



A practical guide to data analysis in general literature reviews

Rebecca Popenoe, Ann Langius-Eklöf, Ewa Stenwall and
Anna Jervaeus

Abstract

Academic theses at the bachelor's and master's levels in nursing, physiotherapy, occupational therapy, public health and related fields often take the form of general reviews of published research on a relevant clinical issue. While there are many guides to doing literature searches and evaluating article quality, there are to our knowledge no published detailed guides of how to do the actual data analysis in such general literature reviews, particularly that are applicable to students. This article seeks to describe a systematic method of data analysis appropriate for undergraduate research theses, where the data consists of the results from available published research. We present a step-by-step guide with authentic examples and practical tips.

Keywords

academic theses, data analysis, general literature reviews, healthcare education, practical guide

Accepted: 13 January 2021

Introduction

This article is a practical guide to conducting data analysis in general literature reviews. The general literature review is a synthesis and analysis of published research on a relevant clinical issue, and is a common format for academic theses at the bachelor's and master's levels in nursing, physiotherapy, occupational therapy, public health and other related fields. Examples of research questions are: What barriers to successful self-care do patients with heart failure describe? What are the effects of physical activity on individuals' experienced stress? How can mindfulness be used in occupational therapy? What strategies work best to prevent postoperative shivering in patients? How do patients experience isolation in infectious disease wards? What strategies are effective for inserting peripheral venous catheters in patients with difficult intravenous access?

The central task for the author of any general literature review is to analyse the results of multiple scientific studies in order to describe the state of knowledge about a particular topic, in order to draw conclusions with clinical applications. This article seeks to describe a systematic method of data analysis appropriate for undergraduate theses, where the data consists of the results from available published research. Although the methods described here are usable by any researchers conducting a general literature review, the article is addressed to students (and their supervisors) in the health sciences.

A general literature review starts with formulating a research question, defining the population, and conducting

a systematic search in scientific databases, steps that are well-described elsewhere.^{1,2,3} Once students feel confident that they have thoroughly combed through relevant databases and found the most relevant research on the topic, however, what is arguably the hardest part of the process remains.

How should one go about synthesizing, analysing and presenting the results of different studies on a topic? This is an especially challenging task when these studies may not themselves ask the exact same research question you the student are asking, even if they contain findings that answer your research question. That the studies may use vastly varying methodologies, from interview studies to randomized trials, also poses a challenge to the task of synthesising and analysing the data.

We have discovered that students lack a structured guide to the actual analysis that a general literature review requires. Below we present a step-by-step guide for analysing data for two different types of research questions. The data analysis methods described here are based on basic content analysis as described by Elo and Kyngäs⁴ and Graneheim and Lundman,⁵ and the integrative review as described by Whittemore and Knafl,⁶ but modified to

Division of Nursing, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

Corresponding author:

Anna Jervaeus, Alfred Nobels allé 23, 141 83 Huddinge, Sweden.
Email: anna.jervaeus@ki.se

be applicable to analysing the results of published studies, rather than empirical data. The methods described here are inductive, that is, they do not describe how to use a pre-existing model or theory to analyse data, but instead describe how to find patterns and answers in the data without preconceived theories or frameworks for analysis.

First step: Making an article matrix

Once you have completed searches of the databases and identified all studies that answer your research question and fit within the inclusion and exclusion criteria you have chosen (for example only studies on people aged over 65 years, or no studies of people with disabilities), it is time to get started on the data analysis. The first step is to get an overview of all identified studies by making an article matrix where you describe each study's aim, methods and results, especially those results that are relevant to your particular research question. Table 1 shows examples of how this might look for an experimental, randomized controlled trial study; for an interview study; and for a study based on a questionnaire. All of these studies will be used in examples later in this article; the first article for a study about inserting peripheral vein catheters (PVCs), and the second two for a study about patient experiences of isolation wards.

The next steps in the data analysis process differ depending upon the type of research question you have and the designs of the studies you will be analysing.

When your research question (aim) is about effect, risk, association or prevalence

Research questions about whether or not a certain treatment, intervention or approach is effective are answered primarily by experimental studies, including randomized controlled trials (RCT) and quasi-experimental studies. Research questions about risk factors, associations and prevalence are answered primarily by cross-sectional and longitudinal studies. The data analysis process for these kinds of quantitative studies involves three steps:

1. Identifying data that answers your research question (aim), in this case largely numerical data that must be 'dug' out of each study's results.
2. Organizing the data in a thematic way.
3. Synthesizing, analysing, and presenting the data.

This section will now go through each of these steps, using the example of the following research question: What strategies are effective for inserting peripheral venous catheters in patients with difficult intravenous access? In the development of this research question you have hopefully already determined whether you only want to examine strategies that lead to success in insertion of a PVC, or whether you also want to include findings about patient experiences of pain during the procedure, how the skill of the provider affects insertion success, or other aspects of what 'success' means for a PVC insertion.

Identifying results that answer your research question

With your research question and its exact terms in mind, you begin the process of data analysis by clearly identifying the actual results of each study that answer your research question. This entails separating the wheat from the chaff; that is, picking out the study results that are relevant to your research question, and ignoring results that are not.

