Scientific Medical Research

Writer: Sarah Basel

S. Corrector: Ameen Alsaras

F. Corrector: Dana Alnasra

Doctor: Mohammad Darawad

❖ Chapter (19)- Surveys and interviews

- 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.

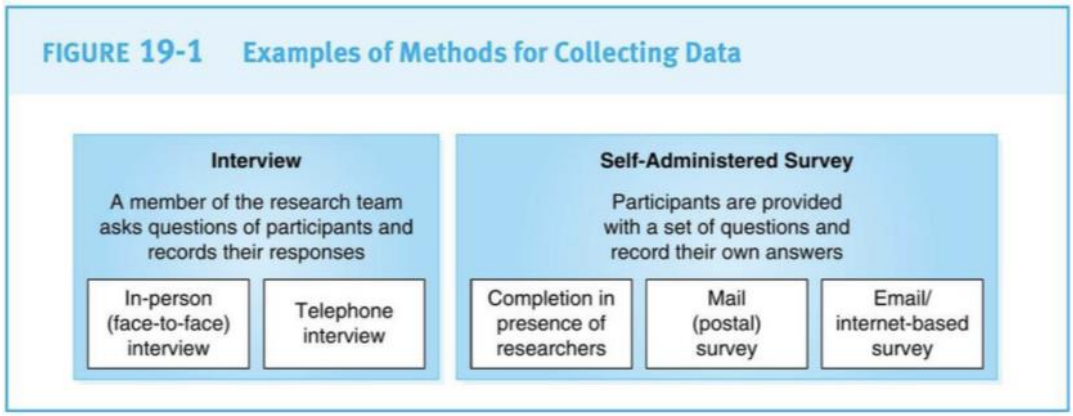
❖ Interviews Vs self-administered surveys:

The first decision to make about data collection is whether to have a member of the research team interview participants or to have participants record their own answers.

Interviews	Self-administered surveys
Conducted in person or via telephone. 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.

- The most important considerations when deciding which approach to use are the goals of the study and the expectations of the sample population members. Additional considerations are **cost, time, and potential barriers to participation**. For example, in terms of financial and time costs:
  - Interviews may require major time commitments from study personnel.
  - Mailed surveys incur direct costs related to photocopying, postage, and data entry.
  - Internet-based surveys may have relatively low costs if a free or low-cost survey-hosting website is used.

- In estimating the cost per participant, consider the likely participation rate. Mailing out 10 surveys may be necessary to receive one completed questionnaire, and the budget and sample size estimates should reflect this expectation.
- Time is another consideration. Asking participants to complete their self-administered questionnaires at the same place and time can generate a lot of data quickly. For example, a school-based survey could gather data from hundreds of students during one 20-minute period. One-on-one interviews may take a considerable amount of time per participant, and it may take months to schedule all of the needed interviews. Mail surveys may take an extended period of time which can be challenging for the researcher.
- The barriers to participation also vary according to the data collection method. Transportation to the interview site may be difficult for some interviewees. Discomfort with the telephone or computer may be a challenge for others.



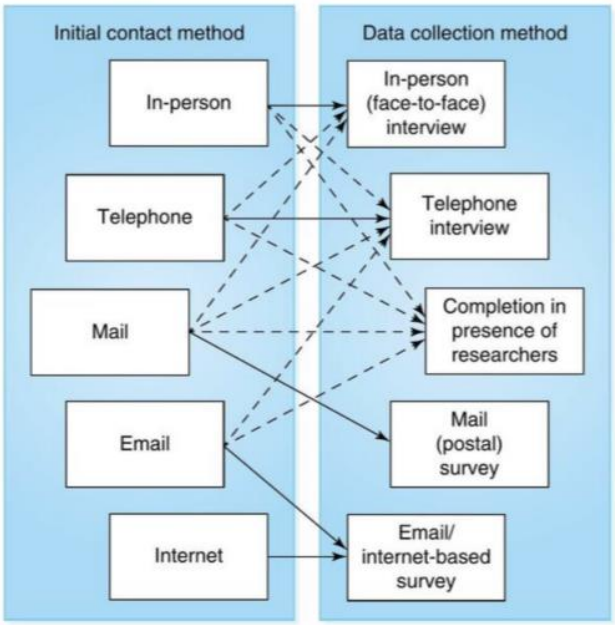
Recruiting Methods

Once a data collection method has been selected, the next step is to decide on an effective method for recruiting members of the sample population to be participants in the study. **The goal of recruiting is to maximize the participation rate among members of the sample population** to yield a study population that is reasonably **representative of the source population**. Ideally, the researcher should try to find a way to compare the characteristics of participants to the demographics of the source population as a whole. For example, in a school-based study the proportion of participants by grade can be compared to the overall distribution of students by grade in the participating schools.

