SCIENTIFIC MEDICAL RESEARCH

WRITER:

إيهاب الرعود، أحمد سامي، أنس عنانزه

LECTURE:

15+16

Chapter 18: Questionnaire development

18.1 Questionnaire Design Overview:

Questionnaire (or **survey instrument**): is a series of questions used as a tool for systematically gathering data from study participants.

- A good questionnaire is carefully crafted for a specific purpose
- Most research studies must design new data collection instruments, but for some research topics, validated question banks are available and questions can be selected from them.

Identify general question categories → Select specific question topics → Choose question and answer types → Check wording → Choose order → Format layout → Pre-test → Revise → Use

18.2 Questionnaire Content:

- **Start** by compiling a list of the topics the survey instrument must cover. (These usually come from your main consent)
➔ **For example**: if you are collecting data about prevalence of smoking among students, you think of the question that makes relationships with this phenomenon that you're interesting in. (is age important for you? Is the income of the family important? Is the living place important? Is educational level important? Etc...)
➔ A source for these types of questions (your content) usually comes from the literature and based on the topic of your research.
- Include questions that confirm eligibility to participate.
➔ You should be asking questions that confirm that such participant is eligible to be included or not.

   For example: if you're aged less than 17, please DON'T answer this questionnaire.

   ❖ **The first set of questionnaire items typically enables the researcher to confirm that participants meet the eligibility criteria for the study**. For example, if only currently registered students are eligible to participate in a university-based cross-sectional survey, one of the first questions should be about school enrolment status. Participants who report that they are not students must be excluded from analysis based on that answer

These are the major questions that usually could be asked in a questionnaire:

Demographics | Key exposures | Key diseases/ outcomes | Related exposures and outcomes

Systems thinking: is the process of identifying the underlying causes of complex problems so that sustainable solutions can be developed and implemented.

➔ May include potential cofounders: factors that might influence the relationships between key exposures and outcomes. Ex: smokers consume more alcohol so higher liver disease rate, so if you're studying the relation between smoking and liver disease alcohol is a cofounder variable.

18.3 Types of questions:
- **Closed ended questions (close-ended questions) allow a limited number of possible answers:**
  - Date/time
  - Numeric
  - **Categorical**: **ordinal** (ranked) or **nominal** (unordered)
  - **Open-ended**, or **free-response**: questions allow participants to explain their answers at length.
  - ❖ Most questionnaires questions are usually closed-ended.
  - ❖ Closed-ended questions are usually easier to statistically analyse than open-ended questions.

    Examples of types of questions →

✔ Yes or no questions are called dichotomous or binary.

✔ Nominal (unordered) questions there is no order in the answers, For example: what's your type of film? Whatever you answer, action, comedy, or documentary, it's still a preference and there is no advantage of selecting any over the other.

✔ Example on ranked questions: educational level.

What's your highest educational level?

Less than high school -> 1 (we give them one point)

High school -> 2

College without degree -> 3

College -> 4

Postgraduate -> 5

So, now there's ranking.

✔ In paired questions, you give them a choice of two, and you give them an alternative as well.

FIGURE 18-3   Examples of Types of Questions

Type	Sample Question	Sample Response Option for the Sample Question
Date	What is your birth date?	_ _ - _ _ - _ _ _ _ m m- d d - y y y y
Numeric	What is your height without shoes (rounded to the nearest half inch)?	_ _ . _ inches
Yes/no	During your lifetime, have you smoked more than 100 cigarettes?	❑ Yes ❑ No
Categorical/ multiple-choice: nominal (no rank)	What is your sex?	❑ Female ❑ Male
	What is your favorite type of film?	❑ Action/drama ❑ Comedy/musical ❑ Documentary ❑ Other:_____
Categorical/ multiple-choice: ordinal (ranked)	What is the highest level of education you have completed?	❑ Less than high school ❑ High school ❑ Some college but no degree ❑ College/university degree or advanced degree
	How much do you agree with this statement?: "No matter how much I exercise, I will not be able to lose weight."	❑ Strongly disagree ❑ Disagree ❑ Neutral ❑ Agree ❑ Strongly agree
	On a scale of 1 to 5, with 1 meaning poor and 5 meaning excellent, how would you rate your hearing (without the use of a hearing aid)?	Poor_____Excellent ❑ 1 ❑ 2 ❑ 3 ❑ 4 ❑ 5
Paired-comparisons	Do you prefer to drink coffee or tea?	❑ I prefer coffee ❑ I prefer tea ❑ I like coffee and tea equally ❑ I do not drink coffee or tea
Rank-ordering	List the following four political issues in order from most important to you (1) to least important to you (4): crime/safety, environment/energy, foreign policy/defense, taxes/revenue	Number from 1 (most important) to 4 (least important): __ Crime/safety __ Environment/energy __ Foreign policy/defense __ Taxes/revenue
Open-ended/ free-response	What is your biggest personal health concern at present?	_____ _____

18.4 Anonymity:
- When anonymity is important, avoid asking questions that could allow the participant's identity to be determined based on his response.
- For many types of studies, there is no need to collect names, contact information, or other identifiable information.
➔ You should not include any question that could lead to the breakdown of the ID of the participants.

18.5 Types of responses:

- Provide clear instructions about the acceptable types of answers for each question.
- Add response for "not applicable" when relevant.

FIGURE 18-4 Examples of Five-Point Responses for Ranked Questions

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Dissatisfied	Somewhat dissatisfied	Neutral	Somewhat satisfied	Satisfied
Very negative	Somewhat negative	Neither negative nor positive	Somewhat positive	Very positive
Poor	Fair	Good	Very good	Excellent
None	Few	Some	Many	Very many
Never	Rarely	Sometimes	Often	Always
Not important	Slightly important	Somewhat important	Very important	Extremely important

➔ This scale is called (**Likert scale)** which presents ordered responses to a questionnaire item that asks participants to rank preferences numerically, such as by using a scale for which 1 indicates strong disagreement and 5 indicates strong agreement.

➔ There's five points possibility, when do we give 1, 2, 3, 4 or 5? It depends...
- Some scales the higher the number the better the answer,
- Some scales the lower the number the better the answer,
- like this question: did you feel tired during the last week? Never here is better than always, those are called negative questions, and need to be recoded when we deal with the data that we're collecting.

18.6 Wording of questions:

- Check each question and all the response items for clarity.

FIGURE 18-5 Problems to Avoid

Problem	Example	Problem with the Example
Big words/jargon	Have you ever had a myocardial infarction?	Participants may not know that a "myocardial infarction" is a technical term for a heart attack.
Undefined abbreviations	Have you ever been told that you have BPH?	Participants may not know that BPH is short for benign prostatic hypertrophy or that BPH means an enlarged prostate.
Ambiguous meanings	What kind of house do you live in?	Without seeing a list of appropriate responses, it is not clear if the answer should be "an apartment," "a rental," "a split-level duplex," or "a single-family home."
Vagueness	Do you exercise regularly?	"Regularly" is not defined. A person who exercises most days of each week might assume that "regularly" means daily and say "no." Another person who exercises once a month may consider that regular. It would be better to ask "In a typical week, how many days do you exercise for at least 30 minutes?"
Double negatives	I did not find this visit with my doctor to be unpleasant. ❑ Disagree ❑ Neutral ❑ Agree	The wording of this question makes it hard to figure out whether a person who was satisfied with a visit should agree or disagree.
Faulty assumptions	Do your gums bleed during regular dental cleanings? ❑ Yes ❑ No	The question assumes that everyone has routine dental cleanings. If "I do not visit the dentist" is not an answer option, a person who does not have dental cleanings is forced to answer no.