Practical tip: Tape your research question to your computer so that you don't lose sight of it. You can also write it on the top margin of your draft during the writing period so that it is visible on every page (just don't forget to remove it at the end).

Be aware that it is not always easy to identify the actual results of any study that answer your particular research question. For example, if you are interested in effective strategies for inserting PVCs in adult patients with difficult vascular access then you need to ignore study results that a) concern children and b) concern other kinds of vascular access or catheterization.

Box 1 shows an example of text from the results section of an article with the information that answers your research question about effective strategies for inserting PVCs in people with difficult intravenous access. The parts of the text that answer your research question are underlined.

Note that figuring out which data is relevant can be difficult. The relevant data will be presented in the Results section, may be mentioned in the Abstract, and is often summarized at the very beginning of the Discussion section of an article. (If you find results conveniently summarized at the beginning of a Discussion section, just make sure to avoid including the actual 'discussion' of these results by the authors, where they compare their results to those of other studies or give their own interpretations and opinions about their results.)

The Abstract seldom describes findings in enough detail; sometimes the very results that are key to answering your particular research question are not even mentioned in the Abstract. Sometimes the findings are explained clearly in the text in the Results section of a study, but sometimes one may need to look at tables, graphs and figures to find the relevant information.

Quantitative studies tend to contain a lot of statistical data, and not all of it is relevant in a general literature review, especially at the undergraduate level. Tables that show the baseline characteristics of the population studied, for example, are important for providing the study with validity (roughly, accuracy and trustworthiness), but may not be relevant to your synthesis of the study's findings.

In general, authors focus on results that are statistically significant. But sometimes they report results where a difference or an effect was found, but the *p*-value is far above 0.05. For example, in Table 2 you can see that Bahl et al.⁷ report that the time it took to insert a PVC was on average 20.7 minutes using ultrasound compared to 15.8 minutes

Table 1. Example of an article matrix summarizing studies for use in a general literature review.

Research question: (Write your research question at the top of the matrix; the articles used in this example answer different research questions.)				
Art. nr	Reference	Aim	Method <i>Describe concisely & use the same format for each article</i>	Results <i>Summarize the study results with a focus on the results of relevance to YOUR literature study</i>
1	Bahl, A., Pandurangadu, A. V., Tucker, J., & Bagan, M. (2016). A randomized controlled trial assessing the use of ultrasound for nurse-performed IV placement in difficult access ED patients. <i>American journal of emergency medicine</i> 34, 1950-1954. ⁷	To analyse outcomes associated with nurse-performed ultrasound (US)-guided intravenous (IV) placement compared to standard of care (SOC) palpation IV technique on poor vascular access patients.	RCT 122 participants 63 randomized to ultrasound 59 to standard of care Trauma center Nurses were trained to use US-guided IV placement in the study's first phase. Successful IV placement and time to placement measured.	Success rate for placement: Ultrasound: 76% (48/63 patients) Standard of care: 56 % (33/59 patients) ($p=.02$) Mean time to placement: Ultrasound: 20.7 min Standard (palpation): 15.8 min ($p=.75$)
2	Mayan, M., Robinson, T., Gokiart R., Tremblay, S., Abonyi; Long (2017). Respiratory isolation for tuberculosis: the experience of indigenous peoples on the Canadian prairies. <i>Public Health Action</i> 7(4), 275–28. ⁸	To understand how indigenous peoples with tuberculosis in Canada experience respiratory isolation.	Semi-structured interviews 48 adults, 17-62 years, men and women Qualitative content analysis	Participants described • Feeling trapped, anxious over being confined, like being in jail • Feeling bored • Missing their families and feeling sad and disconnected • Difficulty being away from their children • Loneliness
3	Guilley-Lerondeau, Bouriguault, C., Guille des Buttes, A-C., Birgand, G, Lepelletier, D. (2017). Adverse effects of isolation: a prospective matched cohort study including 90 direct interviews of hospitalized patients in a French University Hospital. <i>European Journal of Clinical Microbiology and Infectious Disease</i> 36, 75–80. ⁹	To evaluate the satisfaction and psychological impact (anxiety) of isolation precautions in hospitalized patients in a French acute-care university hospital	Cohort study Interviews: 30 patients with isolation precautions 60 patients without isolation precautions Spielberger anxiety scale (20-80 point range)	83% of the 30 isolated patients understood that they were isolated 73% of the 30 isolated patients understood the collective interest behind third isolation 67% of the 30 isolated were not satisfied with the information they had received, compared to 2% of the 60 patients without isolation 52 median score on Spielberger anxiety scale in patients in isolation 31 median score for patients not in isolation

Note that there are more details in the results of each article that may be relevant to your results; the article matrix is merely to give an overall summary.

when only palpation and a tourniquet were used. But then the authors note that the p -value for this finding was 0.75; in other words, the probability that this difference was due merely to chance is 75%. So the difference in the time it took for the two different procedures was very probably due to chance, and not indicative of the ultrasound method being slower. Why would they then write this finding? It may be that they believed that if the study had had many more participants the time difference would have reached statistical significance. In any case, you cannot write in your review that ‘using ultrasound was slower than using traditional methods for inserting a PVC’,

unless you add that ‘but this difference did not reach statistical significance’.