- The best method for initiating contact with potential participants is often related to the intended data collection method:

- 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**, it may be possible to recruit some participants with cold calls (cold = not asked for). However, the participation rate will likely be higher if a letter of invitation is sent first. Sending a letter will also allow for the acquisition of signed informed consent forms prior to the interview, if they are required.
- If the plan is to collect data **via the Internet**, then contacting potential participants via **email** (this might be time and effort consuming) or a website may be the most effective method.

FIGURE 19-2 **Examples of Methods for Contacting Members of the Sample Population**



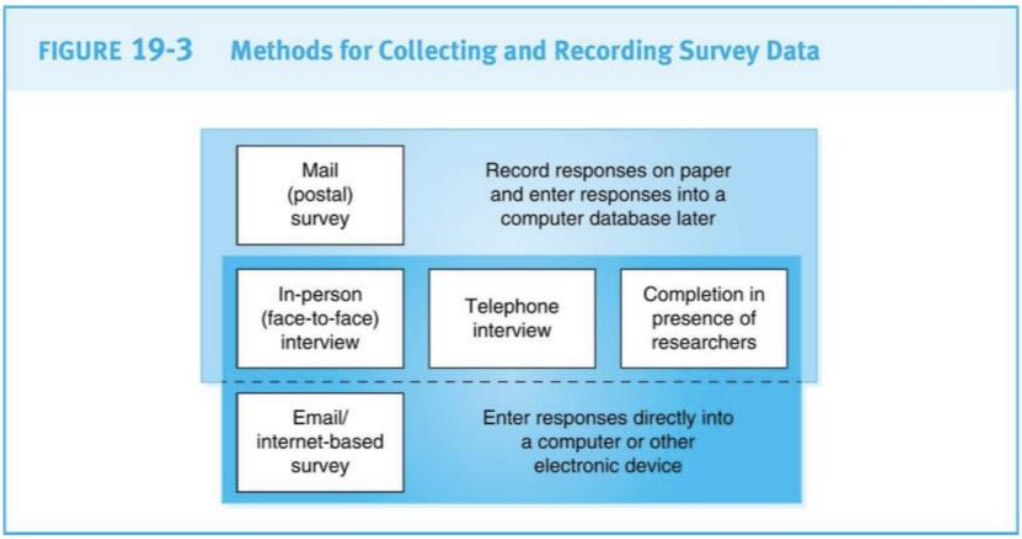
- Participation rates will likely be higher if recruits understand the importance and value of research project.

- Incentives such as small gifts or the opportunity to be entered into a drawing to win a prize may be an effective means of encouraging participation among those invited to be in the study. Any gifts or compensations must be approved by an ethics review committee prior to being offered.

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

- Paper questionnaires have several benefits. In some environments, they are required for the collection of data from a large number of participants at one time, as would be the situation when all students attending a school are asked to complete a questionnaire during the same 20-minute period. Paper instruments allow for the easy collection of signatures on informed consent statements, and some researchers value having paper records as a backup to electronic files. But paper-based surveys have a serious disadvantage: Unless somewhat expensive optical scan forms are used, all responses have to be manually entered into a computer at a later time. Data entry is often a very time-consuming process, and that can become costly.
- The major advantage of computer-assisted surveys is that they eliminate the need for later data entry. They may also simplify the questionnaire by automatically removing any questions not relevant to a particular study participant. For example, they may skip questions specific to females for participants who identify themselves as being male. The main limitation of computer-assisted surveys is that some populations are uncomfortable with computer technology. Discomfort with technology may be expressed in several ways. Older adults who have limited access to the Internet or do not routinely use computers may systematically choose not to participate in an internet-based survey.