FIGURE 18-5 (continued)

Problem	Example	Problem with the Example
Leading answers	What is your impression about the quality of services provided by Center City Hospital? ❑ Fair ❑ Good ❑ Great ❑ Excellent	This question's response options clearly are intended to lead to a positive response; there is no "poor" option.
Answers with a poor scale	How many hours a week do you watch television? ❑ 0 ❑ 1–3 ❑ 4–7 ❑ 8 or more	Even though most people watch more than 1 hour of television daily, which would put them in the "8 or more" response category, they may not want to choose an "extreme" answer. Their inaccurate responses will lead to a false report. Alternatively, these response options may cause respondents to misread the question as how many hours a *day* they watch television.
Lack of specificity	What is your income?	It is not clear if income refers to earnings per hour, week, month, or year, or whether it refers to pre- or post-tax income.
Missing answer options	What color are your eyes? ❑ Brown ❑ Blue	Many possible eye colors are missing.
Overlapping answer options	In a typical week, how many days do you eat fish? ❑ 0 ❑ 1–3 ❑ 3–5 ❑ 5–7	Participants who eat fish 3 days a week or 5 days a week will not know which response to select.

18.7 Order for questions:
- Start with easy questions before moving to more difficult or sensitive questions.
- Avoid **habituation**: once responders have given the same answer to so many questions (agree, agree and agree) that they continue to reply with the same response (to avoid mix up questions).

18.8 Layout and formatting:
- Provide instructions
- Use readable and large fonts and include adequate white space.

FIGURE 18-6   Example of a Self-Response Questionnaire

FIGURE 18-7   Example of a Telephone Interview Script

➔ **A filter question** or **contingency question** is one that determines whether the respondent is eligible to answer a subsequent question or set of questions. For example, participants who indicate that they have never used tobacco products can be prompted to skip a series of questions about smoking habits. For paper-based forms, instructions for skips must be carefully described in words.

18.9 Reliability & validity: مهم
- Reliability/precision ثبات الأداة : consistent answers are given to similar questions.
➔ Look for internal consistency (Cronbach's alpha, KR-20) and test-retest reliability.
➔ Reliable means consistent, to what extent is it precise.
➔ Consistency means if you administer this questionnaire to the same individual more than one time, it will give you similar responses
❖ One aspect of reliability is **internal consistency**, which is present when the items in a survey instrument measure various aspects of the same concept. Some survey instruments ask the same question several different ways, or ask a series of similar questions, in order to confirm the stability of participants' responses.
❖ For example, a questionnaire **might include two questions that are opposites of one another**, like "I enjoy eating most fruits" and "In general, I do not like to eat fruit." The expectation is that all respondents who say the first item is true will say that the second item is false, and vice versa. If a very high proportion of respondents' answers meet this expectation, then that is evidence that the responses are reliable.
❖ internal consistency can be confirmed with tests of intercorrelation such as **Cronbach's alpha and the KR20**.
❖ **Intercorrelation** is present when two or more related items in a survey instrument measure various aspects of the same concept.

- ❖ **Cronbach's alpha:** is a measure of internal consistency that is used with variables that have ordered responses.
- ❖ **The Kuder-Richardson Formula 20 (KR-20):** is a measure of internal consistency that is used with binary variables
- ➔ Both of these statistics are expressed as a number between 0 and 1**. Scores near 1 indicate an assessment tool with minimal random error and high reliability**. Agreement between two or more evaluations is another facet of reliability

- ❖ **Test–retest reliability** is demonstrated when people who take a baseline assessment and then retake the test later have about the same scores each time they are tested.
- ❖ **Interobserver agreement,** **or** **inter-rater agreement** describes the degree of concordance among independent raters assessing the same study participants.
- ❖ **The kappa statistic, also called Cohen's kappa** and represented with the Greek letter κ, determines whether two assessors who evaluated the same study participants agreed more often than is expected by chance. **a valid study will have a value of kappa that is close to 1.**

$$\text{Kappa coefficient} = \kappa = \frac{\text{Observed agreement} - \text{Expected agreement}}{1 - \text{Expected agreement}}$$

"Several approaches are used to evaluate the accuracy of assessment tools that rely on self-reporting, such as psychometric tests and surveys about attitudes and perceptions. These tools are often considered to be proxy measures for an underlying theoretical construct that cannot be directly measured. For example, **happiness and intelligence cannot be directly measured with physical or chemical tests**, but survey instruments can be designed to measure them indirectly. Some researchers use **the word concept** to describe a theory informed by observations and use the term construct to describe a theory informed from more complex abstractions. For example, there is no particular threshold at which an object becomes "heavy" or a person becomes "rich," **because those definitions will vary for different individuals**, but measurements of weight and income can guide the evaluation of these concepts. In contrast, notions about constructs like "trust" and "leadership" are more difficult to quantify. Other researchers consider a concept to be a general abstraction and a construct to be a multidimensional concept that has been carefully defined and crafted for research purposes. In practice, **the terms "concept" and "construct" are often used interchangeably**"
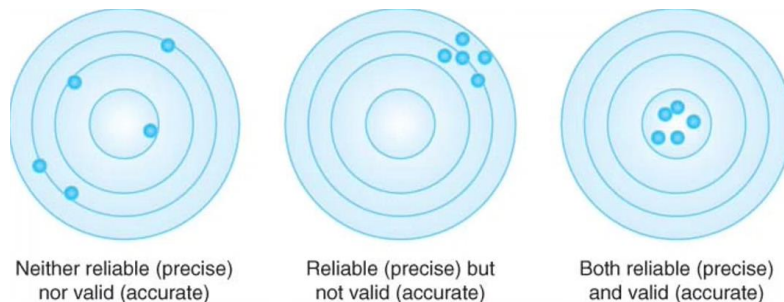
- ▪ Validity/accuracy صدق الأداة : the response or measurements are correct

  Types of validities:

- ▪ **Content (logical) validity:** is present when subject-matter experts agree that the questionnaire captures the most relevant details about the study domain. Usually reviewed by an expert panel.
- ▪ **Face validity:** is present when content experts and users agree that a survey instrument will be easy for study participants to understand and correctly complete.
- ▪ **Construct validity:** the test measures the theoretical construct it intended to assess.
- ▪ **Convergent validity:** is present when two items that the underlying theory says should be related are shown to be correlated
- ▪ **Discriminant validity:** is present when two items that the construct says should not be related are shown not to be associated
- ▪ **Criterion (concrete) validity:** uses an established test or outcome as a standard (or criterion) for confirming the utility of a new test that examines a similar theoretical construct
  Example on criterion validity: a new test of intelligence can be validated against standard IQ tests

- **Concurrent validity:** is evaluated when participants in a pilot study complete both the existing and new tests and the correlation between the test results is calculated. A strong correlation between the tests is evidence that the new test is valid
- **Predictive validity:** is appraised when the new test is correlated with subsequent measures of performance in related domains
- **Factor analysis:** uses measured variables to model a latent variable that represents a construct that cannot be directly measured with one question but appears to have a causal relationship with a set of measured variables.