Answers to your research question may also be expressed in risk ratios, odds ratios, effect sizes, confidence intervals, or a number of other statistical measures. If you need help understanding and interpreting these findings you can consult the literature, search for answers online, or discuss them with your supervisor.

Once you have gone through each article and highlighted or underlined all sentences and all data that seems to answer your research question (Box 1) it is time to organize and evaluate this information.

Box 1. Identifying data that answers your specific research question.

Research question: What strategies are effective for inserting peripheral venous catheters in patients with difficult intravenous access?

An article that speaks to the research question:

Bahl, A., Pandurangadu, A.V., Tucker, J. & Bagan, M. (2016). A randomized controlled trial assessing the use of ultrasound for nurse-performed IV placement in difficult access ED patients. *American Journal of Emergency Medicine* 34: 1950–1954.

Passage from the Results section of the article, with sections underlined or highlighted that are relevant to the research aim:

'A total of 124 patients consented and were enrolled into the study; 63 subjects were randomized into the US [ultrasound] guidance arm, and 61 subjects were initially randomized into the SOC (standard of care) arm, but 2 patients were excluded due to the lack of availability of a study-trained SOC nurse, leaving 59 patients in the study arm. Fig. 1 demonstrates a summary of the results of the 2 study arms and how a functional IV was obtained. Table 2 illustrates the patient demographics and shows that there was no statistical difference in the patient population with regard to age, sex, obesity, IV drug use, ESRD, sickle cell disease, heart rate, or mean arterial pressure.

The success rate for IV placement was 76% (48/63) for the US-guided arm and 56% (33/59) in the SOC arm ($P = .02$). The odds ratio for success for the US guidance arm was 2.52 (95% CI, 1.09–5.92) times the odds of success for the SOC arm.'

The results that directly answer your research question are underlined.

Table 2. Example of data analysis table for use in synthesis of data from randomized controlled trials (RCTs) and other quantitative studies.

Research question:

What nursing strategies are effective for inserting peripheral venous catheters in patients with difficult intravenous access?

Strategy	Results	Relevant details, strengths and limitations of each study
Ultrasound	Bahl et al. (2016)⁷ Success rate for placement: Ultrasound: 76% (48/63 patients) Standard of care: 56% (33/59 patients) ($p = 0.02$) Mean time to placement: Ultrasound: 20.7 min Standard (palpation): 15.8 min ($p = 0.75$) ← <i>Note that this p-value means the difference is NOT significant!</i>	RCT 122 participants Trauma centre Strengths: - RCT - Nurses had uniform level of training in ultrasound (US) (removal of potential confounder) - Narrow carefully applied inclusion criteria for patients with difficult access Weaknesses: - Time to placement of PVC in control group (standard care) is probably underestimated due to study factors
	McCarthy et al. (2016)¹⁰ Success rate for placement, first attempt, patients with difficult access: Ultrasound 81.6 % ($n = 98$) 'Landmark' (palpation w tourniquet): 35.1% ($n = 94$) (Risk ratio 2.32 (1.74–3.11)) Success rate for placement, first attempt, Patients with moderately difficult access: Ultrasound 81.2 % ($n = 202$) 'Landmark' 71.4% ($n = 199$) (Risk ratio 1.14 (1.02–1.27)) More skilled providers were more successful than less skilled providers at IV placement w/US in patients w/ difficult or moderately difficult access	RCT 1189 participants Emergency room (ER) Strengths: - RCT - Sample is large and representative (note that 37% were obese) - Intervention & control groups were similar at baseline Weaknesses: - Healthcare providers of mixed skill levels - Consecutive sampling not fully achieved

(continued)

Table 2. Continued.

Research question: What nursing strategies are effective for inserting peripheral venous catheters in patients with difficult intravenous access?		
Strategy	Results	Relevant details, strengths and limitations of each study
	İsmailoğlu et al. (2015) ¹¹ Success rate for placement: Ultrasound: 70% (21/30 patients) Standard of care: 30% (9/30 patients) ($p = 0.002$) Differences in success rate for first attempt, however, did not reach statistical significance (6 in the treatment group, or 20%, and 3 in the control group, or 10%; $p = 0.278$). Average number of attempts: 2.07 (US), 2.10 (Control) Experienced pain by patient: Mean 4.77/10 in US group Mean 6.0/10 in control group ($p = 0.013$)	RCT 60 participants ER Strengths: - RCT - Consecutive sampling - Intervention & control groups similar at baseline - Nurses trained in US Weaknesses: - Small sample - Few patients in sample had obesity or cancer, common patient groups with difficult venous access - Used inexperienced providers
	Sou et al. (2017) ¹² Median unsuccessful attempts before referral: 2 (IQR 2.4) Median number of attempts with ultrasound: 1 (IQR 1.1) ($p < 0.001$) First-time success rate 93% (348/379 patients) Success for all patients after max 2 attempts Lower pain scores with US-guided insertion: Median 2/10 pain score for attempts with US Median 7/10 pain score for attempts without US ($p < 0.001$)	Cohort study 379 patients Clinic for US-guided insertion to which patients with difficult access were referred from ER Strengths: - Trained healthcare providers - Relatively large number of participants Weaknesses: - Not an RCT - Sampling weaknesses
Heat	Biyik-Bayram and Caliskan (2016) ¹³ Mean rate of successful placement on first attempt in intervention group (heat): 90% Mean rate of successful placement on first attempt in control group: 62.5% ($p = 0.004$) Mean catheterisation time, intervention: 90 seconds Mean catheterisation time, control: 147 seconds ($p = 0.001$)	Experimental study with control group 80 chemotherapy patients Hospital oncology ward Strengths: - Experimental study - Power calculation used to estimate needed number of participants Weaknesses: - Non-randomized study - Risk of bias in patient selection - Patient population limited to those with cancer diagnosis