Training Interviewers:

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. All interviewers should undergo role-specific training and have an opportunity to practice their interview skills. Each interviewer should be given a comprehensive interviewer handbook that provides information about the purpose of the study, details about interview logistics, an annotated script for the in-person or electronic interview, and annotated copies of all study forms. 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

All of this information should also be recorded in the study protocol.

- Interviewers usually feel more prepared for their role after attending one or more training sessions. Facilitators often begin training sessions by explaining the purpose of the research project, emphasizing the importance of strictly following the procedures spelled out in the interviewer handbook, and making clear the absolute necessity of **maintaining the confidentiality of all information that study participants share with them**.
- All paper response forms and/or computer-assisted data entry programs should be closely examined so that every interviewer understands exactly how to record participant responses.
- Each interviewer should have the opportunity to participate in several mock (fake) interviews from start to finish, including the informed consent process. Clear guidelines and lots of practice will help to create skilled, confident, and reliable interviewers. Well-trained interviewers will know how to make participants comfortable, **how not to intentionally or unintentionally guide participants toward particular answers rather than letting participants provide candid responses**, and how to complete all survey forms consistently and completely.
- Inter-rater reliability should be established.

You need to know all the following characteristics

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:

- Surveys and interviews are the most common sources of health data, but other measurements are often important supplements to self-reported information.

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*. Also, they may not know some of their health measures, such as their current weight or blood pressure. Laboratory tests and other objective measures can be used to supplement and validate self-reported data and to quantify attributes that require independent assessment.

Anthropometric Measures

- Anthropometry is the measurement of the human body, and anthropometric measurements are often important in health research, especially in studies of nutritional status.

Some of the most common body measurements are:

- Height (stature)

- Weight
- Waist circumference
- Hip circumference
- Mid-upper-arm circumference (MUAC)
- Skinfold measurements that estimate the body fat percentage

- Standard methods should be used to take all anthropometric measurements- Any tools used for the measurements should be carefully calibrated (marked with a scale of readings, and a zero-point should be set) to ensure accuracy and reliability.
- The individuals taking the measurements should be trained to use all equipment properly and to record results to the appropriate level of precision. They should also ensure privacy for participants while the measurements are being taken.

Vital Signs:

Basic vital signs are physiological measurements that can be quantified accurately after minimal instruction. These include:

- Body temperature
- Blood pressure
- Pulse (heart rate)
- Respiratory rate (breathing frequency)

- 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.

Clinical Examination

A well-trained clinician can make accurate and reliable assessments of many health states that machines are unable to assess well. For example, a clinician can examine:

- 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

- When a clinical examination is part of the data collection process, an assessment form should carefully describe each component of the examination, including the 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.

- **Tests of Physiological Function:**

  Tests of physiological function can provide helpful information about health status. For example, spirometry measures lung function, electrocardiography (ECG) measures heart function, electroencephalography (EEG) measures brain function, and audiometry measures hearing acuity. The costs associated with these tests must be considered when designing primary data collection protocols. Although some medically necessary tests may be covered by patient insurance plans, tests conducted primarily for the benefit of researchers must be paid for by the research team. 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. when tests are conducted as part of a primary research protocol for research purposes rather than clinical purposes, the research team must decide ahead of time, in consultation with specialists in medical ethics, whether the results of the studies will or will not be shared with patients and/or their health care providers. This decision must be disclosed to participants during the informed consent process prior to any measurements being taken.

- **Laboratory Analysis of Biological Specimens:**

  Tests of blood, urine, stool, saliva, and/or other biological specimens may be helpful for identifying the presence of a disease or markers for a disease, the characteristics associated with having a disease, and the risk factors for a disease. 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. The use of existing samples also requires ethics committee review and approval. Participants may have a right to know the results of the laboratory tests conducted on their own biological specimens, and the protocol should discuss how notification will occur.

- **Medical Imaging:**

  Medical imaging techniques are sometimes used to visualize parts of the human body. Examples are radiography (X-rays), computed tomography (CT) scans, magnetic resonance imaging (MRI), and ultrasound. The resulting images may be useful to researchers for purposes of diagnosis and/or for the assessment of responses to therapies.

- **Tests of Physical Fitness:**

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

  - Cardiorespiratory fitness can be assessed using a I-mile walking test, a I .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.