> "Some statistical methods can provide information about which items in an assessment tool might be redundant or unnecessary and therefore can be removed without compromising the validity of the survey instrument. For example, **principal component analysis (PCA)** creates one or more index variables (called components) from a larger set of measured variables. The index variable is generated from the linear combination of measured variables, so it is a weighted average of the contributing variables. The PCA process determines the optimal number of components, the best measured variables to combine, and the best weights to use for the calculation"

Neither reliable (precise) nor valid (accurate)   Reliable (precise) but not valid (accurate)   Both reliable (precise) and valid (accurate)

18.10 Commercial research tools:

- Already-validated surveys instruments may be useful for incorporation into a new study.
- These may need to be purchased.

➔ The scale should be reliable, consistent, and accurate.

> several widely used and validated tests are available to researchers, such as:
>
> - The Beck Depression Inventory and the General Health Questionnaire (GHQ), which assesses psychological status.
> - The Mini-Mental State Examination (MMSE), which evaluates cognitive function.
> - The SF-36 and SF-12, which both measure health-related quality of life, a multidimensional construct that captures an individual's perceived physical, mental, emotional, and social well-being and the perceived impact of health status on the quality of daily life.

18.11 Translation:

Translation of the survey instrument into one or more additional languages may be necessary if the source population contains speakers of more than one language. Researchers using multiple languages must be certain that the translated version expresses the same meaning as the original survey. Accuracy may require the rephrasing of whole sentences, not just direct word-for-word translations.

There are two ways to ensure that the correct meaning is being conveyed:

1. **Back translation (double translation): in which one person translates the questionnaire from the original language to a new language and then a second person translates the survey instrument in the new language back into the original language.** ensures that the meanings of the original survey are maintained in the translated version.
2. A second approach is **to have two translators independently translate the survey instrument from the original to the new language**. The two translations are compared, and a consensus process is used to decide which words and phrases best convey the precise meaning and complexity of the original questionnaire.

18.12 Pilot testing:

- A **pilot test (or pretest):** is a small-scale preliminary study conducted to evaluate the feasibility of a full-scale research project**.**
➔ allows the researcher to correct problems with the survey instrument prior to data collection from participants.

A pilot test of a questionnaire is helpful for checking, among other issues:

1. The wording and clarity of the questions
2. The order of the questions
3. The ability and willingness of participants to answer the questions
4. The responses given, and whether the responses match the intended
5. The amount of time it takes to complete the survey

The researcher should ask several volunteers to help with the pilot test. These volunteers should be from the target population and meet the eligibility criteria for the study (in terms of age, exposure and disease status, and other key factors), but they should not be members of the sample population

Chapter 19: Surveys and Interviews

19.1 Interviews versus Self-Administered Surveys:

Interviews are very popular and valid method of data collection, and they could be used by researcher to improve response rate and improve accuracy of data.

Most primary studies collect data from individual participants using an interview method or self-administered questionnaire. **Self-reported surveys are usually the least costly and least time-consuming way to gather information.** However, interviews may allow for more detailed information to be gathered and can be accompanied by laboratory and other tests.

- ❖ The first decision to make about collection of quantitative data is whether to have a member of the research team interview participants or to have participants record their own answers.
- ❖ Interviews may promote accuracy and completeness buy may be expensive because of personnel costs.
- ❖ Self-administered surveys may yield a higher sample size, be more cost-effective, and be preferable for sensitive questions.

Interview	Self-Administered Survey
A member of the research team asks questions of participants and records their responses	Participants are provided with a set of questions and record their own answers

In-person (face-to-face) interview	Telephone interview	Completion in presence of researchers	Mail (postal) survey	Email/ internet-based survey
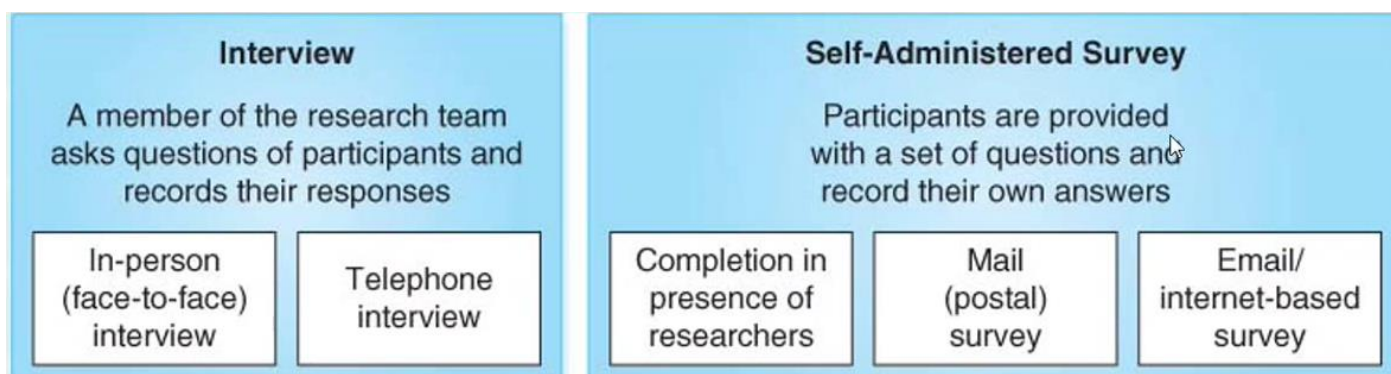
An Interview: is the process of a researcher verbally asking a participant questions and recording that person's responses

A self-administered survey: uses a questionnaire form that participants complete by themselves, using either a paper-and-pencil version or an online version of the survey instrument

Interview	Self-administered surveys
Advantages: ❖ trained interviewers ensure the accuracy and completeness of each questionnaire while recording the responses. **Disadvantage:** ❖ may be expensive because of personnel costs	**Advantages:** ❖ allows for the cost-effective collection of data from a large number of participants. ❖ May be the best way to get honest answers to sensitive questions (like questions about mental or sexual health). ❖ Can be completed at specific study site, such as a workplace or school or hospital, or they can be delivered by mail or the internet.

