Systematic Review Process

DATA EXTRACTION

By Atefeh Davoudian

Zanjan University Of Medical Sciences

Spring 1401

Why Is Data Extraction Important?

- To summarize studies in a common format to facilitate synthesis and coherent presentation of data
- To identify numerical data for meta-analyses
- To obtain information to assess more objectively the risk of bias in and applicability of studies
- To identify systematically missing or incorrectly assessed data, outcomes that are never studied, and underrepresented populations

On Data Extraction (I)

- Extracted data should:
  - Accurately reflect information reported in the publication
  - Remain in a form close to the original reporting, so that disputes can be easily resolved
  - Provide sufficient information to understand the studies and to perform analyses
- Extract only the data needed, because the extraction process:
  - Is labor intensive
  - Can be costly and error prone
- Different research questions may have different data needs

On Data Extraction (II)

- Data extraction involves more than copying words and numbers from the publication to a form.
- Clinical domain, methodological, and statistical knowledge is needed to ensure the right information is captured.
- Interpretation of published data is often needed.
- What is reported is sometimes not what was carried out.
- Data extraction and evaluation of risk of bias and of applicability typically occur at the same time.

Comparative Effectiveness Reviews: Clarifying Research Terminology

- In the Evidence-based Practice Center Program, we often refer to two types of tables:
- Evidence Tables
  - Essentially are data extraction forms
  - Typically are study specific, with data from each study extracted into a set of such tables
  - Are detailed and typically not included in main reports
- Summary Tables
  - Are used in main reports facilitate the presentation of the synthesis of the studies
  - Typically contain context-relevant pieces of the information included in study-specific evidence tables
  - Address particular research questions

What Data To Collect?

- Use key questions and eligibility criteria as a guide
- Anticipate what data summary tables should include:
  - To describe studies
  - To assess outcomes, risk of bias, and applicability
  - To conduct meta-analyses
- Use the PICOTS framework to choose data elements:
  - Population
  - Intervention (or exposure)
  - Comparator (when applicable)
  - Outcome (remember numerical data)
  - Timing
  - Study design (study setting)

Data Elements: Population, Intervention, and Comparator

- Population-generic elements may include patient characteristics, such as age, gender distribution, and disease stage.
  - More specific items may be needed, depending upon the topic.
- Intervention or exposure and comparator items depend upon the extracted study.
  - Study types include randomized trial, observational study, diagnostic test study, prognostic factor study, family-based or population-based genetic study, et cetera.

Data Elements: Outcome (I)

- Outcomes should be determined a priori with the Technical Expert Panel.
- Criteria often are unclear about which outcomes to include and which to discard.
  - Example: mean change in ejection fraction versus the proportion of subjects with an increase in ejection fraction by ≥5 percent
- Record different definitions of “outcome” and consult with content experts before making a decision about which definition to use.

Data Elements: Outcome (II)

- Apart from outcome definitions, quantitative data are needed for meta-analysis:
  - Dichotomous variables (e.g., deaths, patients with at least one stroke)
  - Count data (e.g., number of strokes, counting multiple ones)
  - Continuous variables (e.g., mm Hg, pain score)
  - Survival data
  - Sensitivity, specificity, receiver operating characteristic
  - Correlations
  - Slopes

Data Elements: Timing and Study Design

- The data elements to be extracted vary by type of study.
- Consider collecting this information when recording study characteristics for randomized trials:
  - Number of centers (multicenter studies)
  - Method of randomization (adequacy of allocation concealment)
  - Blinding
  - Funding source
  - Whether or not an intention-to-treat analysis was used

Always Provide Instructions

- Provide “operational definitions” (instructions) indicating exactly what should be extracted in each field of the form.
- Make sure that all data extractors understand the operational definitions the same way.
  - Pilot-test the forms on several published papers.
  - Encourage communication to clarify even apparently mundane questions.

Single Versus Double Extraction

- Independent extraction of data by at least two experienced reviewers is ideal but is also resource intensive.
- There is a tradeoff between cost and the quality of data extraction.
  - Data extraction often takes longer than 2 hours per paper.
  - A reduction in the scope of the work may be necessary if independent data extraction is desired.
- Careful single extraction by experienced reviewers, with or without crosschecking of selected items by a second reviewer, is a good compromise.

