2020.1838706; [24]Tran VT, Porcher R, Tran VC, Ravaud P. Predicting data saturation in qualitative surveys with mathematical models from ecological research. *J Clin Epidemiol*. 2017;82:71-78.e2. doi:10.1016/j.jclinepi.2016.10.001; [25]Turner-Bowker DM, Lamoureux RE, Stokes J, et al. Informing a priori sample size estimation in qualitative concept elicitation interview studies for clinical outcome assessment instrument development. *Value Health*. 2018;21(7):839-842. doi:10.1016/j.jval.2017.11.014; [26]Van Rijnsoever FJ. (I Can’t Get No) Saturation: A simulation and guidelines for sample sizes in qualitative research. Derrick GE, ed. *PLOS ONE*. 2017;12(7):e0181689. doi:10.1371/journal.pone.0181689; [27]Vignato J, Inman M, Patsais M, Conley V. Computer-assisted qualitative data analysis software, phenomenology, and Colaizzi’s method. *West J Nurs Res*. 2022;44(12):1117-1123. doi:10.1177/0193945921030335; [28]FDA. Patient-focused drug development: Methods to identify what is important to patients. Published online February 25, 2022.; [29]Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity—Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 1—Eliciting Concepts for a New PRO Instrument. *Value Health*. 2011;14(8):967-977. doi:10.1016/j.jval.2011.06.014