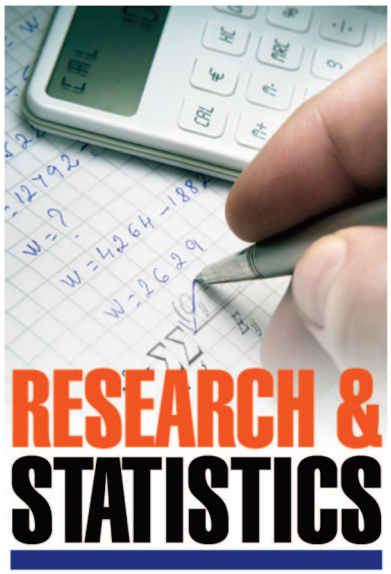
Research and Statistics : Demystifying Type I and Type II Errors

Jacky M. Jennings and Erica Sibinga

Pediatrics in Review 2010;31;209
DOI: 10.1542/pir.31-5-209



Author Disclosure
Drs Jennings and Sibinga have disclosed no financial relationships relevant to this article. This commentary does not contain a discussion of an unapproved/investigative use of a commercial product/device.

Demystifying Type I and Type II Errors

Jacky M. Jennings, PhD, MPH,* Erica Sibinga, MD*

Introduction

Early in the fall of 2009 a colleague evaluated a 22-month-old boy. The patient was afebrile when examined, but his mother brought him in because at 1:30 AM, he had developed a temperature of 38.4°C and appeared uncomfortable, experiencing a number of crying spells. She gave acetaminophen at 2 AM and again at 6 PM. At the time of the visit, the boy was comfortable and smiling. Your colleague had seen the patient's 5-year-old sister last week and had diagnosed group A Streptococcus (GAS) pharyngitis. Because of the recent household case of GAS pharyngitis, your colleague obtained a rapid test for the boy, which yielded negative results, and a backup throat culture. She treated the 22-month-old with oseltamivir in case he had a novel influenza A (H1N1) viral infection. You ask your colleague why she did not use the rapid influenza diagnostic test (RIDT) to detect influenza viral nucleoprotein antigen in the patient's respiratory specimens before treating. She says that she looked at the package insert, which stated that the sensitivity of RIDTs for detecting novel influenza A (H1N1) virus infections ranged from 10% to 51% (1)(2) and that she was worried about a type II error.

Type I and Type II Errors Defined

When evaluating for infection, a type II error occurs when a diagnostic test result indicates that an individual is not infected and the individual truly is infected. (3) A type II error also is called a false-negative (Table 1). A type I error, on the other hand, occurs when a diagnostic test result is positive, indicating that the individual is infected, when, in fact, the individual is not infected. (3) A type I error also is called a false-positive (Table 1). In either case, a type II or type I error leads to an erroneous conclusion. Although infection is used as an example, these errors are applicable to a broad range of clinical situations.

Examples of Type I and Type II Errors

As highlighted in the example opening this article, the sensitivity of RIDTs for detecting novel influenza A (H1N1) can be as low as 10% (although there have been few studies to date). (1)(2) The low sensitivity results in a high likelihood of a false-negative result. Thus, any negative test result would have a high likelihood of being an erroneous result, that is, a type II error. In the example, the clinician chose not to conduct the RIDT for detecting novel influenza A (H1N1), in part, because of the likelihood of a type II error, choosing instead to treat the 22-month-old patient presumptively because he has been classified as a high-risk case due to his age. (4)

The likelihood of a type II error, or false-negative, is the same reason that a backup throat culture was performed for GAS. Although the rapid test for GAS typically has a sensitivity of 80% or higher (and, thus, the likelihood of a false-negative is less than that associated with the RIDT), a backup culture is performed to ensure that the negative result is, indeed, a true negative.

In testing for GAS and influenza A (H1N1), the clinician was less concerned about the likelihood of a false-positive or a type I error. She was more concerned about missing a true infection, which might lead to

*Department of Pediatrics, Johns Hopkins School of Medicine, Baltimore, Md.