Developing Data Extraction Forms (Evidence Tables)

- To address all needs, a generic data extraction form will have to be very comprehensive.
- Although there are common generic elements, forms need to be adapted to each topic or study design to be most efficient.
- Organization of information in the PICOTS (population, intervention, comparator, outcome, timing, and setting) format is highly desirable.
- Balance the structure of the form with the flexibility of its use.
- Anticipate the need to capture unanticipated data.
- Use an iterative process and have several individuals test the form on multiple studies.

Common Problems Encountered When Creating Data Extraction Forms (Evidence Tables) (I)

- Forms have to be constructed before any serious data extraction is underway.
  - Original fields may turn out to be inefficient or unusable when coding begins.
- Reviewers must:
  - be as thorough as possible in the initial set-up,
  - reconfigure the tables as needed, and
  - use a dual review process to fill in gaps.

Common Problems Encountered When Creating Data Extraction Forms (Evidence Tables) (II)

- Lack of uniformity among outside reviewers:
  - No matter how clear and detailed are the instructions, data will not be entered identically by one reviewer to the next.
- Solutions:
  - Develop an evidence table guidance document—instructions on how to input data.
  - Limit the number of core members handling the evidence tables to avoid discrepancies in presentation.

Sample Fields From a Table Guidance Document: Vanderbilt University Evidence-based Practice Center

- In the “country, setting” field, data extractors should list possible settings that could be encountered in the literature:
  - Academic medical center(s), community, database, tertiary care hospital(s), specialty care treatment center(s), substance abuse center(s), level I trauma center(s), et cetera.
- In the “study design” field, data extractors should list one of the following:
  - Randomized controlled trial, cross-sectional study, longitudinal study, case-control study, et cetera.

Samples of Final Data Extraction Forms (Evidence Tables)

- For evidence reports or technology assessments that have many key questions, data extraction forms may be several pages long.
- The next few slides are examples of data extraction forms.
- Remember, there is more than one way to structure a data extraction form.

Tools Available for Data Extraction and Collection

- Pencil and paper
- Word processing software (e.g., Microsoft Word)
- Spreadsheet (e.g., Microsoft Excel)
- Database software (e.g., Microsoft Access, Epi Info™)
- Dedicated off-the-shelf commercial software
- Homegrown software

Data Extraction Sample

Guideline elements[a]	Author (year of publication)	Design/ level of evidence	Sample	Setting	Intervention measure	Outcome/ recommendation
Skin care	Pittman et al. (2012)	RCT/ II	59	ICU (USA)	To compare (1) bowel management system (BMS) catheter; (2) rectal trumpet (RT) utilized as a rectal fecal incontinence device; and (3) usual care (UC) consisting of barrier creams and/or a fecal pouch collector.	No significant difference in HAPU prevalence ($p = .63$)
Emerging therapies	Brindle and Wegelin (2012)	Two group quasi-experimental/ III-1	85	CSICU (USA)	To evaluate the silicone border foam dressing in the sacrum area	No significant differences in the incidence between both groups ($p = .3$)
	Park (2014)	Quasi-experimental/ III-1	102	ICU (South Korea)	To evaluate the silicone foam dressing in the sacrum area	Significant decrease in HAPUs incidence ($p < .001$)
	Santamaria et al. (2015b)	RCT/ II	313	ICU (Australia)	To evaluate silicone foam dressings when applied to the sacrum and heel in the emergency department and maintained throughout their ICU stay	Significant decrease: Overall incidence of HAPUs ($p = .001$), sacral event ($p = .05$), heel event ($p = .002$)
	Santamaria et al. (2015a)	Pre-post quasi-experimental/ III-1	341	ICU (Australia)	To evaluate silicone foam dressings when applied on heel in the emergency department and maintained in ICU	Significant decrease heel HAPUs incidence ($p < .001$)

Data Extraction Sample

Table 1. Quality improvement project description and quality assessment (n=12)

