

2.4: Experimental Design and rise of statistics in medical research

Introduction

We were in the fifth month after [WHO](#) had declared the [Covid-19 pandemic](#) when I last updated this page. If you followed the news at that time you would know of the appeal from some (including the then-President of the United States) for use of [hydroxychloroquine](#), an anti-parasite drug, as a prophylactic or treatment for active Covid-19 infection (cf. Liu et al 2020). The [FDA](#) as well as other institutions advised against its use, in part because experimental design concerns were raised for early studies (Kupferschmidt, 2020).

We will spend some time later in the semester with experimental design ([Chapter 5](#)), but we can start here. For a number of reasons the **Randomized Control Trial** or RCT — experimental, prospective, double-blind clinical trial with **random selection** of subjects from a **reference population** and **random assignment** of subjects to a treatment group or an appropriate placebo treatment **control group** — is considered the gold standard for producing knowledge (Kaptchuk 2001).

Note:

Experimental control implies researcher imposes conditions to remove possibly confounding effects on the dependent variable — outcome of the experiment. **Placebo**, derived from Latin *placere*, to please (but see Aronson 1999), refers to an inert substance (“sugar pill”), or to a substance with known activity but without effect on the target condition or “wrong indication” (e.g., antibiotics administered for viral infection), given to research subjects in lieu of active treatment. Thus, placebos are examples of treatment controls. The **placebo effect** is improvement of subjects who received the placebo and not active treatment (Pardo-Cabello et al 2022). In contrast, **nocebo effects** are adverse effects attributed to placebo treatment. Under most circumstances, placebo applies to humans only because placebo effects are thought to be product of psychological factors, although mechanisms of action are in dispute. Is sentence necessary for a placebo effect (cf. McMillan 1999)?

Experimental studies imply that the researcher imposed treatments or controls onto subjects. The subjects are followed and outcomes are recorded. Thus, experiments by definition are also **prospective studies** — the outcome is recorded for subjects after some period of time. With a well-designed experiment, the researcher may have evidence to support the claim that, for example, Treatment A causes the outcome.

In contrast, **observational studies** are those in which treatments arise by acts of nature. In both experiments and observational studies, there can be treatment and control groups; the distinction between the types of studies is how assignment of subjects to treatments were affected. Observational studies generally are **retrospective studies** — the outcome has already occurred, the researcher follows up to identify differences among the groups that may account for different outcomes. Examples of observational, retrospective study designs include **cross-sectional** and **case control**; **cohort studies** are prospective studies. Observational studies are discussed further in [Chapter 5.4: Clinical trials](#).

Compared to observational studies, in principle, experiments can establish **cause and effect**. Cause and effect refers to an explanation about relationship between two events or objects. In biology, Ernst Mayr (1904 – 2005) distinguished between two levels of explanation: **proximate** (how) **explanations** and **ultimate** (why) **explanations** (Mayr 1961, cf. Laland et al. 2011). As you know, our mechanism for identifying cause and effect is application of the [Scientific Method \(Chapter 2.5\)](#). Discussions of how to detect cause and effect are provided throughout this book, but emphasized in a few sections ([Chapter 16.2](#) and [16.3](#)).

The principles of good experiments include many steps beyond simply choosing treatments and controls. In [Chapter 5](#) we’ll go into more depth, but I wished to list for you some of the key principles of good experimental design. With respect to human-subject research, the researcher needs to protect against many sources of potential bias.

- **Randomization** of subjects assigned to treatment groups controls for individual differences.
- **Placebos** are a means to establish controls in a study so that effects may be attributed to the **active treatment**.
- **Single-blind** implies that the subject does not know what treatment was given. **Double-blind** implies that not only is the subject unaware of the treatment received, but, crucially, neither does the researcher.
- The double-blind design — neither the patient-subject nor the researchers know who received the placebo or the treatment — controls for subtle **biases**.

The experimenter may influence the outcome of the experiment if knowledge about who received the placebo or the new drug; the subject may respond differently with knowledge that they received the placebo and not the new drug. The key intent in this

experimental design is to avoid systematic error, errors in studies that may occur because of our conscious and unconscious beliefs and biases. Placebos are used as treatments because people (and animals!) sometimes get better (or worse) with or without treatment; thus, to be effective, subjects receiving a new drug must get better more frequently than do subjects on placebo. Importantly, the well-designed placebo allows the researcher to gain insight into the mechanism of action by the new drug.

