CLINICAL OUTCOMES SOLUTIONS

Innovative science to capture the patient voice

Quantitative Approaches for Estimating Sample Size for Qualitative Research in COA Development and Validation

Helen Doll, MSc DPhil
Strategic Lead, Quantitative Science
Helen.doll@clinoutsolutions.com

CLINICAL OUTCOMES SOLUTIONS

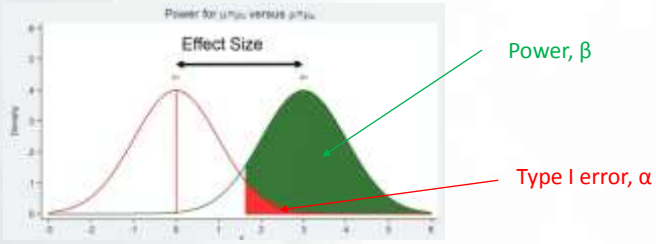
Quantitative Approaches to Sample Size Selection for COA Development

PRESENTATION SUMMARY

- Calculating sample size for **quantitative** studies
- Sample size for **qualitative** studies
  - Random sampling
  - Determining concept saturation quantitatively
- Sample size for estimating meaningful change thresholds (MCTs) for COAs using **Exit Interviews**
  - Robustness and measurement saturation

Calculating Sample Size for Quantitative Studies

- Sample size is calculated using a power analysis.
- A power analysis calculates, for varying sample sizes, a probability (power, β) of finding a statistically significant result (at chosen Type I error, α) for a given population effect size (Cohen, 1988).



- Other factors are important: ethical concerns, practical matters such as participant availability and other resources such as researcher time.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Sample Size for Qualitative Studies

- Generally, the sample sizes used in qualitative research are not justified (Marshall et al, 2013) even though researchers are concerned about using the right sample size (Dworkin, 2012).
- Need to ensure there is *enough*, but not too much, data (>30 too large; Boddy, 2016).
- One review identified that samples of 20 and 30 (and multiples of 10) were most common (Mason, 2010), with 25-30 being a typical recommendation (Dworkin, 2012).

Marshall, B., Cardon, P., Poddar, A. and Fontenot, R. (2013), "Does sample size matter in qualitative research?: a review of qualitative interviews in IS research", Journal of Computer Information Systems, Vol. 54 No. 1, pp. 11-22

Dworkin, S. L. (2012). Sample size policy for qualitative studies using in-depth interviews. Arch Sex Behav (2012) 41:1319–1320

Boddy CR. Sample size for qualitative research. Qualitative Market Research: An International Journal, Vol. 19 Issue: 4, pp.426-432

Mason, M. (2010, August). Sample size and saturation in PhD studies using qualitative interviews. In Forum qualitative Sozialforschung/Forum: qualitative social research (Vol. 11, No. 3).

Calculating Sample Size for Qualitative Studies

- There are no well-established published guidelines to allow formal estimation of sample size *a priori* for qualitative research.
- Qualitative studies generally do not aim to estimate magnitudes and generalise to a larger population, rather to evaluate patterns in a data set (*how* and *why* rather than *what*).
- Is it appropriate to use a quantitative model for a qualitative study?
  - Can we apply the concept of random sampling to qualitative studies?
  - Can we estimate the number of interviews needed to achieve concept saturation?

Random Sampling and Qualitative Research

- A random sample is used in quantitative studies when we want to generalize the results to the wider population.
  - Every person from the population has equal chance of being selected.
- Theoretical and practical reasons why random sampling is not appropriate for qualitative studies:
  - Qualitative samples tend to be small and sampling error would be large.
  - The characteristics under study in the population are often not known.
  - The research characteristics, values, beliefs and attitudes, are unlikely to be normally distributed in the population.
  - Some people may provide more information than others.

Marshall, MN 1996. Sampling for qualitative research. Family Practice 1996; 13(6):522-525.

Concept Saturation

- Concept Saturation is a critical concept in qualitative research.
  - It is assessed based on the qualitative researcher's interpretation, not using any quantitative methodology and this cannot be used to estimate sample sizes *a priori* (Guest et al, 2006).
  - The sample size to achieve saturation depends on how many concepts are present/to be identified, and on the heterogeneity of the population.
- Most qualitative researchers who aim for concept saturation do not rely on random sampling
  - Sampling may be purposive, aiming to select individuals whose responses will provide particularly useful information
- "*Often, researchers invoke the criterion of saturation to justify small samples — very small samples with thin data*". (Charmaz, 2005)

Guest, G., Bunce, A. and Johnson, L. (2006), "How many interviews are enough? An experiment with data saturation and variability" Field Methods, Vol. 18 No. 1, pp. 59-82.