Note: PVC: peripheral venous catheter.

Organizing quantitative data in a thematic way

Some types of quantitative data lend themselves easily to one kind of organization. For example, in our review of research on PVC insertion, it makes sense to divide studies according to the strategy they focused on: 1) ultrasound, 2) heat, or 3) alternative therapeutic methods (Table 2). In other kinds of studies the organizing principle may be less obvious. If one is measuring the effect of nurse-led interventions to reduce smoking, for example, one could organize the data by the type of intervention (individual counselling vs. group counselling), by the type of study (RCTs vs. cohort studies), by how long the follow-up period was (studies with longer follow-up are, as a

rule, more clinically meaningful), or by another relevant factor.

The other type of information that should now be taken into account and recorded in the data analysis table (Table 2) is information relating to the strengths and weaknesses of each study, which you will need for the next step in the analysis process. A study's strengths and weaknesses might relate to the number of participants in the study, how randomization was carried out, or confounders (factors 'outside' the study that might have affected the results, such as if poor quality catheters were used in a study, potentially explaining the study's poor results). A study's strengths and weaknesses are usually

described in a study's Discussion section, but you can also look for aspects of a study that the authors themselves do not specifically mention, but that you notice may have had an effect on their results. A weakness could be, for example, that a high number of participants had multiple health problems affecting the results, or that a company that sold ultrasound equipment financed the study, or other aspects that might detract from (or add to) the validity and reliability (roughly, accuracy of measurement) of the results.

Table 2 shows an example of how data from our PVC study could be organized into a table to create a good overview of the studies' results as well as the studies' strengths and weaknesses.

Synthesising, analysing and presenting quantitative data

The challenge of synthesising and analysing quantitative data is that studies of the same phenomenon tend to study the phenomenon in different ways and to measure their results in different ways. For example, to measure outcomes, two studies measured the percentage of successful attempts,^{7,11} whereas another study divided patients into groups with different degrees of difficulty of access.¹⁰ Yet another study, a cohort study, lacking a control group, found merely that 93% of patients, deemed to have difficult access, were able to have an IV line inserted when ultrasound was used.¹² We do not know what the success rate would have been for a matched control group where ultrasound was not used; only that previous studies have rarely found such a high success-of-insertion rate for patients with difficult intravenous access when ultrasound is not used.

In cases where studies do have similar designs and ways of measuring outcomes, the results can easily be summarized. For example, 'In three RCT studies insertion of an IV line using ultrasound was found to be successful in between 64% and 81% of patients with difficult venous access, compared to only 39% to 45% of similar patients in a control group where palpation with or without a tourniquet was used'. But the more common situation is that you have to summarize disparate types of results. In these cases, you can summarize the overarching results in your own words, and then describe the results of each study one by one. For example, to summarize what the research shows about the use of ultrasound to insert IV catheters in people with difficult intravenous access, the summary could be as shown in Box 2.

The same process is followed for each of the studied procedures: heat, alternative therapies, etc. That is, the overall results of the studies of heat's effects on easing catheter insertion, for example, are presented, and then the specifics of each study that examined heat's effects are presented. Depending on the wording of your research aim, you might want to include other types of results, such as how much pain patients felt when different procedures were used, or whether the use of ultrasound was more successful with experienced and trained providers, which some studies examined.

It is also important to note each study's validity, that is, the validity of the study's results. Not all results are equal, because some studies are better than others. For example, a study with a large sample, careful randomization, and few confounding factors, will produce a more trustworthy result than a study with a smaller sample, selection bias, no randomization, and many confounding factors. The validity and reliability of particular studies can be woven into your synthesis of results, as in the example in Box 2, or discussed separately at the end of your results section.

When your research question (aim) is about experiences, attitudes or perceptions

Research questions about experiences, attitudes and perceptions can be answered by primarily two types of studies: 1) qualitative studies (based on interviews, focus-group discussions, or observations) and 2) questionnaire-based studies. Occasionally other types of quantitative studies use qualitative methods to assess participants' experiences and results from such studies can also be used when answering research questions about experiences, attitudes and perceptions. For example, for a general literature review on patient experiences of diabetes support groups a student might find an RCT study that investigated whether group support sessions in diabetes self-care were more motivating for patients than individual appointments with a healthcare provider. This study used a questionnaire to gauge participants' experiences, and the results regarding experiences of group support sessions could be relevant to the literature review.