- Researchers must make the safety of participants their top priority. Appropriate precautions must be taken to ensure a safe environment.

- **Environmental Assessment:**

- Both the natural and built environments can have an impact on human health. Consider just a few of the many environmental factors that may affect the safety of the home:

  - Is the entrance to the home accessible, or are there stairs or other barriers to access for people with mobility limitations? Are any stairs in the home loose or uneven? Do all stairs have handrails? Is any carpeting in a stairway firmly affixed to each step? Are all stairways free of clutter? Do exterior and interior stairs have adequate lighting?

- Does the home have adequate temperature control to prevent extreme heat and extreme cold?
- DO residents have reliable access to clean drinking water?
- Is the kitchen free of pests and rubbish?
- 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?
- 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?
- 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?

- This kind of studies require a lot of resources and that's why they are not commonly performed in our region.
- Similar lists of questions could be developed for schools, healthcare facilities, workplaces, and other locations.
  - Some of these questions can be answered by trained observers. These assessors may describe findings qualitatively, assigning ratings like "high" or "low" to observed conditions based on predefined lists of rating criteria. Other assessments require quantitative measurement of environmental contaminants, such as tests of paint chips for lead or tests of basement radon levels. For some types of hazards, the exposure dose, frequency, and duration must be ascertained. Risk assessments may be conducted at one point in time or at several time points. Researchers must have the permission of owners and/or residents before they enter a building or conduct environmental assessments of a structure.

- **GIS (Geographic Information Systems):**

  Sometimes a map and/or spatial analysis of important features in the study area helps answer the study question. If so, a GPS (global positioning system) receiver can be used to acquire the geographic coordinates (in latitude, longitude, and altitude) for relevant locations, such as the homes of participants, nearby hospitals and other health care facilities, roads, schools,

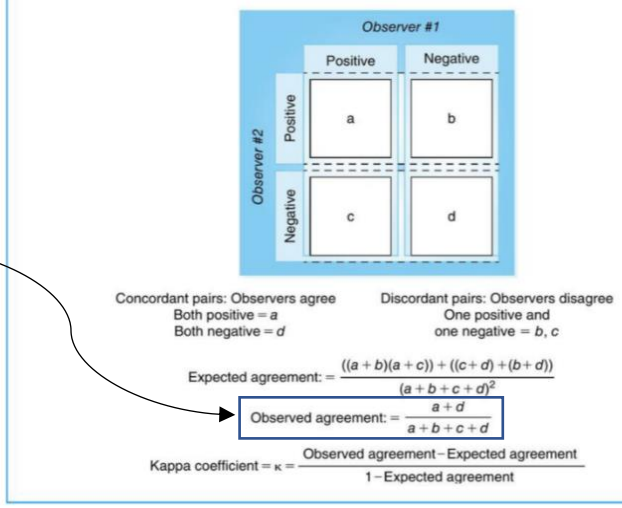
religious and social organizations, grocery stores, recreation facilities, water sources, and industrial sites. The coordinates for public locations can be collected by anyone, but permission from the owners or residents of private land may be needed before entering their property to take a GPS reading. The GPS coordinates for the homes of participants is individually identifying information, so precautions must be taken to protect geographically linked personal data.

- 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.
- **Inter-Rater Reliability:**

  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 = O, If they agree on the interpretation of 100% of the X-rays shown to both of them, K = l. If they agree more often than expected by chance, kappa will have a positive value somewhere between 0 and l. Although complete agreement is rare, a valid study will have a value of kappa that is close to l. 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.

- We are required to calculate the "observed agreement" only.
- Observed agreement should be at least 90%

FIGURE 20-1 Inter-Rater Agreement



Chapter (21)- secondary analysis:

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

Overview of Secondary Analysis:

For some studies, the data collection stage of the 5-stage research process is the step of acquiring existing data sets for secondary analysis. These data files may be publicly available individual-level or population-level data, privately held survey data, or electronic or paper health records. Whatever the data source, what makes a project a **secondary analysis is that the researcher conducting the statistical analysis has not had (and does not have) any contact with the individuals whose data are being examined.**

- A researcher conducting a secondary analysis contributes to scientific knowledge by analyzing and interpreting accumulated data that might otherwise remain untapped. Sometimes a researcher can download an entire data set from an Internet website or have it sent by email. Such files often contain already cleaned data that are ready to analyze within minutes of receipt. At other times, the data are available only as paper records or electronic files from which the relevant information must be extracted and entered into a new computer database prior to analysis.