19.2 Recruiting Methods:

- - The goal is to maximize the participation rate among sampled individuals.
- - The recruitment methods are often paired with data collection method (such as using online methods to recruit for an Internet-based survey).
- - Provide multiple opportunities to participate, such as follow-up mailing (some participants don't respond from the first time, so you should follow with them).
- - Consider appropriate incentives. (Incentives such as small gifts or the opportunity to be entered into a drawing to win a prize)
- - Participation rates will likely be higher if recruits understand the importance and value of the research project.
- - several hundred calls made by **random-digit dialing**—calls to a computer-generated list of unscreened telephone numbers—may yield only a few people willing to participate in a study

Examples of methods for contacting members of the sample population:

- If the plan is to **interview people in person**, the best recruiting method may be to visit potential recruits at work, at school, at home, at a public venue, or at another appropriate location. Alternatively, if the contact information for sampled individuals is available, which would be true if recruiting patients from a collaborating clinic or recruiting employees from a cooperating corporation, then interviews could be set up by sending a letter or an email of invitation and then following up with calls to all of the sampled individuals
- If the plan is to **interview by telephone**, the participation rate will be higher if a letter invitation is sent first.
- If the plan is to collect data **via the internet**, the most effective method is contacting potential participants via email or website.
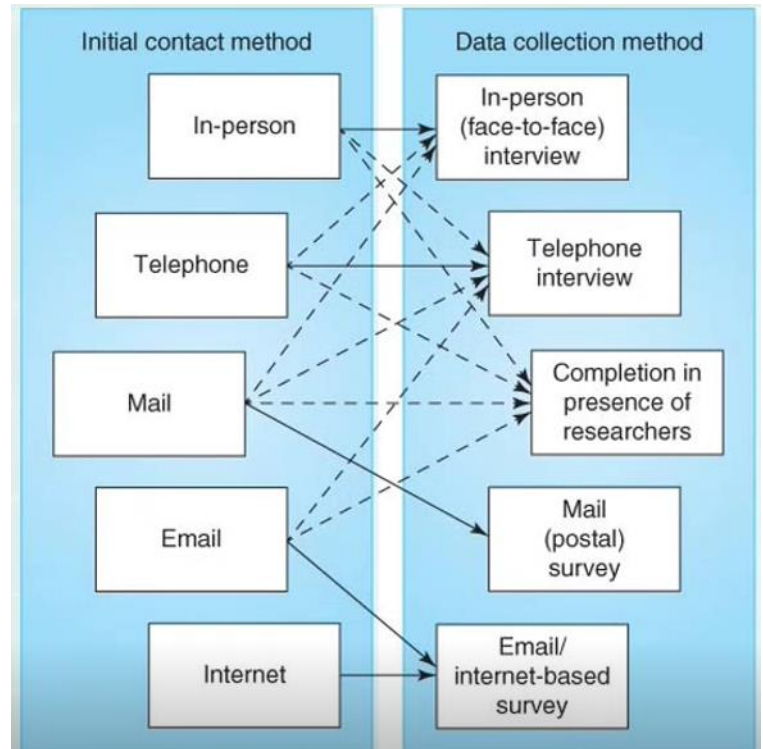
19.3 Data Recording Methods:

A decision must also be made about how responses will be recorded and when they will be entered into a computer database. There are two basic options.:

1. One is to record the responses on **paper** and to enter or scan them into a computer database later.
2. The other is to have interviewers or participants enter responses **directly into a database**.

❖ Paper-based surveys must later be scanned or typed into computers**; this can be expensive and time-consuming**, but they also **allow a large number of people to take the survey at one time** and place. (Also, error increases by this method, like missing data or data entry errors)

❖ Computer-assisted surveys input survey responses directly into a computer and can have built-in checks and "skips", but **not all population are comfortable with computers**. (**This method eliminates the need for later data entry, also missing of data is very minimal**)
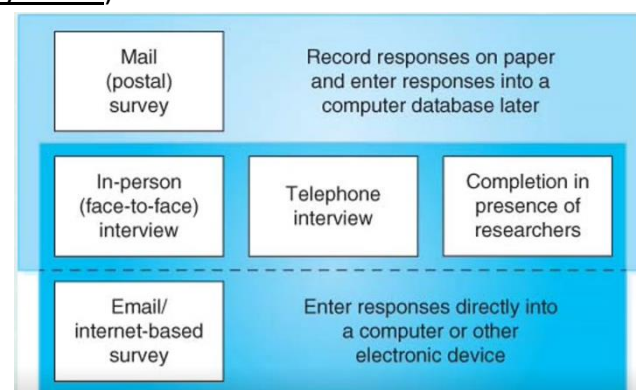
19.4 Training Interviewers:

- **Uniformity** in data collection procedures is important.
- Interviewers need access to comprehensive handbooks and training sessions when they can practice their skills.
- The interview process should be the same for all participants in a study. Uniformity is easiest to accomplish when all interviewers are provided with the tools they need to follow a standardized set of procedures.

- **Interviewer bias** is a form of information bias that occurs when interviewers systematically question cases and controls or exposed and unexposed members of a study population differently, such as probing individuals they believe to have the disease or exposure of interest for more information but not doing the same probing for participants they believe to be unexposed controls.
- interviewer handbook that provides information about the purpose of the study, details about interview logistics
➔ The training and handbook should:
- Explain the interview process step-by-step
- Specify exactly how to ask questions and record responses
- Identify any prompts or follow-up questions that the interviewer must use or is allowed to use
- Emphasize any restrictions against asking for clarification about particular items
- Provide checklists for handling problems that might arise during an interview, such as interruptions

FIGURE 19-4 Characteristics of Well-Trained Interviewers	
Characteristic	**Actions That Demonstrate the Characteristic**
Respectful	• Communicates pleasantly and professionally with all study participants and members of the research team • Has practiced interviewing enough to be comfortable with both the script and the interview process • Asks supervisors for assistance when it is needed
Organized	• Begins each scheduled interview session on time • Has all necessary materials on hand prior to the start of each interview session • Maintains meticulous records and completes all files and paperwork promptly
Considerate	• Dresses and grooms appropriately for in-person interviews • Is alert to modifiable conditions that may make interviewees uncomfortable, such as loud background noises or dim lighting • Allows adequate time for participants to respond to each question
Articulate	• Speaks at an appropriate pace and volume • Enunciates clearly • Uses an appropriate tone of voice (and, for in-person interviews, appropriate facial expressions and gestures) • Rereads questions and/or the list of closed-ended responses when a participant does not understand the question or the acceptable responses

FIGURE 19-4 Characteristics of Well-Trained Interviewers (continued)	
Characteristic	**Actions That Demonstrate the Characteristic**
Consistent	• Reads the script exactly as it is written • Probes for answers only when the script indicates that probing is approved • Does not provide explanations for any question unless an explanation is provided in the script or approved in the interviewer handbook
Impartial	• Avoids verbal and nonverbal expressions of approval or disapproval • Does not express personal opinions • Avoids leading interviewees toward a particular answer (for example, by placing special emphasis on particular words in a question or by probing until receiving a particular desired response)
Honest	• Does not fabricate or falsify reports • Records responses to open-ended questions verbatim, without rephrasing, paraphrasing, "correcting," or interpreting them
Careful	• Completes all steps of the interview process in the correct order, as prescribed by the interviewer handbook • Documents informed consent prior to conducting an interview • Does not skip any component of the interview • Completes all response forms correctly

Chapter 20: Additional Assessments

20.1 Supplementing Self-Reported Information:

Self-reports, such as those made during interviews and the completion of questionnaires, are essential data sources, but they have significant limitations. Respondents may not tell the truth, either because they do not accurately remember the answers or because they want to provide answers that are thought to be correct.

