CHAPTER 7

Overview of Study Designs

Many types of observational and experimental study approaches are useful for clinical and population health research.

7.1 Types of Study Approaches

Eight common study designs used for population health research are briefly described in **Figure 7-1**. The figure does not represent a comprehensive list of all types of study approaches. Many research projects use variations of one of these approaches, and in others a hybrid of two approaches might be suitable. A diversity of designs can be valid and helpful for collecting and analyzing new data, analyzing existing data, and reviewing the literature in the health sciences.

The study design selected for a particular project must be appropriate for the goals of the study. A series of questions can help identify the most suitable approach:

- Do new data need to be collected, or are there existing data sources that can be used to answer the study question?
- If new data will be collected, is an experimental study required or can the study question be answered using non-experimental methods?
- Is the research question based primarily on exposure status, disease status, or membership in a particular population?
- Are there time constraints? Some studies allow for rapid data collection, while others require months or even years of follow-up.

7.2 Primary, Secondary, and Tertiary Studies

The first step in selecting an appropriate research approach is deciding whether new data will be gathered. A **primary study** collects new data from individuals. A **secondary study** analyzes an existing data set or existing health records. A **tertiary study** reviews and synthesizes the existing literature on a topic. Each of these three major study approaches has

FIGURE 7-1 Summary of Study Approaches

Study Approach	Goal
Case series	Describe a group of individuals with a disease
Cross-sectional study	Describe exposure and/or disease status in a population
Case–control study	Compare exposure histories in people with a disease (cases) and people without a disease (controls)
Cohort study	Compare rates of a new (incident) disease in people with different exposure histories or follow a population forward in time to look for incident cases of a disease
Experimental study	Compare outcomes in participants assigned to an intervention or control group
Qualitative study	Seek to understand how individuals and communities perceive and make sense of the world and their experiences
Correlational (ecological) study	Compare average levels of exposure and disease in several populations
Review/meta-analysis	Synthesize existing knowledge

FIGURE 7-2 Key Considerations for Primary, Secondary, and Tertiary Studies

Study Approach	Study Plan	Key Questions
Primary	Collect and analyze new data	■ What are possible source populations? ■ Will it be possible to recruit enough participants?
Secondary	Analyze existing data	■ What are possible sources of usable data files? ■ What research questions can be explored with the available data?
Tertiary	Review and synthesize the literature	■ Does the researcher have access to adequate library resources? ■ Can the researcher reasonably expect to access *all* of the needed articles?

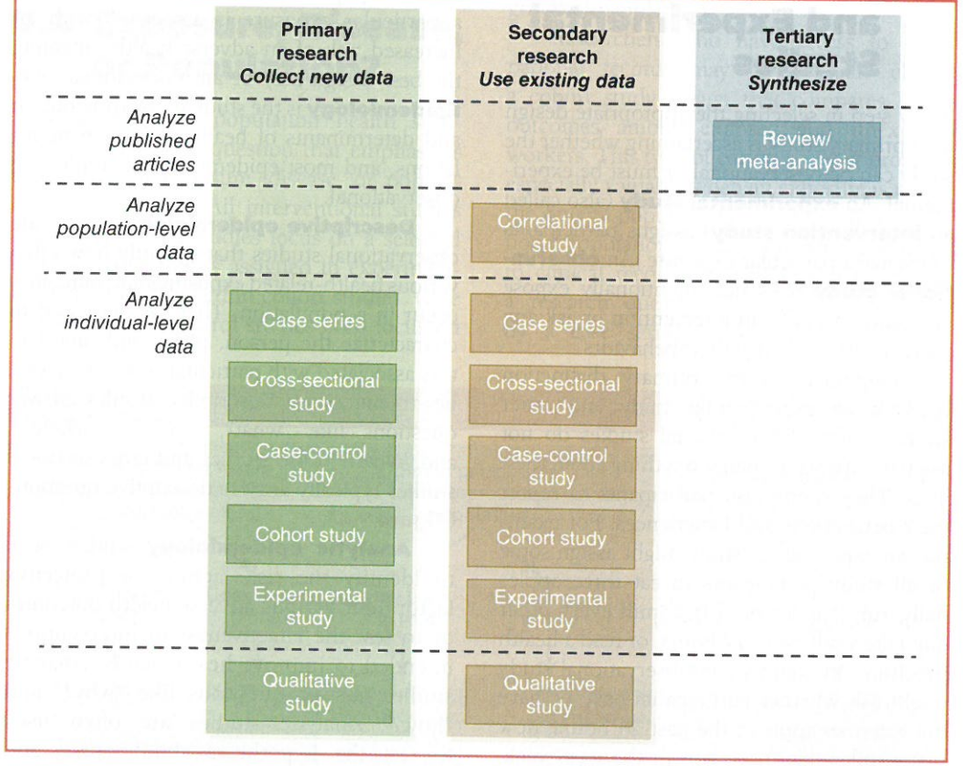
its own critical considerations (**Figure 7-2**). Being aware of these likely challenges at the start of the study design process enables the researcher to make informed decisions about how to proceed.