As with data from experimental, longitudinal and cross-sectional studies, the data analysis process in studies about subjective experiences involves three steps:

Box 2. Example of how to synthesize results from quantitative studies.

'Four studies, of which three were randomized controlled trials, have found that the use of ultrasound is more effective than palpation alone or palpation with a tourniquet when inserting an IV catheter in patients with difficult intravenous access. In the largest and most rigorous study (McCarthy et al., 2016), 81.6% of patients with difficult access were successfully given IV catheters using ultrasound, compared to 35.1% of patients with difficult access for whom only palpation and tourniquet were used. The two smaller RCT studies, with 122 and 60 patients respectively, also showed significantly greater success for insertion of IV catheters in patients with difficult access when ultrasound was used: 76% vs. 56% (Bahl et al., 2016) and 70% vs. 30% (İsmailoğlu et al., 2015).'

Box 3. Identifying data that answers your specific research question.

Research question: How do adult patients experience living in isolation on an infectious disease ward?

Article: Mayan, M., Robinson, T., Gokiart, R., Tremblay, M., Abonyi, S and Long, R. (2017). Respiratory isolation for tuberculosis: the experience of indigenous peoples on the Canadian prairies. *Public Health Action* 7(4): 275–228.

Passage from article, with sections underlined or highlighted that are relevant to the research aim (p. 278):

'Upon entering hospital isolation, participants described experiencing a great deal of sadness and disconnect over being separated from their family members. They acknowledged that their friends and family were often fearful of visiting because of the diagnosis of TB and the isolation setting. One participant described how her sister, who was the only family member in her life, drove her to the hospital, but had not returned, as she was "too scared". Being isolated in hospital caused many participants to feel as though people no longer wished to be around them, "like nobody don't want you no more. You got a sickness and you can stay there until that sickness goes away. If it doesn't go away you're by yourself there".

Participants who were mothers noted that it was especially difficult to be separated from their children. Most often, children, who were being pre-emptively treated for TB infections, lived far away, and/or were too young to visit their parent in isolation. A mother described leaving her husband and eight children behind: "They don't listen to their dad as much, that's why when I'm here my husband phones me and tells me to tell them to do this, do that, because they don't listen to him and they listen to me more".

Note that the other parts of this passage, however interesting and peripherally relevant, do not answer the specific research question. They may be relevant as context to other statements, but they do not answer the research question itself directly.

1. Identifying the findings in each study that answer *your* research question (aim).
2. Condensing and summarizing these findings.
3. Grouping the findings thematically into categories that speak to your research question.

This section will now go through each of these steps, using the example of the following research question: How do adult patients experience living in isolation in an infectious disease ward?

Identifying results that answer your research question

The first task, again, is to identify what each study actually found. This sounds easy but it is not always.

Practical tip: Tape your research question to your computer so that you don't lose sight of it. You can also write it on the top margin of your draft during the writing period so that it is visible on every page (just don't forget to remove it at the end).

As you search through the Method and Results sections of your articles, keep in mind your inclusion and exclusion criteria. If you have chosen to write about adults, for example, but an article discusses experiences of both children and adults, you will need to ignore the results that only relate to children. Or, if your research question is about patient experiences of living in isolation in an infectious disease ward, you need to be careful not to include descriptions of other family members' experiences, or their experiences of the disease itself, or their experiences of being released from isolation. As you search for answers to your research question you need to be strict with

yourself and remember that, however interesting the patients' descriptions of the disease itself are, or how it felt to be released from isolation, that is not what you are interested in now.

Another factor to watch out for when you identify data that answers your research question is your own biases and preconceived ideas of what you may find. You may expect, for example, that people find living in isolation to be very difficult, but if a participant in one study states that they appreciated the break from the daily grind, you have to include that as well! Being clear with yourself about your own assumptions and subjectivity before embarking on the analysis will help you to maintain your objectivity.

In practical terms, a good way to identify study results that answer your research question is to use a pencil or highlighter pen to mark all relevant findings. Again, keeping your specific research question as well as your inclusion and exclusion criteria squarely in front of you, go through each article's results carefully, marking the phrases, sentences and paragraphs that answer your research question (see Box 3).

Practical tip: Go through each article in your literature review methodically with a pencil or highlighter pen and underline or highlight the paragraphs, sentences, parts of sentences or survey results that answer your research question specifically.

The specific findings of each study that answer your research question can be called 'data units'. (In qualitative studies these are often called 'meaning units'.) Note that several types of data or information may be included

Table 3. Example of data analysis table.

Research question: How do patients experience isolation in infectious disease wards?