❖ Laboratory tests and other objective measures can be used to supplement and validate self-reported data. These types of scientific data are usually collected in person by a member of the research team so that the measurements can be collected according to precise protocols

20.2 Anthropometric Measures:

Anthropometry **is the measurement of the human body.**

- **Examples include**: height, weight, waist circumference, Hip circumference, Mid-upper-arm circumference (MUAC) and skinfold measurement that estimate the body fat percentage.
- Ensure privacy for participants is being measured.
- It is often best for **two members** of the research team to be present when measurements are being taken.
- When a child is being measured, it may be necessary for a parent or guardian to be present with the child.
- ➔ **Standard methods** should be used to take all anthropometric measurements. Any tools used for the measurements should be carefully calibrated **to ensure accuracy and reliability**.

20.3 Vital Signs:

Vital signs: are physiological measurements that provide clinical data about an individual's essential body functions. Most basic vital signs can be quantified accurately after minimal instruction.

- **Examples include:** body temperature, blood pressure, pulse (heart rate), and respiratory rate.
  (A thermometer measures body temperature. A manual or electronic sphygmomanometer (a blood pressure cuff) measures systolic and diastolic blood pressure. Measuring resting pulse and respiratory rates does not require any instruments other than timekeeping devices)
- The research protocol should state exactly how each measurement should be taken.
- ➔ **Standardization increases the precision and validity of the measurements**. Additionally, tests of inter-rater reliability (discussed later) can be used to confirm that all assessors generate similar or identical results when they measure the same person.

20.4 Clinical examination:

Examples include:

- Heart sounds
- Breath sounds and other respiratory functions
- Bowel sounds and the condition of the abdomen
- The range of motion (ROM) and the condition of the joints
- The condition of the skin, hair, and nails
- The health of the eyes, ears, nose, and mouth
- Mental status
- The ability to conduct activities of daily living
- Other signs of health or disease
- ➔ **An assessment form can guide the procedures that will be used for all participants.**

Assessment form includes each component of examination, including exact procedures to be used and the specific diagnostic criteria for each item on the assessment form, as well as the order in which these elements should be examined. Care should be taken to ensure the comfort, privacy, and safety of each person being assessed.

20.5 Tests of Physiological Function:

Tests of physiological function can provide helpful data about health status.

- **Examples include:** spirometry for lung function, electrocardiography for heart function, electroencephalography for brain function, and audiometry for hearing acuity.
- Consider the costs of these test
  ➔ Because of cost considerations, secondary analyses of existing medical records may be the best option for researchers whose study questions require the use of expensive equipment

20.6 Laboratory Analysis of Biological Specimens:

- **Examples include:** immunologic, genetic, and other tests for blood, urine, saliva, and other body fluids.
- Decide ahead of time whether results will be shared with participants.
  ➔ Some immunologic, genetic, and other studies require the collection of new body fluids or tissue biopsies, either as part of routine clinical practice or specifically for the purposes of the research project. Before new specimens are collected, a research ethics committee must verify that the potential physical risks to participants caused by the collection of the sample will be minimized. Some studies may be able to make use of existing specimen banks. These samples may be fully anonymous, or they may be linked to other information about the donor.

20.7 Medical Imaging:

- Medical imaging techniques are sometimes used to visualize parts of the human body.
- Examples include X-rays, CT scans, MRIs, and ultrasound.

20.8 Tests of Physical Fitness:

Kinesiology is the study of the mechanics, physiology, and psychology of body movement, function, and performance.

- Examples include test of strength, endurance, and flexibility.
- Safety must be the top priority.

Many different tests can be used to measure physical fitness levels:

- Cardiorespiratory fitness can be assessed using a 1-mile walking test, a 1.5- mile run test, or some other test of aerobic fitness.
- Measures of muscle strength and endurance include timed curl-ups, push-ups, pull-ups, flexed arm hangs, bench presses, leg presses, and grip tests (using a handgrip dynamometer).
- Flexibility can be measured using a sit-and-reach test (often measured with a flexometer) and other activities that stretch the lower back, hamstrings, or other muscle groups.
- Additional tests of fitness may assess agility, balance, coordination, speed, power, and reaction time.

20.9 Environmental Assessment:

- Environmental risk to human health can be assessed by trained observers with checklists and by laboratory tests.

The natural and built environments can have short- and long-term impacts on human health. Consider just a few of the many environmental factors that may affect the safety of a home:

- Is the entrance to the home accessible, or are there stairs or other barriers to access for people with mobility limitations? Is there adequate outdoor and indoor lighting?
- Are any stairs in the home loose or uneven? Do all stairs have handrails? Are all stairways free of clutter? Is any carpeting firmly affixed to the floor so that it will not slip?
- Does the bathtub or shower have a nonslip surface to prevent falls? Is the water heater set to prevent scalding and burns? Is the bathroom free of water damage, moisture, and mold?
- Do residents have reliable access to clean and safe drinking water?
- Is the kitchen free of pests and rubbish?
- Has the home been tested for toxic substances such as lead paint and asbestos? Is the home ventilated to prevent the buildup of radon gas? Are household chemicals, such as cleaning supplies, safely stored?
- Is the home equipped with working smoke alarms and carbon monoxide detectors?
- Does the home have adequate temperature control to prevent extreme heat and extreme cold?
- Are there sidewalks that facilitate safe walking near the home? Is the home located near a park, a playground, or another place where residents can safely engage in physical activity and recreation?
  Similar lists of questions could be developed for schools, healthcare facilities, workplaces, and other locations

➔ This kind of studies require a lot of resources and that's why they are not commonly performed in our region

20.10 GIS (Geographic Information Systems):

A geographic information system (GIS): is a computer-based platform for mapping the locations of events, identifying spatial clusters, and testing complex spatial associations
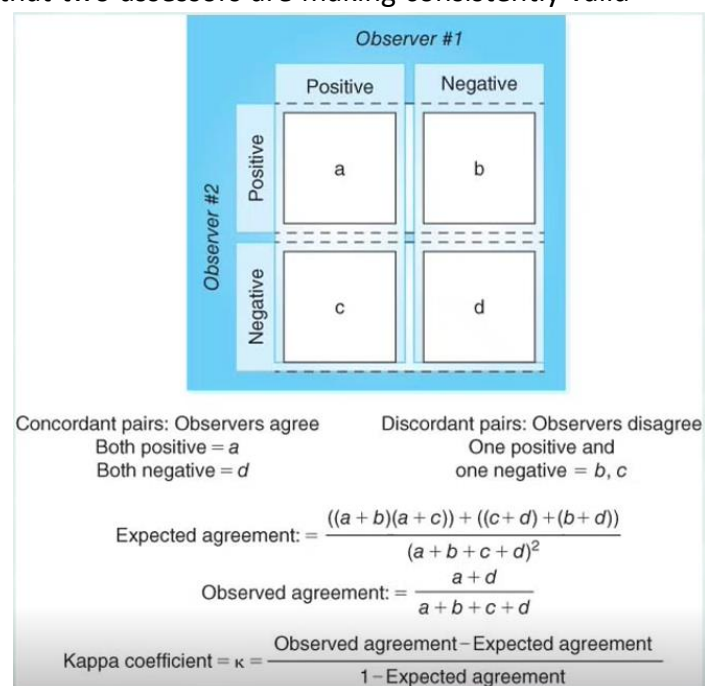
- Use a GPS (globe positioning system) to gather the latitude and longitude for key places, so that they can be mapped and incorporated into spatial analysis.