Author Year Country	Sample and setting	PIP Programs and components (in dot points)	Main results (only ICU results are listed here in projects that included non-ICU samples)	QI-MQCS* criteria met
Azuh et al. 2016 USA	68-bed Medical ICU Sample: patients admitted to MICU with Braden Scale score < 19 n=3233	Early mobilisation program • Assessing patients' mobility level • Introducing a new staff role: patient mobility assistant • Mobilising patient: repositioning, sitting on edge of bed/on chair, walk with/without assistance • Staff education • Patients/family education	• PI cumulative incidence: significant decrease from 9.2% in 2011 to 6.2% in 2013 (p=0.041) • Hospital readmission of MICU: significant decrease from 17.1% in 2011 to 11.5% in 2013 (p=0.001) • Significant care processes improvements in repositioning; assistance with daily living activities; bed to chair mobility (P<0.001) • Patient satisfaction: 97% (207/213). Total of 213 patients were surveyed for this item.	13/16
Baldelli et al. 2008 USA	An ICU in a university medical centre Sample and sample size: no information	Pressure ulcer prevention bundle • PI risk and skin assessment • Elevating bed head at ≤ 30^0 • Moisture prevention • Repositioning • Elevating heels • Optimising nutrition • Using pressure relief mattress • Staff education	• PI period prevalence: decrease from 42% in 2005 to 20% in 2006 (significance not reported) • PI cumulative incidence: decrease from18% in 2005 to 10% in 2006 (significance not reported)	11/16
Ballard et al. 2008 USA	2 ICUs: a 26-bed ICU (trauma, neurosurgical, and surgical combined); and an 18-bed medical ICU in a hospital Sample and sample size: no information	A multicomponent program • Redesigned risk assessment and documentation chart • Staff education • Repositioning • Revising existing protocols • Weekly prevalence assessment • Using a new skin care wipe	• PI period prevalence: reduced from >30% to <10% over 18 months (significance not reported)	11/16

Data Extraction Sample

Author	Design, sample and setting	Primary outcome	Secondary outcome	Classification PU	Follow up
Lazzara and Buschmann, 1991[17]	RCT $n = 74$ Nursing homes	Exp.: Air -filled overlay (SOF CARE) PU Categories I-II $n = 32.2\%$ ($n = 10/31$) PU Categories I $n = 16.1\%$ (n = 5/31) PU Categories II n = 16.1% ($n = 5/31$) Contr. Gel mattress (No specification) PU Categories I-II $n = 31.7\%$ ($n = 8/26$) PU Categories I $n = 15.8\%$ ($n = 4/26$) PU Categories II $n = 15.8\%$ ($n = 4/26$) P = not reported	*PU Healing:* Air-filled overlay 58% ($n = 7/12$) PU improved or decreased. 41.7% ($n = 5/12$) increased or unchanged Gel mattress 60% ($n = 9/15$) improved, 40% ($n = 6/15$) worsened or unchanged P = not reported	Not reported	6 months Weekly skin assessments
Sideranko et al, 1992[20]	RCT $n = 57$ Hospital Surgical intensive care units	Exp.: Static air mattress overlay (4-in. thick, Gay Mar Sof Care bed, Unikion Gay Mare industries Inc. Orchard park, New York) PU $n = 5\%$ ($n = 1/20$) Contr.1: Alternating air mattress overlay (1 ½ inch thick, alternating air mattress Lapidus Airfloat system, American Hospital supply Corp., Valencia, California) PU $n = 25\%$ ($n = 5/20$) Contr.2: Water mattress overlay (4 in. thick water mattress Lotus RXM 3666, Connecticut Artcraft Corp. Naugatuck, Connecticut) PU $n = 12\%$ ($n = 2/17$) NS	*Mean pressure position:* Alt 3800 mmHg Static air 2500 mmHg Water 2330 mmHg	Not reported	Total days not reported Mean follow up 9.4 days
Takala et al, 1996[21]	RCT $n = 40$ Hospital Acute respiratory failure, intensive care, no traumatised patients	Exp.: Carital air-float system (series of 21 double air cells). (Carital Optima, Carital Ltd.) PU Categories I-II $n = 0\%$ Contr.: Standard hospital mattress 10 cm thick foam mattress, density 35 kg/m^3 (Espe Inc. Kouvola, Finland) PU Categories I-II $n = 36.8\%$ ($n = 7/19$) PU Categories I $n = 12$ PU Categories II $n = 1$ P <.005	*SkinT° at area pressure expose* was lower at air mattress ($P < .001$) *Pressure interface* between skin—mattress was lower in air mattress, most prominent at sacrum (different days ranging from $P < .001$ to NS)	Grading of Shea	1-year study 14 days follow up