1	2	3	4
Data units (and the study they came from, listed in a separate article matrix)	Summary or condensed version of column 1	Codes = the essence of the meaning in each data unit (make sure they are relevant to the study's research question)	Subcategories = groupings of codes according to common denominators (should be in the form of an answer to the study's research question)
Upon entering hospital isolation, participants described experiencing a great deal of sadness and disconnect over being separated from their family members. (2)	Sadness and feeling of disconnect over separation from family	Missed family	Patients longed for their families
One woman commented that, 'If my husband and children had been able to stay in the room with me the whole time would have been a lot easier'. (2)	'If I had family in the room it would have been easier'.	Felt family would have made it easier	
'I felt like a pariah. I felt that my whole body was dirty and somehow shameful'. (4)	Felt dirty and ashamed and like a pariah	Felt dirty, ashamed	Patients felt dirty and ashamed
'When the nurses and doctors came in with their gloves and caps on I felt like I was myself a big infectious bacteria – it was not a good feeling, let me tell you'. (4)	Made to feel like an infectious bacteria by heavily clothed healthcare workers	Felt dirty due to heavily clothed healthcare workers	
Several patients said in interviews that they didn't understand why they were in isolation. (7)	Lacked understanding of why they were in isolation	Lacked understanding of why in isolation	Patients lacked understanding of why they were in isolation
Among 238 patients in an isolation ward, 31% stated that they did not know why they were in isolation. ($p < 0.05$) (8)	Almost 1/3 of patients lacked understanding of why in isolation	Lacked understanding of why in isolation	
67% of 30 isolated patients were not satisfied with the information they had received, compared to 98% of 60 patients not in isolation. (3)	67% of isolated patients weren't satisfied with information compared to 98% of non-isolated patients	Dissatisfied with information received about condition	Patients were dissatisfied with information received about condition

The number after each data unit refers to a number you have given the article each data unit came from. This is important so that you can later go back and see the context of each data unit, which may contain information you will need when you write up your results.

that answer your research question (see Table 3, column 1):

- Direct quotations from participants in a study
- The authors' descriptions, interpretations and summaries of their results
- Results from a questionnaire or survey (including tables and/or figures; see the last two data units in Table 3, column 1)

Condensing and coding: Figuring out the 'gist' of each data unit

It would be lovely if the human brain could synthesize data all by itself, but unfortunately our brains like to play tricks on us. If you think after marking up all your articles that you have the 'gist' of what they say and what the answer to your research question is, think again. The only way to do research with integrity and according to

ethical guidelines, and to faithfully represent research findings from multiple studies, is to systematically tabulate and summarize, and sometimes even count, your data: on your computer, on paper, using post-it notes, or even on a whiteboard if you want a large overview.

Start by making a table with several columns and many rows and copying over all the relevant results you have highlighted or underlined in your articles to the left-most column (see Table 3, column 1). These are the answers to your research question that the studies you are using have found. These are your 'data units'.

When reviewing studies on a topic, the data units may all be very similar, or they may be very different and even contradictory. For example, perhaps all of the studies you are summarizing found that people in isolation on infectious disease wards missed their family and friends. But perhaps some patients felt secure and protected by being in isolation, while others felt anxious and in danger. Your job is to record what the studies showed and what people said as faithfully as possible, without skipping over any data units because they do not agree with others, or because they are just like all the others.

Once you have filled a long column with all the findings or data units from the published studies that speak to your research question (keeping track of which study each data unit comes from), it is time to simplify what you have written in each square so that you can easily group together your data into overarching categories. If your data units consist of long descriptions you should have a column that condenses their content, as in Table 3, column 2, so that you can get an easy overview. If the data units are brief, however, you can go directly to the next column, where you 'code' your data. This means that you write a word or brief phrase that summarizes each piece of data and tries to capture its essence. Depending on what kind of research question you have these may be phrases like 'Felt anxiety' or 'Wanted more information' or 'Appreciated when nurses used humour' or 'Found mobilization after surgery to be scary'.

Note that a table for data analysis should always include the research question at the top of the page. Otherwise it is too easy to start listing data that is interesting, but not relevant to your research question.

The actual analysis of data really begins when you begin to code the data, and questions of interpretation come into play. For example, should 'participants described experiencing a great deal of sadness and disconnect over being separated from their family members' be separated into two codes, one that has to do with sadness and one that has to do with feeling disconnected? Another example is determining whether the essence of 'I felt like I was myself a big infectious bacteria' is that the person felt 'dirty and ashamed', or whether the person just had a vivid imagination and the description had nothing to do with feeling ashamed. Which is the accurate way to interpret the patient's statement? One way to make these judgments is to look at the context of the quotation in the original article. Another is to discuss with others how they interpret the essence of a data unit.

Synthesis and analysis: Finding patterns in the data

Once you have filled in the 'code' column you can move on to the next step of data analysis, which is to find overarching categories in your data. Of course, some patients feel one way, and some feel another, and an isolation ward at one hospital will not be the same as an isolation ward at another hospital. Your goal now is to find both the unifying categories and the variations in patient experiences, and even to take note of the 'outlier' findings, that maybe only came up in one study or for one patient.

Now you need to put on your categorizing glasses and step back from the details in the data units in order to start seeing the big picture. In practical terms, look for codes that say essentially the same thing or are about the same general topic, and group them in a first stage into a subcategory. For example, 'missed family' and 'felt family would have made it easier' are both about the patient's longing for family. So 'Patients longed for their families' might be the subcategory you choose to group these into (see Table 3, column 4).