Example: To study the geographic distribution of coffee-houses that offer waterpipes (shisha) around universities in Jordan, this kind of studies offer possible interventions for policy makers to minimize the prevalence of water-pipe smoking among students.

20.11 Inter-Rater Reliability:
- Use the kappa statistic or another test to demonstrate that two assessors are making consistently valid measurements.

Statistical tests can be used **to determine the extent of agreement between two assessors who are evaluating the same study participants.** For example, a measurement known as **the kappa statistic** can indicate whether two radiologists examining the same set of X-rays reach the same conclusion about the presence or absence of a fracture more or less often than expected by chance. If the two radiologists agree as often as expected by chance, K = 0, If they agree on the interpretation of 100% of the X-rays shown to both of them, K = 1. If they agree more often than expected by chance, kappa will have a positive value somewhere between 0 and 1. Although complete agreement is rare, **a valid study will have a value of kappa that is close to 1**. Other measurements of inter-observer agreement or inter-rater agreement (also called concordance) can also be used to assess the validity and consistency of other assessment tools and procedures.

Observer #1 / Observer #2 table with cells a, b, c, d

Concordant pairs: Observers agree
Both positive = a
Both negative = d

Discordant pairs: Observers disagree
One positive and one negative = b, c

$$\text{Expected agreement:} = \frac{((a+b)(a+c)) + ((c+d) + (b+d))}{(a+b+c+d)^2}$$

$$\text{Observed agreement:} = \frac{a+d}{a+b+c+d}$$

$$\text{Kappa coefficient} = \kappa = \frac{\text{Observed agreement} - \text{Expected agreement}}{1 - \text{Expected agreement}}$$

Chapter 21: Secondary analyses

Some health research studies analyse existing clinical records, survey data, or population data rather than collecting new information.

21.1 Overview of secondary analysis:

- A **secondary analysis** is a study in which a researcher analyzes data collected by another entity.
- In a secondary analysis, **the researcher conducting the statistical analysis has not had (and does not have) any contact with the individuals whose data are being examined.**
    - ➔ Researchers should be aware and know how data was collected exactly.
    - ➔ Secondary analysis is often an excellent option for researchers with strong statistical skills but limited time and/or data collection resources.

21.2 Publicly Available Data Sets:

- ❖ **An anonymized data set (also called a deidentified data set):** is one that has been stripped of all potentially identifying information, such as names, street addresses, and personal identification numbers.
- ❖ **Deidentification:** is the process of removing potentially identifying information from a data file so that the data can be shared with others without violating the privacy of the individuals whose data are included in the file.

Many datasets are available online, such as a diversity of cross-sectional studies from the CDC:

- National Health and Nutrition Examination Survey (NHANES)
- National Health Interview Survey (NHIS)
- Behavioral Risk Factor Surveillance System (BRFSS)
    - ➔ The data may be downloadable to all, or there may be an application process for acquiring the relevant data.
    - ➔ Read all relevant supporting documentation and be aware of possible costs and authorship issues.

21.3 Private Data Sets:

- Professors (and others) may make **not-yet-analysed** data files available to students and other investigators.
- Ethical approval, careful review of the data collection methods, and frank conversations about authorship are required.
    - ➔ A request for access to a private data set is most likely to be granted when the new researcher has some existing connection to the original researcher.

21.4 Clinical Records:

- **Clinical records are a common source of data for case series.** Individuals working in clinical settings often can apply to gain access to patient records for research purposes.
- Ethical review of a research protocol is required before data are accessed (eg., **HIPAA** Act in USA).
    - ➔ **HIPAA:** a set of regulations about patient protection that must be carefully followed.
- Some data are electronic; others may be on paper and require data entry
- Records are often incomplete; absence of information in a file cannot be assumed to mean that the symptom or sign was not present in the patient.
- Some secondary analyses are conducted using deidentified records from registries. **A registry** is a centralized database containing information about people who have had a particular exposure or been diagnosed with a particular disease

- Most clinical sites require researchers to submit an application form to an oversight committee for review and approval prior to being authorized to access the data. The application must explain the goals of the study, the process that will be used to identify eligible patient records, the specific information that will be extracted from each patient's files, the steps that will be taken to protect the confidentiality of the data file, and the analysis plan. Applicants must also provide evidence of having successfully completed both research ethics training and specific instruction about patient privacy laws and policies. For example, researchers working with patient records in the United States must be prepared to comply with the **Health Insurance Portability and Accountability Act (HIPAA)** Privacy Rule.

21.5 Health Informatics, Big Data, and Data Mining:

- **Health informatics:** is the application of advanced techniques from information science and computer science to the compilation and analysis of health data.

- **Bioinformatics:** is the use of computer technologies to manage biological data.
   -> Bioinformatics typically focuses on analysis of molecular-level data (or, less often, tissue-level data). Clinical informatics and public health informatics usually focus on patient or population-level data. The tools of health informatics can be used to create novel data sets for research purposes.

- **Big data** refers to the analysis of data sets that are so large and complex that they require access to powerful hardware and special statistical software applications. These data sets may include data for many thousands or even millions of individuals from **data sources** such as:
   - Electronic health records (EHRs) or electronic medical records (EMRs), some of which use SNOMED CT (Systematized Nomenclature of Medicine Clinical Terms) as a standard terminology
   - Billing records, which often use ICD codes (Intonational Classification of Diseases codes) based on diagnoses or CPT codes (Current Procedural Terminology codes) based on procedures
   - Laboratory records, which often use LOINC codes (Logical Observation Identifiers Names and Codes)
   - Medication records, which often use NDC codes (National Drug Code identifiers)
   - Social media posts and other sources of information derived from the Internet
   - A diversity of other sources

- **Data mining** can be used to extract particular phrases from large sets of records

21.6 Ethics Committee Review:

- Any <u>data file containing possibly identifiable information</u> requires review by an ethics committee prior to beginning analysis.
- For all other data files, check with the appropriate committees about what review is required.
- Use of hospital records for research purposes always requires review by one or more research ethics committees. If the data for a secondary analysis come from a private source, then, prior to even looking at the data set, the analyst usually must obtain clearance from his or her own institution and perhaps also from the institution that houses the data. This permission must be secured prior to even looking at the data set. The application for permission to analyse existing data is often shorter than the application required for primary studies, and review usually can be expedited. It is better to err on the side of submitting an unnecessary proposal than to erroneously presume that a project is exempt from review without confirming the validity of this assumption.

- ❖ Most publicly available data, especially those collected by government agencies or federally sponsored researchers, were collected under protocols approved by one or several research ethics committees and then stripped of all personal identifiers prior to being shared. Additional approval by an ethics committee at the institution where the secondary analysis will be conducted is often not required <u>when several conditions are met</u>:
  - The data were collected after approval by a trusted organization's research ethics committee.
  - The data set contains no individually identifying information.
  - The data to be analysed are publicly available.
- ❖ However, researchers are responsible for becoming familiar with the requirements of their own institutions and ensuring that their work is compliant with all institutional policies. When there is any doubt about whether review is required, the institutional review board should be consulted.