A case to consider

Consider the following experiment (Diener et al 2006; see also Liu et al 2018): subjects who had several migraines per month were treated with acupuncture, sham-acupuncture, or standard treatments including beta blockers, calcium channel blockers, or antiepileptic drugs. After 26 weeks the reductions in reported migraines was compared. The authors reported that there was no difference in numbers of migraines among patients who received the different therapy treatments. The authors conclude that because acupuncture lacks side-effects that may occur with standard therapies, acupuncture may be a good choice for patients seeking relief from migraine.

Another case to consider

Consider the following example. My dad was diagnosed with lung cancer in his early 80s; his left lung showed many spots when imaged and biopsy confirmed the cancer diagnosis. Surgeons removed half of the lung and after five years he was considered cancer free. Why did he develop cancer in the first place? If you immediately think, “He’s a smoker,” that’s not a good explanation: he last smoked tobacco in his early thirties (latency smoking-lung cancer link is about 20 years, Lipfert et al 2019). Tobacco smoking is not the only environmental trigger for lung cancer. Long term exposure to **radon gas**, a naturally occurring, radioactive noble gas, has been linked to lung cancer (EPA). Cancer of the lung in non-smokers is the seventh leading cause of cancer mortality worldwide (Field and Withers 2012). I grew up on Vashon Island, Washington, in a non-smoking home environment. Radon levels on Vashon Island and other areas around Puget Sound are low (source: [Washington State Department of Health](#)). Vashon Island is rural, but, as it turns out, within range of a larger copper smelter located in nearby Ruston (Fig. 2.4.1; my home was a 17-km distance from the smelter). The smelter was last in operation in 1986 and was torn down in 1993 ([EPA publication number 910R94001](#)). The smelter stack rose more than 500 feet, dispersing smoke laden with heavy metals, notably **arsenic** and **lead**, into the air (Bromenshenk et al. 1985). Over the smelter’s 68 years of service, winds carried away the smoke to my island and to other areas known now as the “[Ruston-Vashon Island Exposure Pathway](#)” (Kalman et al 1990). Thus, tens of thousands of people were (and continue to be) exposed to the heavy metals deposited into the soils, forming a distinct treatment group (Milham & Strong 1974; Kalman et al 1990; EPA 2000). Is arsenic exposure a plausible mechanism for lung cancer? Workers exposed to arsenic have higher rates of lung cancer (Sullivan 2007, Wei et al 2019). Cultured lung cells exposed to arsenic associated with changes in gene expression (Clancy et al 2012). Coincidentally, two of the family dogs developed and died of cancer, as did one female goat. Perhaps my dad’s lung cancer was attributed to long exposure to arsenic (his blood readings for arsenic were in the range of 11 ug/L).

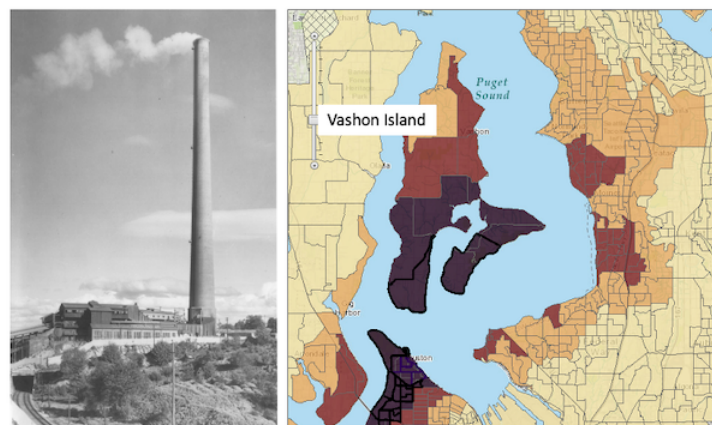


Figure 2.4.1: Left: ASARCO smelter, Ruston, Washington, image from [Department of Ecology, State of Washington](#). Direction of smoke from the stack is north, toward Vashon Island. Right: Heat map of arsenic and lead affected areas. image from [kingcounty.gov](#). Darker regions correspond to heavier arsenic and lead contamination of soils.