If two people go through the same set of data codes they will not necessarily arrive at the same ways of coding and then categorizing those data units, just as you might organize your shirts according to sleeve length but someone else might organize theirs according to colour. Two people's subcategories after analysing the same data might be similar – you might have labelled a number of similar data units as 'feeling lonely', whereas I might have summarized them as 'wishing for more company'. If you do the analysis with another person you can compare notes and discuss and try to come to a consensus about what the best terms are in order to faithfully represent the data.

If you are doing the data analysis alone you should be sure to keep an open mind and to go back and check again and again that the way you are categorizing the data and labelling the categories stays true to the data and reflects what the data says. If, for example, one article found that, 'patients didn't like always being touched through gloves', it is not correct to label this 'feeling stigma', if the context doesn't support this interpretation. It could be that they did not like the actual sensation of being touched by latex gloves. Your categorizing of the data should stay as true to the actual words of study participants and study authors as possible. Do not abstract to too high a level.

The data is king and you have to bow to it at all times. You serve it; not it you. Even if your hunch going into the study was that patients in isolation suffer from stigma, you cannot interpret statements to satisfy your predetermined ideas about what you will find. Objectivity rules, however unkind a dictator it may be. This is true of all inductive research where you do not apply a pre-existing model or theory to your data (deductive methods), but let the data speak for itself.

Usually there will be some categories that come up again and again, whereas maybe some other experience of isolation will be mentioned in only one study. You need to faithfully present all the data, but also give some

Table 4. Further categorizing and abstracting from the data.

Research question:

How do patients experience isolation in infectious disease wards?

1	3	4	5
Data units	Codes = the essence of the meaning in each data unit (make sure they are relevant to the study's research question)	Subcategories = groupings of codes according to common denominators (should be in the form of an answer to the study's research question)	Category = groupings of subcategories according to common denominators (should be in the form of an answer to the research question)
Upon entering hospital isolation, participants described experiencing a great deal of sadness and disconnect over being separated from their family members. (2)	Missed family		
One woman commented that, 'If my husband and children had been able to stay in the room with me the whole time would have been a lot easier'. (2)	Felt family would have made it easier	Patients longed for their families	
'I felt like a pariah. I felt that my whole body was dirty and somehow shameful'. (4)	Felt dirty, ashamed		Distressing feelings and emotional reactions
'When the nurses and doctors came in with their gloves and caps on I felt like I was myself a big infectious bacteria – it was not a good feeling, let me tell you'. (4)	Felt dirty due to heavily clothed healthcare workers	Patients felt dirty and ashamed	
Several patients said in interviews that they didn't understand why they were in isolation. (7)	Lacked understanding of why in isolation		
Among 238 patients in an isolation ward, 31% stated that they did not know why they were in isolation. ($p < 0.05$) (8)	Lacked understanding of why in isolation	Patients lacked understanding of why they were in isolation	
67% of 30 isolated patients were not satisfied with the information they had received, compared to 98% of patients not in isolation (3)	Dissatisfied with information received about condition	Patients were dissatisfied with information received about condition	Lack of comprehension of condition

sense of what experiences were more prevalent and which were less common.

Sometimes your categorizing of the data will require only one level of categorizing or grouping of codes, but sometimes you might want to abstract up one more level, that is, group together categories at an even higher level. For example, you might think it helpful to group together the two subcategories, 'patients longed for their families', and 'patients felt dirty and ashamed' into the overarching category **distressing feelings and emotional reactions** (Table 4). Or, instead, perhaps it seems more logical to group

'patients longed for their families' with another subcategory, say, 'patients missed friends and social interaction' into an overarching category such as 'longings for social interaction'.

Depending on your data and the scope of your research question, you may even need a fifth column/level of categories to adequately organize your data. There is no pre-set number of columns you should have – the data in conjunction with your particular research question will govern how much synthesizing and abstraction is required. Beware, however, of creating categories that are so

Box 4 Structure of results.

Results

Write a brief summary of the overall results in a first paragraph.

Category (subheading): Distressing feelings and emotional reactions

Subcategory (one paragraph): Patients longed for their families

Subcategory (one paragraph): Patients felt dirty and ashamed

Category (subheading): Lack of comprehension of condition

Subcategory (one paragraph): Patients lacked understanding of why they were in isolation

Subcategory (one paragraph): Patients were dissatisfied with information received about condition

overarching and abstract as to be meaningless; for example, grouping ‘social situation’ and ‘fresh air’ into a category ‘environment’. The highest level groupings of data should be both a clear reflection of the concrete data, and in the form of answers to the research question.

Practical tip: Let your data be the guide to how many columns – that is how many levels of abstraction – you need. You may have captured adequately what the data says at the subcategory level, or the organization of your data may require a fifth column in your table with a higher level of groupings of categories.

When it is time to write your results, the categories in the final, right-most column will usually form the subheadings in the results section of your literature review. Lower-level categories will form the paragraphs under each subheading.

So the outline of your results might look like the example shown in Box 4.

Final advice

One of the biggest challenges in writing a literature review is to actually draw a conclusion, rather than to merely list the results of many different studies. It takes time to synthesize diverse data, to summarize statistical findings, or to draw out key categories of an experience. But keep in mind that the whole point of a literature review is precisely to draw overarching conclusions for the reader, as well as, ultimately, to point out their clinical relevance.