Charmaz, K. "Grounded theory for the 21st century: Applications for advancing social justice studies," In N.K. Denzin & Y.S. Lincoln (Eds.), The Sage Handbook of Qualitative Research 3rd edition, Sage, Thousand Oaks, CA, 2005, 507-535.

Quantitative Determination of Saturation: Example Approaches

A mixed method approach to saturation: applying partial least square regression to qualitative data

ISOQOL 20th Annual Conference
Miami, Florida, USA
October 9-12, 2013

Supporting thinking on sample sizes for thematic analyses: a quantitative tool

International Journal of Social Research Methodology, 2015
Vol. 18, No. 6, 669–684, http://dx.doi.org/10.1080/13645579.2015.1005453

Andrew J.B. Fugard & Henry W.W. Potts

RESEARCH ARTICLE

(I Can't Get No) Saturation: A simulation and guidelines for sample sizes in qualitative research

PLoS ONE 2017; 12(7): e0181689

Frank J. van Rijnsoever[1,2]*

Quantitative Determination of Saturation: Fofana et al, 2013

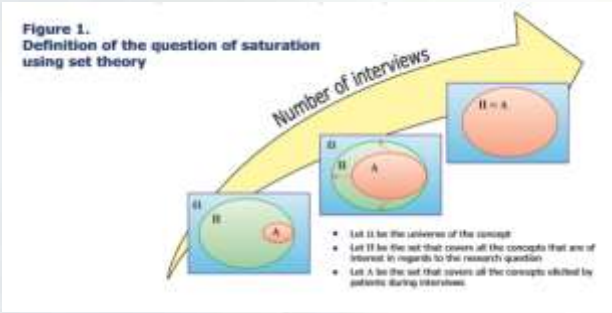
A mixed method approach to saturation: applying partial least square regression to qualitative data

- **Fofana et al (2013)** used **set theory** to determine whether it is likely that, in a series of qualitative interviews, saturation had been achieved.
- Analyses based on **12 interviews** aiming to investigate the impact of *Clostridium Difficile* infection on nurses' work in the hospital; **67 concepts** were spontaneously elicited.
- Used Partial Least Squares (PLS) regression to assess how the number of concepts elicited in the last set of 3 interviews (y variable) compared to the number of concepts elicited during the first set of 9 interviews (x variable).

Fofana F, Bonnaud-Antignac A, Guillemin I, Marrel A, Beriot-Mathiot A, Regnault A. A mixed method approach to saturation: applying partial least square regression to qualitative data. ISOQOL 20th Annual Conference Miami, Florida, USA, October 9-12, 2013

Quantitative Determination of Saturation: Fofana et al, 2013

A mixed method approach to saturation: applying partial least square regression to qualitative data



Using set theory, saturation is achieved when the number of concepts elicited by patients during the interviews (**Λ**) is equal to the number of concepts relevant to the research question (**Π**).

Fofana F, Bonnaud-Antignac A, Guillemin I, Marrel A, Beriot-Mathiot A, Regnault A. A mixed method approach to saturation: applying partial least square regression to qualitative data. ISOQOL 20th Annual Conference Miami, Florida, USA, October 9-12, 2013

Quantitative Determination of Saturation: Fofana et al, 2013

A mixed method approach to saturation: applying partial least square regression to qualitative data

PLS regression results showed that the first 9 nurse interviews were able to explain only 50.9% of the variability of the last 3 nurse interviews (i.e., 49.1% is not explained). This questions the cut-off of acceptability for saturation.

Further research is needed to gain a better understanding of the interpretation of PLS results with regards to saturation.

