1

The Basis of Statistical Reasoning in Medicine

1.1 What Is Statistical Reasoning?

“Well, do you believe the article, or not?” your impatient colleague demands yet again.

A 34-year-old woman is brought to the emergency room by her anxious husband. You’re her doctor.

“She was fine until four hours ago!” her frightened husband blurts out. “She suddenly collapsed!”

You learn that your new patient, a cigarette smoker, didn’t lose consciousness or convulse. She does use birth control pills. Your exam reveals a frightened but well-groomed adult female who is unable to communicate and can’t control the right side of her body. Finishing your workup, you correctly diagnose an acute stroke and orchestrate her emergency care.

“Are you going to give her tPA* to help reduce the size of the stroke?” asks your colleague, who, observing the new case, wanders over to where you are now writing up the patient’s record.

“No,” you reply. “Current protocols say to give tPA within three hours of symptom onset.” Checking your watch, you add, “It’s been almost four and a half hours now according to the husband. Giving it now would risk a major bleeding event.”

“Hey!” your colleague retorts. “Didn’t you see the article this month describing the latest research experience? It found that its safe to give tPA for more than three hours after the stroke. Actually,” she adds, rummaging through a collection of papers, “I have it…right…Ah! Here it is.” Your stomach squirms restlessly as she plops the manuscript on your desk right where you’re writing the patient’s orders.

“It’s in a good journal,” she persists. Turning to the relevant section, she says, “See? This is what I was talking about.”

“I didn’t see this,” you confess thinking Current tPA practice says that giving it now risks a major bleed. Yet if it could help…

* Tissue plasminogen activator. This therapy is believed to reduce the magnitude of a stroke.
“Right here!” your colleague interjects, pointing to the abstract. “The study says the risk of a bleed is low and that patients do better with tPA administration even out to five hours after stroke symptoms.”

Struggling to integrate the new material, you ask, “How many patients in the study?”

“A lot for a stroke trial,” she responds at once. “It’s a large subgroup analysis from a randomized, double-blind controlled clinical study with a combined primary endpoint and multiple secondaries. While the primary endpoint just missed stat significance, the secondary endpoint of reduced hospitalization…”.

On and on your colleague goes, spewing out the study details. The clock is ticking as you struggle to decide if you believe the manuscript. Is the analysis sound? you wonder. Does is apply to my patient? Can I believe the conclusion? Is it too late to apply it’s results…

“Doctor!” a nurse calls out, jolting you from your thoughts. “Just picked up a call from the paramedics. A motor vehicle accident with head trauma coming in. Any instructions?”

“Time’s a’wastin,” your colleague reminds you.

Statistical reasoning in medicine is, at its heart, the process by which we determine if healthcare research results identified in one sample apply to others. For the emergency room physician in the previous example, the urgent question is whether the results can be generalized from the new manuscript to the stricken patient. Since this generalization is not automatic and commonly inappropriate, we must know when to generalize sample results versus when not to.

Before we delve into the statistical issues, we should first acknowledge that physicians have a built-in difficulty with these concepts. It’s not our fault—we were engineered this way!

1.1.2 Physicians and the Patient Perspective

Physician training focuses on the individual patient to the exclusion of the population. We may call this the “single-patient perspective.” The single patient perspective is inculcated in medical students from their first year introductory courses. The detailed instruction in anatomy, histology, biochemistry, and physiology we received was not so much to enhance our understanding of the constitution of the population, but to deepen our understanding of the structure of the individual. It was a discussion about facts. Exams were on facts, not on variability.

As medical students, we are trained to develop thorough and exhaustive differential diagnoses, compiling a complete list of possible explanations for the symptoms and signs the individual patient presented, with little thought given to the frequency of the disease in the population. In my medical education, a seemingly disproportionate emphasis was placed on rare diseases (e.g., Whipple’s disease, or Tsushugamuchi fever). This exposure trained us to identify the one unexpected patient who might have that unusual disease. In essence we were taught to bring “the rare to bear” in developing a patient’s diagnosis.