Practical tip: Synthesize and draw conclusions – don’t just list studies. Write category by category (‘ultrasound helped reduce pain’ or ‘patients want more information’), rather than study by study.

In practical terms this means that each paragraph in your results section should be about a particular result, for example, that people in isolation often feel they lack

understanding of why they are in isolation, rather than about a particular, individual study.

Make sure to describe how you conducted the data analysis in your thesis’ Methods section. It is important that readers understand how you went about the process and what method for data analysis you used. Describe the steps you took in concrete terms. How did you go about extracting key data – underlining, discussing in a group, putting it in a table? How did you go about summarizing quantitative data or coding data units – alone or with another person, according to what organizational scheme? How did you go about categorizing codes and comparing studies? And, ultimately, how did you organize your presentation of your findings?

As you struggle through and coding and the entire data analysis process – and it is usually a struggle! – do not forget to think a bit about the bigger picture. A well-executed general literature review can help you as students, soon-to-be clinicians or experienced clinicians keep abreast of developments in your respective fields, and can be a vital step in increasing the use of evidence-based practice in healthcare, ultimately ensuring that patients get optimal, up-to-date care.

Some final practical tips:

- Use the imperfect tense when writing the results: “Three studies showed . . .”, “Patients described . . .” (Just because the studies you examined found certain results does not mean you can claim that it IS so.)
- Begin each paragraph with a sentence that summarizes the results you will describe in that paragraph.
- Keep references at the end of sentences, when possible, so that they do not distract the reader in the text. Instead of writing ‘McCarthy et al. (2015) found that technicians with more experience. . .’ write ‘Technicians with more experience had higher PVC-insertion success rates than technicians with less experience in study at an American emergency department (McCarthy et al., 2015)’.
- Do not overgeneralize. Instead of writing ‘Patients in isolation missed their families’, write ‘In most studies patients’ most salient experience was missing their families’.
- Be careful not to draw too large conclusions. That three studies show that X had positive effects does not mean that X works. It means that X ‘may’ work, or ‘can’ work, or ‘might’ work.

Acknowledgements

Thanks to Malin Sandström and Helena Welch for the example research project of PVC insertion in patients with difficult venous access, and to our many students whose struggles and successes with data analysis in general literature reviews have inspired this article.

Declaration of conflicting interests

The authors declare that there is no conflict of interest.

Funding

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

ORCID iD

Anna Jervaeus  <https://orcid.org/0000-0002-5582-8142>

References

1. Bettany-Saltikov J and McSherry R. *How to do a systematic literature review in nursing: a step-by-step guide*. 2nd ed. London: McGraw-Hill Education/Open University, 2016.
2. Coughlin M and Cronin P. *Doing a literature review in nursing, health and social care*. London: Sage, 2017.
3. Hart C. *Doing a literature review: releasing the research imagination*. 2nd ed. London: Sage, 2018.
4. Elo S and Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008; 62(1): 107–115. DOI: 10.1111/j.1365-2648.2007.04569.x.
5. Graneheim UH and Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurs Educ Today* 2004; 24: 105–112. DOI: 10.1016/j.nedt.2003.10.001.
6. Whittemore R and Knafl K. The integrative review: updated methodology. *J Adv Nurs* 2005; 52(5): 546–553. DOI: 10.1111/j.1365-2648.2005.03621.x.
7. Bahl A, Pandurangadu AV, Tucker J and Bagan M. A randomized controlled trial assessing the use of ultrasound for nurse-performed IV placement in difficult access ED patients. *Am J Emerg Med* 2016; 34: 1950–1954. DOI: 10.1016/j.ajem.2016.06.098.
8. Mayan M, Robinson T, Gokiart R, et al. Respiratory isolation for tuberculosis: the experience of Indigenous peoples on the Canadian prairies. *Public Health Action* 2017; 7: 275–281.
9. Guilley-Lerondeau B, Bourigault C, Guille des Buttes A-C, et al. Adverse effects of isolation: a prospective matched cohort study including 90 direct interviews of hospitalized patients in a French University Hospital. *Eur J Clin Microbiol Infect Dis* 2017; 36: 75–80. DOI: 10.1007/s10096-016-2772-z.
10. McCarthy ML, Shokoohi H, Boniface KS, et al. Ultrasonography versus landmark for peripheral intravenous cannulation: a randomized controlled trial. *Ann Emerg Med* 2016; 68(1): 10–18. DOI: 10.1016/j.annemergmed.2015.09.009.
11. İsmailoğlu GE, Zaybak A, Akarca FK and Kiyan S. The effect of the use of ultrasound in the success of peripheral venous catheterization. *Int Emerg Nurs* 2015; 23: 89–93. DOI: 10.1016/j.ienj.2014.07.010.
12. Sou V, McManus C, Mifflin N, et al. A clinical pathway for the management of difficult venous access. *BMC Nurs* 2017; 16. DOI: 10.1186/s12912-017-0261-z.
13. Biyik-Bayram S and Caliskan N. Effects of local heat application before intravenous catheter insertion in chemotherapy patients. *J Clin Nurs* 2016; 25: 1740–1747. DOI: 10.1111/jocn.13193.