Publicly Available Data Sets:

A growing number of governmental agencies allow researchers access to their anonymized data sets (also called deidentified data sets) that have had all potentially identifying information removed from the files. These organizations are experts at collecting data but often do not have the resources to conduct a thorough statistical analysis of an entire data set before it becomes relatively obsolete. Sharing data with external researchers is therefore a cost-efficient way to extract as much information as possible out of data sets, especially when the data were expensive to collect. Some research teams supported by federal funding agencies and some private organizations are also required to make their data available to researchers upon request, and others voluntarily share their data.

- Available data sets are often listed on the websites of government health agencies.

- Many datasets are available online, such as a diversity of cross-sectional studies from the CDC (these names are not for memorization):
  - National Health and Nutrition Examination Survey (NHANES)
  - National Health Interview Survey (NHIS)
  - Behavioral Risk Factor Surveillance System (BRFSS)

Private Data Sets:

- Individual researchers and small research teams may have data available that have not yet been analyzed. The researchers may have computerized data files that have not yet been fully explored, or paper records may have been set aside because they are not a current priority of the research team. Sometimes the original researcher or research team may have published the results of some portion of the data set, but left unanalyzed some of the other potentially significant, interesting, and novel aspects of the data set. In these situations, the original researchers may be open to a new researcher taking the lead on analyzing an underexplored portion of the data set and writing up the results for possible publication,
- 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. Students are most likely to have success asking their own professors for data sets to analyze. If students are interested in the work of a research group at another university or hospital, they may find it helpful to ask their professors to reach out to friends at the other institution. The ethics review committees of both institutions may need to approve the data sharing plan, especially if identifiable information might be included in the data file.
- When privately held data are shared with a new investigator, the original researchers usually expect to be coauthors on any resulting publication. The roles and responsibilities of each party should be agreed on as early as possible in the research process, preferably before the data files are shared.

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. 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. Sometimes the relevant information can be extracted from an electronic database. When electronic records are not available, a data extraction form can be created and used to compile the relevant information from each patient file. The extracted information can be entered directly into a computer database or recorded on paper for later data entry. Whenever possible, the data files should not contain any individually identifying information.

A major limitation of using existing clinical records is that patient records are often incomplete. Researchers cannot make any assumptions about the missing information. For example, researchers cannot assume that the absence of information about a symptom means that the patient did not experience the symptom. The patient might have had the symptom but failed to mention it to the clinician- Perhaps the clinician did not specifically ask whether the symptom was occurring. Maybe the patient did mention the symptom but the clinician did not record it, perhaps because the symptom did not seem especially relevant. Similarly, researchers cannot assume that the information in the medical records of one health care provider tells a complete story about those patients' health status. Consider medication use. Researchers cannot assume that a patient is not taking a particular medication just because that patient's records at one clinical site do not mention that the patient has been prescribed that drug. The patient might have been prescribed the medication by a clinician at some other site, And, even if the patient's records show that a prescription was written for a particular medication, that does not mean that the patient filled the prescription and took the drug. If the research question requires complete information about symptoms or medication usage or other details, a primary study design may be necessary.

Health Informatics, Big Data, and Data Mining:

- Health informatics applies advanced techniques from information science and computer science to the compilation and analysis of health 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

- Text mining and other forms of data mining can be used to extract particular phrases from large sets of records. Clinical informatics projects might use data mining techniques to explore hospital records. Public health informatics projects might use data mining and computational linguistics to explore social media events. Big data approaches have the power to reveal patterns and trends that are not apparent in smaller data sets analyzed with traditional statistical methods. Specialized training is usually required before researchers are prepared to implement data mining and other big data methods.
- **Ethics Committee Review:** any data file containing possibly identifiable information requires review by an ethics committee prior to beginning analysis.

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. The application for permission to analyze existing data is often shorter than the application required for primary studies, and review is usually able to be expedited.

- 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 when several conditions are met:
- 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 analyzed are publicly available.

However, researchers are responsible for becoming familiar with the requirements of their host 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.