Chapter 22: Systemic Reviews and Meta-Analysis

A **systematic review** is the careful compilation and summary of all publications relevant to a particular research topic, and a **meta-analysis** creates a summary statistic for the results of systematically identified articles.

22.1 Overview:

- ❖ Most scientific research projects seek to identify new findings derived from a single study population, but the goal of tertiary analyses is to engage in the scholarship of integration.
- ❖ **Synthesis research:** <u>integrates existing knowledge from previous research projects</u>. The common types of synthesis research in the health sciences include narrative reviews, systematic reviews, and meta-analyses.
- ❖ **<u>The goal of a review article</u>** is to synthesize what is already known about a topic by connecting previous studies and offering new interpretations of their contributions to scientific knowledge.
- ❖ However, review articles have limitations. Not all journals publish review articles as they are regarded as exhibiting less originality than other types of research.
- ❖ **All reviews require:**
  - An extensive search of the literature.
  - The extraction of key information from relevant articles.
  - The clear and concise presentation of this information.
- ➔ They all have the same objective and population.

FIGURE 22-1 Key Characteristics of Reviews and Meta-Analyses

Approach	Narrative Review	Systematic Review	Meta-Analysis
Objective	Synthesize existing knowledge	Synthesize existing knowledge	Synthesize existing knowledge
Primary study question	What conclusions about this topic are supported by previous studies?	When all previously published studies on this topic are examined, what conclusions can be drawn?	When the results of all previously published studies on this topic are merged, what is the summary statistic?
Population	Published literature	Published literature	Published literature
When to use the approach	The goal is to describe a new perspective on a topic that can be supported by the existing literature.	The goal is to compare the findings of previous studies on a well-defined topic.	The goal is to summarize previous findings using pooled statistics.
Requirements	The researcher has excellent library access. The researcher has a unique perspective on the topic.	The researcher has excellent library access. The researcher can obtain every relevant article.	The researcher has excellent library access. The researcher has strong quantitative skills.

FIGURE 22-1 (continued)

Approach	Narrative Review	Systematic Review	Meta-Analysis
First steps	1. Decide what story the article will tell.	1. Decide on the specific objectives of the review. 2. Select the search methods that will be used to find potentially relevant articles. 3. Select inclusion and exclusion criteria for articles.	1. Decide on the specific objectives of the review. 2. Select the search methods that will be used to find potentially relevant articles. 3. Select the inclusion and exclusion criteria for the articles. 4. Decide how to assess the quality of the studies. 5. Decide how the results of the studies will be combined into one summary statistic.
What to watch out for	Limited publication venues	Publication bias	Studies that cannot be fairly compared
Key statistical measure	No statistics are required.	No statistics are required, but providing some results from included studies may be helpful.	Summary measures for included studies must be reported.

22.2 Selecting a topic:

❖ Each review needs to have an appropriate scope. So, when starting a tertiary analysis, the most important decision is the selection of a topic that is narrow enough that all the relevant publications can be acquired.

❖ Most successful reviews have more than just a few articles & less than hundreds of articles.

❖ The topic may need to be modified if a preliminary literature search does not yield an appropriate number of articles. <u>If an initial search of an abstract database yields only 8 possibly relevant articles, the topic probably needs to be expanded.</u> <u>If a search produces 352 articles, the topic needs to be narrowed</u> to a more specific disease condition, a smaller geographic area, or a reduced scope.

➔ **For example**: a review of risk factors for cardiovascular disease would be cumbersome. A very long book would be required to cover all the identified risk factors, and an article-length summary would provide such a superficial level of information that it would not be useful. There is a greater likelihood of success for a review article with a narrower scope—one that limits the types of risk factors, the particular cardiovascular diseases, and the population groups included in the analysis.

22.3 Library access:

❖ No review article can be written without excellent library access because **every relevant article** must be identified and obtained during a systematic review

❖ The article acquisition process usually requires access to a university library that allows affiliates to make numerous interlibrary loan requests.

❖ Check with an institutional librarian about policies and prices for accessing articles that are not part of the library's collection.

❖ The researcher must also prepare to maintain a meticulous system for tracking articles that have already been acquired, those that have been requested but not yet received, and those that need to be requested.

22.4 Narrative Reviews:

❖ Narrative reviews tell a story about a topic using evidence from the literature to support the "plot".

❖ A narrative review might summarize important clinical aspects of a disease or summarize the epidemiological profile of a well-defined population.

❖ Narrative reviews must be carefully organized by theme, methodology, chronology, or some other guiding principle, because they are intended to convey a perspective and not merely compile facts.

❖ A narrative may also be appropriate when the researcher has developed a unique conceptual framework or theory that can be illustrated with examples from the literature.

❖ A narrative review works best when the researcher has a unique perspective on a topic and/or a particular expertise in the field.

➔ However, narrative reviews are becoming less common as editors and reviewers push for the use of systematic methods. So, Researchers must be prepared to justify their selection of a narrative approach.
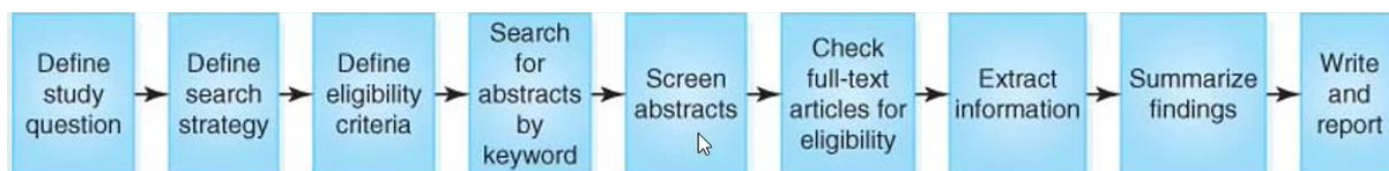
22.5 Systematic Reviews:

Systematic reviews use a predetermined & comprehensive searching & screening method to identify relevant articles while minimizing bias that occur when researchers handpick the articles they want to highlight.

➔ Therefore, after the identification of a focused study question, the most important decisions in a systematic review **are the selection of keywords and inclusion criteria.**

➔ The goal is to craft a search strategy that identifies all the articles ever published on the narrow, well-defined area covered by the review.

The systematic review process:

1. Identification of an appropriately narrow study question.

2. Selection of a well-defined & valid search strategy [from where, & what keywords]

3. Screening of all potentially relevant articles [using the inclusion criteria].

4. Extraction of relevant information from all eligible articles.

5. Summarization of the findings of these articles

22.6 Search Strategy:

The first section to be written in the methodology of systematic review is search strategy

- The MeSH dictionary (available through PubMed.org) can help with focusing or expanding search terms by identifying the definitions of key terms as well as synonyms and related terms.