Fofana F, Bonnaud-Antignac A, Guillemin I, Marrel A, Beriot-Mathiot A, Regnault A. A mixed method approach to saturation: applying partial least square regression to qualitative data. ISOQOL 20th Annual Conference Miami, Florida, USA, October 9-12, 2013

Quantitative Determination of Saturation: Fugard and Potts, 2015

Supporting thinking on sample sizes for thematic analyses: a quantitative tool

- **Fugard and Potts (2015)** quantified the process of saturation by fitting a model (assuming random sampling) in terms of:
  a) the **expected population theme prevalence** of the **least prevalent theme**, derived either from prior knowledge or based on the prevalence of the rarest themes considered worth uncovering, e.g. 1 in 10, 1 in 100;
  b) the **number of desired instances** of the theme (the negative binomial probability distribution used to quantify the probability of identification); and
  c) the **power of the study**. An adequately powered study will have a high likelihood of finding sufficient themes of the desired prevalence. “This calculation can then be used alongside other considerations”.

Fugard AJB and Potts HWW. Supporting thinking on sample sizes for thematic analyses: a quantitative tool. Int J Soc Res Methodol 2015; 18(6):669-684.

Quantitative Determination of Saturation: Fugard and Potts, 2015

Table 1. Sample size required as a function of the population theme prevalence and how many instances of the theme are desired, with 80% chance of observing the desired number of instances (power).

Population theme prevalence (%)	Desired number of theme instances							
	1	2	3	4	5	10	20	30
5	32	59	85	110	134	249	471	687
10	16	29	42	54	66	124	234	343
15	10	19	28	36	44	82	156	228
20	8	14	21	27	33	61	116	170
25	6	11	16	21	26	49	93	136
30	5	9	14	18	21	40	77	113
35	4	8	12	15	18	34	66	96
40	4	7	10	13	16	30	57	84
45	3	6	9	11	14	26	50	74
50	3	5	8	10	12	24	45	66
55	3	5	7	9	11	21	41	60
60	2	4	6	8	10	19	37	55
65	2	4	6	7	9	18	34	50
70	2	4	5	7	8	16	31	46
75	2	3	5	6	8	15	29	43
80	1	3	4	6	7	14	27	40
85	1	3	4	5	7	13	25	37
90	1	2	4	5	6	12	23	35
95	1	2	3	4	6	11	22	33

Fugard AJB and Potts HWW. Supporting thinking on sample sizes for thematic analyses: a quantitative tool. Int J Soc Res Methodol 2015; 18(6):669-684.

Quantitative Determination of Saturation: Rijnsoever, 2017

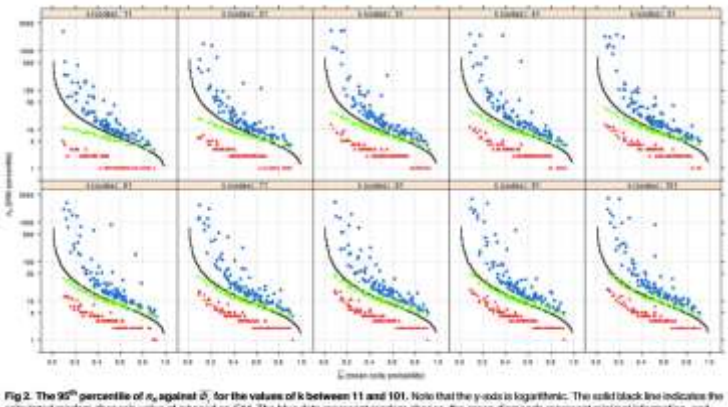
(I Can't Get No) Saturation: A simulation and guidelines for sample sizes in qualitative research

- Explored the sample size in qualitative research required to reach theoretical saturation (building on work by Tran et al, 2017)
  - A population consists of sub-populations containing different types of information sources holding a number of codes
  - Theoretical saturation reached when all codes have been observed once in the sample
  - Three different scenarios: random sampling, minimal information (≥1 new code per sampling step), maximum information (largest number of new codes per sampling step)
    - Used simulations, systematically varying hypothetical populations, to assess minimum sample size for each scenario

Tran V-T, Porcher R, Tran V-C, Ravaud P. Predicting data saturation in qualitative surveys with mathematical models from ecological research. Journal of clinical epidemiology. Elsevier; 2017; 82: 71–78.

Quantitative Determination of Saturation: Rijnsoever, 2017

Sample size, n, to achieve saturation (with 95% confidence) vs probability of observing codes in the interviews, for differing number of codes in the population



random sampling
minimal information
maximal information

Fig 2. The 95th percentile of n_c against $\bar{p}$, for the values of k between 11 and 101. Note that the y-axis is logarithmic. The solid black line indicates the calculated random chance's value of n based on F71. The blue dots represent random chance, the green diamonds represent minimal information, and the red triangles represent maximal information.