If this scenario seems plausible, I hope you immediately recognize it as a case of **conformation bias** (see [Chapter 2.6](#)). Putting aside for a moment the different arsenic species, each with different LD₅₀ (the lethal dose needed to kill half the population — see [Chapter 20.10](#)), the difficulty ascribing arsenic as a causal agent for my Dad’s cancer is that many other exposures happened simultaneously. For example, indoor carpets are a primary source of several volatile organic compounds (Haines et al 2020). Prior

to 1980, carpets may have contained formaldehyde and other known carcinogenic agents. My dad also commuted by car between work and home for decades, this during the early years of the Clean Air Act of Environmental Protection Agency of the United States (it wasn't until 1981 that new cars met EPA emission standards: [Clean Air Act timeline here](#)). Thus, all commuters including my Dad were exposed to gasoline combustion emissions, many known to be carcinogenic (Parent et al 2007). Moreover, a limited study by Public Health of Seattle and King County (2001) found that rates of cancer on Vashon between 1980 and 1988 were similar to those in other areas of King County.

Note:

While we “know” tobacco cigarette smoking increases lung cancer risk, and many experiments with animal models convincingly show the link (e.g., Hutt et al 2005), no experiment in the strict sense, i.e., a prospective, randomized control trial, has ever been conducted (hint: it would be unethical; see discussion in Allmark and Tod 2016). Instead, the cumulated evidence from observational studies on exposures of different populations over the years overwhelmingly points to smoking as a leading cause of lung and other cancers.

Questions

1. Was my Dad's lung cancer attributable to his 40-plus year of exposure to soil arsenic (he's a non-smoker)? How should we approach this question?
2. In Diener et al (2006), the authors concluded that because acupuncture lacks side-effects that may occur with standard therapies, acupuncture may be a good choice for patients seeking relief from migraine. Do you agree with the authors?
3. Ethical standards evolve with time. An ongoing debate in research is whether and how placebos are to be used in human subjects research. Placebos are a means to establish controls in a study so that effects may be attributed to the active treatment. The “gold standard” of clinical trials is considered to be the randomized double-blind design — neither the patient-subject nor the researchers know who receives the placebo or the treatment. Following review of the WHO report on *Use of Placebos in Vaccine Trials*, pick one study and evaluate whether or not the decision to use placebos was warranted in your opinion.
4. I searched PUBMED for “double-blind” by decade and found the following results (August 2018) (Table 2.4.1). Open R and/or R Commander and create two variables, then generate a scatter plot. Describe the shape of the relationship between number of publications citing “double-blind” and time (e.g., 1950 – 1959, 1960 – 1969, and so on).

Table 2.4.1. PUBMED results for “double-blind” by decade.

Decade	Publications
1950	60
1960	995
1970	7184
1980	24737
1990	39643
2000	53965
2010	69265

5. Here's one way to enter this data into R. At the R prompt (or in the R Script window of R Commander), create two variables, `Decade` and `Pubs` `Decade <- c(seq(1950, 2020, by=10))`
`Pubs <- c(59, 995, 7161, 24728, 39670, 54011, 57043)` Make an XY scatter plot `plot(Decade, Pubs)`
6. Repeat the PUBMED search as above but search for “placebo”. Make a table like the one above and provide a scatterplot of your results.
7. Is the concept of a placebo relevant if the subjects in your experiment are yeast cells, not humans?
 - Similarly, if your subjects are yeast cells, how does the concept of performing experiments “blind” apply?
8. Ethical standards change with time. An ongoing debate in research is whether and how placebos are to be used in human subjects research.
 - If placebos are so important, why is their use a concern in clinical trials?

- Following review of the WHO report on Use of Placebos in Vaccine Trials (see Readings at the end of this chapter), pick one study and evaluate whether or not the decision to use placebos was warranted in your opinion.

R notes for question 5:

- `<-` is an assignment operator (`assignOP`); everything to the right of `<-` is assigned to the object named to the left of the `<-` operator. You can instead use `=` in place of `<-` , but because `=` is also used in other contexts besides assignment, a quick look at blogs by data scientists will find a preference to use `<-` for clarity and consistency.
- `c()` “combines” arguments into a vector.
- `seq()` is used to generate a sequence of numbers between a lower and an upper limit; if `by = n` is included, the sequence will be increased by the value n. If omitted, then the sequence is increased by 1.

This page titled [2.4: Experimental Design and rise of statistics in medical research](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Michael R Dohm](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.