> **For example:** a search for "health care costs" shows that synonyms for this term include "treatment cost" and "medical care costs." **Sub-headers** (sometimes called child terms) for "health care costs" include "direct service costs," "drug costs," "employer health costs," and "hospital costs." And "health care costs" fall under the headers (sometimes called parent terms of "health care economics and organizations," "health care quality, access, and evaluation," and "delivery of health care."

- Use **Boolean operators** such as AND, OR, and NOT.
- Confirm that the selected search string will capture several articles known to meet the eligibility criteria
- → Examples of using Boolean operators:
  - [a or b] → will find any abstract that includes "a" or "b" or both.
  - [a AND b] → will yield only those abstracts that include both terms.
  - [a NOT b] → will yield only those abstracts that include "a".
  - More complex search strings may use parentheses, such as [a AND (b OR c)], which will find any abstract that includes both "a" and "b" or includes both "a" and "c."

FIGURE 22-3 Examples of Using Boolean Operators to Expand or Restrict the Number of Abstracts Identified in a Database

Search String	Approximate Number of "Hits" in PubMed
cancer	3.5 million
bladder cancer	70,000
schistosomiasis	25,000
"schistosomiasis"[Mesh]	21,000
Schistosomiasis mansoni	10,000
cancer AND schistosomiasis	1500
bladder cancer AND schistosomiasis	650
bladder cancer AND *Schistosomiasis mansoni*	40
bladder cancer OR schistosomiasis	90,000
bladder cancer NOT schistosomiasis	65,000
colorectal cancer	200,000
colorectal cancer AND schistosomiasis	150
bladder cancer AND colorectal cancer AND schistosomiasis	20
(bladder cancer OR colorectal cancer) AND schistosomiasis	800

Note: Because the PubMed database is constantly adding new abstracts, the numbers in this table will not exactly match the results of a new search.

- ❖ Understanding the language used by MEDLINE and other databases allows for the design of a database appropriate search string. **For example:**
  - In MeSH language a "child" is defined as a person who is 6 to 12 years old. Individuals who are 2 to 5 years old are classified as "preschool children," and those who are 13 to 18 years old are "adolescents."
  - A keyword search of [child]—that is, a search for the word "child" in all of the titles and abstracts of articles indexed in PubMed—will yield hundreds of thousands more hits than a search for ["child"[Mesh]] that only searches for articles indexed with "child" as a MeSH keyword. The researcher must be attuned to these particularities when designing search procedures.

To check the appropriateness of search terms, the researcher can identify several articles known to be relevant to the study question and then confirm that the search string captures all of those articles. If the search misses one or more of those key references, then the search strategy needs to be modified.

However, this process must not be used to exclude disliked articles, which would cause the inclusion bias that systematic reviews seek to minimize.

Once a validated system for identifying all of the potentially eligible articles is in place, the selected abstract databases are systematically searched for articles that might meet the inclusion criteria. If the topic is appropriately narrow, then keyword searches can often reduce the number of abstracts and/or articles that must be screened for eligibility to a reasonable number, often several hundred articles rather than many thousands of articles.

22.7 Search Limiters:

❖ Be careful about decisions to limit the search database screened, the languages or publication years of articles, and other choices that may reduce the number of articles identified by the search.

➔ For example, if you decided to search from midline only, or to choose articles in English language only, **you will have to justify** why you took these decisions because this can be considered as inclusion bias as they limit the number of identified articles.

➔ Or if you searched for broad keyword terms like "adult" or "United States" they may limit the search as many papers reporting about these populations may not include these terms.

➔ Researchers should be cautious about using the built-in filters available in some abstract databases. For example, PubMed allows researchers to use filters to restrict results to particular types of articles (such as clinical trials or reviews), particular species (such as human-only studies), and particular age groups (such as infants or adults aged 65+ years). These limiters work only if an article was indexed appropriately by the submitting journal. Because many articles about humans do not add "human" as a keyword and many studies do not include keywords for the ages of participants or the study design, the built-in limiters often exclude many studies that would otherwise be eligible for the review. It is usually better to use study-specific exclusion checklists to screen out abstracts that are ineligible rather than to artificially limit the number of screened abstracts using filters.

22.8 Eligibility criteria:

As the database searches are being conducted, each identified article's title and abstract are reviewed to determine if they are eligible. The decision about an article's eligibility for inclusion in a systematic review or meta-analysis is based on predetermined lists of inclusion and exclusion criteria. These eligibility criteria should ensure that all of the included studies pertain to the main research question.

- create both a list of inclusion criteria and a list of exclusion criteria prior to screening the articles to avoid bias.
- Be prepared to justify all criteria, especially those related to quality screening.
- Consider whether to use snowballing and the gray literature to expand the search.

snowballing: to look up every article cited in the eligible articles in order to identify other not indexed relevant articles.

The gray literature: articles which are available to reviewers not published yet.

hand searching: all articles in the tables of contents of selected volumes of relevant journals are scanned [any other source of information about your topic like magazines].

the figure summarizes the process of systematic review steps ->
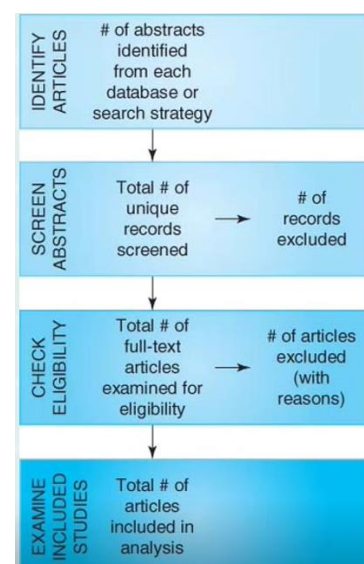
- ❖ If the study question includes specific exposures, diseases or outcomes, and/or populations, the eligibility criteria should ensure that all of the included studies match those "EDPs."
- ❖ The eligibility criteria can also impose some requirements about the study design and sample size, if those restrictions are scientifically justifiable.
- ❖ Studies of causality are often appropriately limited to reviews of experimental studies, but reviews that are not focused on causation generally do not need to exclude observational studies.

22.9 Data extraction:

After identifying eligible articles, Data extraction tables list descriptive characteristics about the contents of these articles like: Author, location, year, population, sample size...etc.

➔ The **data extraction table allows for easy compilation & comparison of observations** relevant to the study question. And it's usually included in the search report.

22.10 Systematic review results:

- ❖ The researcher should record and report both statistically significant findings ($p < .05$) and statistically insignificant findings ($p \geq .05$) that are related to the main study question.

Why both?

➔ Let's say you studied 20 articles. when you say 10 articles has shown an association between specific variables, it gives an idea that there is an association, but when you mention both; that 10 articles did show an association and other 10 didn't show any association, this would be more clear that it's 50% of all articles did show an association.

➔ **So the studies that do not show statistically significant results <u>are just as valuable as those that find a significant association</u>.** Actually, one of the primary contributors of the systematic reviews to the health literature is the ability to identify both: areas of consensus and areas of (disagreement & uncertainty).

- ❖ **Publication bias** occurs when articles with statistically significant results are more likely to be published than those with null results.

(Proving that publication bias has occurred may not be possible, but the presence of consensus should be cautiously interpreted when only limited number of studies are published on the topic or the results are mixed.)

We'll continue this chapter in the next lecutre...