Finding: Theoretical saturation is more dependent on the mean probability of observing codes in the interviews than on the number of codes in a population

Sample size for estimating meaningful change thresholds (MCTs) for COAs using Exit Interviews

Exit Interviews: Estimating sample size

- Objective: To derive meaningful change thresholds (MCTs) in clinical trial participants
- Exit interviews to be conducted in the study participants after they have completed the study.
  - Conduct interviews to explore participant perceptions on any changes experienced during the study and whether or not any changes are meaningful.
  - Qualitative findings will be triangulated with data obtained from other sources (e.g. quantitative data, clinical review and interpretation).
  - 300 estimated study participants. How many should participate in the exit interviews?
- **Conduct a power calculation to define the (minimum) number of participants**

Exit Interviews: 1) Difference in proportions

- Base the sample size calculation on a **test of proportions**.
  - are the study arms equivalent in terms of the proportions achieving meaningful change?
- For example, the equivalence of proportion means is tested with:
  - 50%/55% of patients on treatment observing any level of meaningful improvement (at least minimal or higher) compared with 20%/25% of patients on placebo
  - significance (alpha) set at e.g. 0.05 and power of e.g. 80%
  - These calculations would lead to a sample size requirement of approximately **80 treatment completers (40 in each study arm)**.

Exit Interviews: 2) Expected size of MCT

- Base the sample size calculation on **the expected size of the MCT** from previous research, as well as distributional parameters such as measure standard deviation (SD).
  - COA measure scores for patients who experience improvement (on a PGI, for example) evaluated using a paired sample t-test.
  - The equivalence of paired means tested with significance (alpha) set at 0.05, power 0.80, a specific correlation between assessments (e.g. 0.40), specific population SD with 95% CIs, and MCTs estimated from previous research.
  - A **minimum** sample size of participants is estimated (e.g. N=60-70).
  - MCT estimated using these participants.

Exit Interviews: 3) Robustness and Measurement Saturation

- The **robustness** of this estimate is evaluated in terms of the sample SD and stability of the MCT with increasing sample size.
  - If the sample SD > population value then a further 5% of patients in each treatment arm will be interviewed and the MCT recalculated.
  - A robust estimate is defined as an MCT which does not change with 2 successive sample size increases with instead the 95% CI narrowing (**measurement saturation**). A moderate effect size of 0.50 may also be used as a criterion.
- To ensure the **generalizability** of exit interview results, participants can be sampled based on a hierarchical cluster design. Included patients sampled to be proportionate to the clinical trial sample distribution with respect to demographic and clinical characteristics.

Sample size estimation in qualitative research: Conclusions

1) Specific approaches can be used to estimate sample size in qualitative research, e.g. to assess concept saturation.
   - These need to be considered alongside other issues, and may also only be able to be applied once data have been collected.
2) Sample size calculation for small samples, e.g. for exit interviews, is facilitated with the use of standard quantitative techniques, such as having the power to identify a difference in the proportion of subjects reaching the MCT between groups, or having the power to identify a specific MCT value.
   - The robustness of the estimate of MCT can be assessed through measurement saturation, an extension of concept saturation.
   - Generalizability issues can also be incorporated into sampling.
3) The Clinical Outcomes Solutions team, led by Stacie Hudgens, has applied similar power calculations in other types of small sample work.

How many qualitative interviews is enough?

"That is, of course, a perennial question if not a great one. The answer, as with all things qualitative, is "**it depends**." It depends on your **resources**, how **important** the question is to the research, and even to how many respondents are enough to satisfy committee members for a dissertation. For many qualitative studies one respondent is all you need – your person of interest. But in general the old 4 rule seems to hold that you **keep asking as long as you are getting different answers**, and that is a reminder that with our little samples we can't establish frequencies but we should be able to find the RANGE of responses. Whatever the way the question is handled, the best answer is to **report fully how it was resolved**." Harry Wolcott

Baker, S. E., & Edwards, R. (2012). How many qualitative interviews is enough? Expert voices and early career reflections on sampling and cases in qualitative research. Retrieved from http://eprints.ncrm.ac.uk/2273/