Primary studies allow researchers to design studies that will answer their preferred research questions. A researcher collecting new data gets to design data collection protocols that are optimal for the study goals and resources, including selecting the source population, the sampling and recruiting methods, and the content and wording of the questionnaire. However, there are also some downsides to collecting new data. It may take months for a rigorous protocol to be developed by the research team and then reviewed by one or more research ethics committees.

Some primary studies can collect data quickly after a protocol is approved, but some require months or years of recruitment and data collection. There is a risk of failure if a sufficient number of participants cannot be recruited. Primary studies can become expensive, especially if there is a per-participant cost for laboratory testing or the use of proprietary survey instruments. The types of data that can be collected and the number of people who can participate in a study might be limited by resource constraints.

Data from any type of primary study may be made available for secondary analysis (**Figure 7-3**). The obvious advantage of secondary studies is that a researcher may be able to move very quickly from the definition of the study question to the analysis of related data. The major disadvantage of analyzing existing data is that the available data files might not include the exact variables of greatest interest to the researcher. When a retrospective review of clinical records is conducted, the files might have incomplete information about patient histories or lack confirmation that particular signs and symptoms were not present in cases. These omissions may require the study question to be revised. The data sets generated from national health surveys and made available to researchers might include a very limited number of questions about any particular topic. Some of the questions a researcher might be most interested in exploring may have already been answered by others who had earlier access to

FIGURE 7-3 Primary, Secondary, and Tertiary Study Approaches



the data. The researcher must identify a valid and accessible source of data, then be prepared to select a study question based on the content of the available data files.

Tertiary studies also allow a researcher to move relatively quickly to the analysis stage, because the data collection process consists of conducting a comprehensive literature review and tracking down the full text of all the potentially relevant articles. Researchers with university affiliations need to check with their institutions' libraries about their policies (and possible fees) for acquiring articles that are not part of their collections. Researchers without university affiliations must consider the costs involved in accessing all of the required articles.

7.3 Observational and Experimental Studies

A key step in selecting the appropriate design for a primary study is ascertaining whether the study can be observational or must be experimental. An **experimental study** (also called an **intervention study**) assigns participants to receive a particular exposure. An **observational study** does not intentionally expose any participants to an intervention or ask any participants to change their behaviors.

Assignment is the primary distinction between an experimental study and other study designs. Observational studies do not ask participants to change anything about their lives. They simply ask participants to report their perceptions and experiences. For example, an experimental study might assign some or all study participants to eat three apples daily, run 2 miles on a treadmill every other day, take a pill every 12 hours, or read a health brochure. In contrast, an observational study might ask whether participants have or have not eaten an apple in the past 24 hours, how many miles they have run in the past week, whether they take any medications routinely, and whether they have seen an ad for a health promotion campaign.