Table 1. True Condition Versus Diagnostic Test Results and Type I and Type II Errors

		True Condition	
		Infected	Not-Infected
Diagnostic Test Results	Positive	True positive	False positive Type I error
	Negative	False negative Type II error	True negative

complications of either acute rheumatic fever in the case of GAS or significant morbidity or complications due to the patient's young age in the case of novel influenza A (H1N1).

The opposite is true when testing for human immunodeficiency virus (HIV). When testing for HIV, there is great concern for making a type I error, that is, telling patients that they are infected when, in fact, they are not. Because HIV is a chronic disease that is not curable, a false-positive result can be extremely troubling to a patient. In contrast to the GAS diagnostic procedure, in which every negative test result prompts a confirmatory culture, every positive HIV result from a Western blot receives a confirmatory enzyme linked immunosorbent assay test. This second procedure is performed to reduce type I errors. Because an undiagnosed HIV infection will result in no treatment and may increase transmission risk, clinicians also have great concerns about type II errors, in which the diagnosis is missed.

Type I and Type II Errors in Research

When conducting research, the equivalent of a type II error in statistical terms is failing to reject the null hypothesis when it should have been rejected (Table 2). (3)(5) For example, in the case of a randomized, controlled trial (RCT) designed to test the difference between therapy A and therapy B, a type II error would result in concluding that the two therapies were not different from each other when, in fact, they were different. The probability of a type II error often is denoted with the Greek letter beta. The probability that a study will conclude correctly that two therapies are different when they are, indeed, different equals 1−beta or the power of a study.

In statistical terms, the equivalent of a type I error is rejecting the null hypothesis when it should not have been rejected (Table 2). (3)(5) A type I error in an RCT results in the conclusion that there was a difference between therapy A and therapy B when, in fact, the two therapies were not different. The probability of a type I error is often denoted with the Greek letter alpha. The alpha is probably most familiar as the *P* value of a study, which, by common convention, has been set at a significant value of 0.05. Setting the *P* value of a study at 0.05 means that the researchers are willing to accept a 5% probability of a type I error or a 5% probability that the results showing a difference could have occurred by chance alone.

Table 2. True Condition Versus Randomized, Controlled Trial (RCT) Study Results and Type I and Type II Errors

		True Condition	
		Therapies are different	Therapies are not different
RCT Study Results	Therapies are different	Correct decision	False decision Type I error (Probability=α)
	Therapies are not different	False decision Type II error (Probability=β)	Correct decision

Summary

- For diagnostic test results and for research study results, type I and type II errors are an expression of the possibility that the test result or the research conclusion does not reflect the true condition. In both instances, it is important to have an understanding of the likelihood of these types of errors.

References

1. Faix DJ, Sherman SS, Waterman SH. Rapid-test sensitivity for novel swine-origin influenza A (H1N1) virus in humans. *N Engl J Med*. 2009;361:728–729
2. Ginocchio CC, Zhang F, Manji R, et al. Evaluation of multiple test methods for the detection of the novel 2009 influenza A (H1N1) during the New York City outbreak. *J Clin Virol*. 2009;45:191–205
3. Gordis L. *Epidemiology*. Philadelphia, Pa: WB Saunders Company; 1996
4. Centers for Disease Control and Prevention. *Updated Interim Recommendations for the Use of Antiviral Medications in the Treatment and Prevention of Influenza for the 2009–2010 Season*. 2009. Accessed February 2010 at: http://www.cdc.gov/H1N1flu/recommendations.htm
5. Last JM. *A Dictionary of Epidemiology*. 3rd ed. Oxford, United Kingdom: Oxford University Press; 1995

Research and Statistics : Demystifying Type I and Type II Errors

Jacky M. Jennings and Erica Sibinga

Pediatrics in Review 2010;31;209
DOI: 10.1542/pir.31-5-209