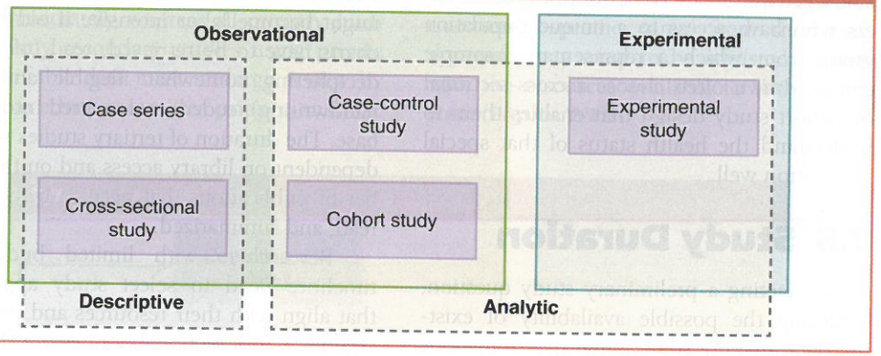
If the goal is to examine whether an intervention is effective, an experimental design is likely to be the only suitable one. Experimental designs are used to test the efficacy of new medications, vaccines, and medical devices. They are also used to test the effectiveness of preventive interventions, diagnostic methods, treatment protocols, and rehabilitative therapies. Most experimental studies randomly assign some participants to an active intervention group and others to a comparison group, so that the results in the two populations can be compared. Experimental studies require careful design and oversight to ensure that they minimize the risks to participants and do not violate other standards for ethical research and practice.

If the goal is to describe the health profile of a population or to examine whether a particular exposure is associated with an increased risk of an adverse health outcome, the best design may be an observational one. **Epidemiology** is the study of the distribution and determinants of health in human populations, and most epidemiological studies are observational.

Descriptive epidemiology studies are observational studies that quantify how often various health-related exposures and outcomes occur in a population. They are also used to characterize the person, place, and time factors associated with particular types of adverse health outcomes. Descriptive studies answer questions like "what?," "who?," "where?," and "when?" Case studies and cross-sectional studies typically answer descriptive questions (**Figure 7-4**).

Analytic epidemiology studies seek to identify the risk factors (or protective factors) for various adverse health outcomes or to test the effectiveness of interventions intended to improve health status. Analytic studies answer questions like "why?" and "how?" Analytic studies are often used to test the hypotheses about causal and protective factors that were generated by

FIGURE 7-4 Epidemiological Study Approaches



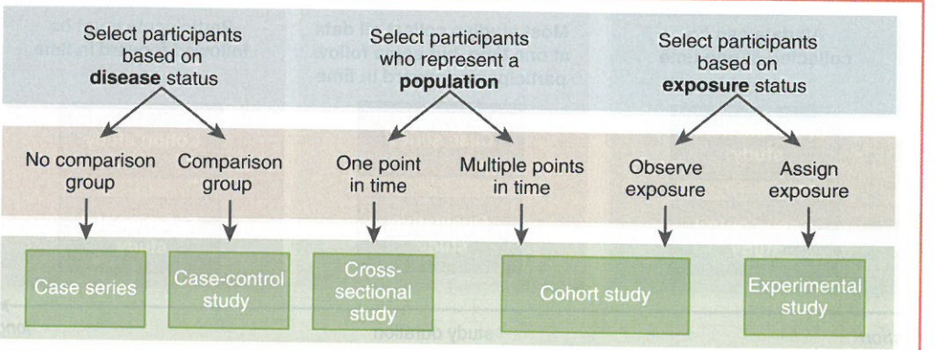
descriptive studies. Case–control, cohort, and experimental studies typically answer analytic questions.

7.4 Exposure, Disease, or Population?

Most clinical and population health studies have a research question that emphasizes one particular exposure, disease, or population (**Figure 7-5**). All interventional studies and many cohort studies focus on a selected exposure, one that is assigned in experimental trials and observed in cohort studies. Case series and case–control studies both focus on "cases," individuals with a particular disease. Cross-sectional studies and some types of cohort studies seek to recruit a study population that is representative of a well-defined group of people.

Researchers who have access to occupational records may preferentially choose a cohort study design that compares health outcomes among exposed and unexposed workers. This type of cohort study approach is especially valuable when an exposure is rare in the general population but high in the occupational population. Researchers who have access to clinical records may preferentially choose a research question that can be answered by a case series or a case–control study.

FIGURE 7-5 Population Selection for Each Study Approach



Those study approaches are particularly valuable when the disease is relatively rare. Researchers who have access to a unique population group from which a representative sample can be drawn often choose a cross-sectional or cohort study design that enables them to understand the health status of that special population well.

7.5 Study Duration

After selecting a preliminary study question, exploring the possible availability of existing data, making some initial determinations about the types of study approaches that might be appropriate, and identifying the special populations or data sources that might be accessible to the researcher, the final decision about which study design to use requires an evaluation of the expected duration and cost of the possible approaches.

The time required for collecting and analyzing data varies from study to study. Some primary studies allow for the collection of all needed data from participants at one point in time. Others require participants to be followed for weeks, months, or even years (**Figure 7-6**). The timeline for a secondary study might be very short if an entire data file and the relevant supporting documentation (such as copies of the questionnaire and codebook) can be downloaded from a website. However, secondary data collection might become labor intensive if old hospital charts have to be retrieved, read (often after deciphering somewhat illegible and faded handwriting), coded, and entered into a database. The duration of tertiary studies is highly dependent on library access and on the number of publications that need to be acquired, read, and summarized.

Researchers with limited budgets or timelines need to select study approaches that align with their resources and time constraints. Cross-sectional studies are the most popular primary study design because they collect data from each participant at one point in time, typically using a simple questionnaire. Case series, case–control studies, and qualitative studies may also allow data to be collected quickly. Cohort and experimental studies require data to be collected from each participant at least twice, first during a baseline examination and then at one or more follow-up times, so they typically take longer than other primary study designs. The fastest option for secondary analyses is accessing a data set that is already clean and ready to use, because compiling, cleaning, and coding new data sets from patient records, client records, and other types of information can take a long time. Population-level data (used during correlational studies) and publicly available, deidentified individual-level data can be analyzed immediately, whereas all primary studies and any secondary studies using private or identifiable data must undergo review by a research ethics committee prior to analysis.

Flexibility is necessary at this stage of the research planning process, because research questions, specific aims, and the study approach may need to be adjusted to align with the resources available to the research team (**Figure 7-7**).

FIGURE 7-6 Time Frame for Primary Data Collection

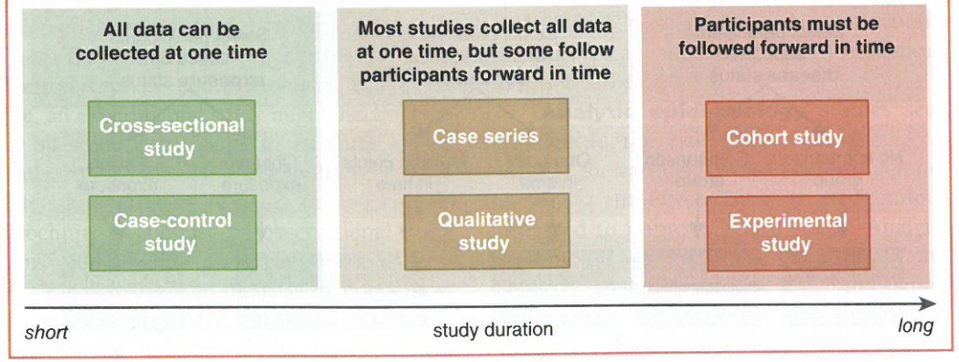


FIGURE 7-7 Flow Diagram for Selecting a Study Approach