Sensitivity and specificity: alien edition

A light-hearted look at statistics

Roger Suss MD CCFP(EM) FCFP

Our screenplay begins with a meeting of 3 members of the Murspian scientific expeditionary force. Their mission: to collect live specimens of pregnant Homo sapiens *and bring them back to their home planet for further study.*

Zoltan: Captain, since Org "enthusiastically landed" our ship, I have successfully repaired all our equipment except for the tricorders. While they are highly reliable in identifying which humans are pregnant, their compromised power supply won't last until the planned abduction date. We are, therefore, evaluating other potential identification methods. We will use the tricorders as the gold standard to measure the effectiveness of the other methods. Then we can use the other methods on abduction day.

Org: And I have been working on perfecting the identification tools.

Captain: I know. I keep hearing about your "alien abductions" every other day. One minute humans are lying in bed and the next you paralyse them and "probe" them to determine whether they are pregnant. Can't you be more discreet?

Org: That wasn't my fault. And in any case, lately I have been working on less invasive methods. I call my new technique "observation." I believe that observation can separate pregnant from non-pregnant humans. My strategy is to use a sensitive test to identify who might be pregnant and then to use a specific test to confirm who really is pregnant. So I started with a highly reproductive population on a Pacific island. I noticed that having breasts is 100% sensitive as a test for pregnancy. In other words, out of the 25 pregnant humans on the island, all of them had breasts. Here is a chart that shows my findings ....

	Pregnant	Not pregnant	
Breasts	A = 25	B = 25	A/(A+B) = 50% Positive predictive value
No breasts	C = 0	D = 50	D/(C+D) = 100% Negative predictive value
	A/(A + C) = 100% Sensitivity	D/(B + D) = 66% Specificity	

Captain: Ummm ... Org ... I also see that only half of the breasted humans were pregnant, so the positive predictive value is only 50%. We can't afford to abduct twice as many humans as needed. We will have to feed them all the way home.

Org: Ahhh, but of the humans without breasts, not one of them was pregnant, so the negative predictive value is 100%.

Zoltan: Org's observational test of looking for breasts is sensitive, so it is good for ruling out pregnancy. May I suggest Org's nose as a mnemonic: *SNout, a highly sensitive test is good for ruling out*. But it is not specific enough to rule in pregnancy. That takes us to Org's eyes for a second mnemonic: *SPin, a highly specific test is good for ruling in*.

Org: Wait! That is only step 1. Now I will use a specific test to rule in pregnancy on the remaining humans as follows ....

	Pregnant	Not pregnant	
Extra head between the legs	A = 1	B = 0	A/(A + B) = 100% Positive predictive value
Only 1 head	C = 24	D = 25	D/(C + D) = 50% Negative predictive value
	A/(A + C) = 4% Sensitivity	D/(B + D) = 100% Specificity	

Captain: That is less than helpful, Org.

Org: But it has stellar specificity! Every human with 2 heads was pregnant!

Captain: As Zoltan says, SPin—your test has high specificity, so a positive result rules in pregnancy. SNout—but your test has such lousy sensitivity that a negative test is almost useless to rule it out. Your test missed 96% of the pregnant humans. I think I will have to give your snout a spin, Org.

Zoltan: I have a suggestion. I have noticed that pregnant humans tend to be round in the middle.

Org: Yes, you are right!! They do have an unusually high waist-to-hip ratio!

Zoltan: These are the results you would get combining those observations ....

	Pregnant	Not pregnant	
Round with breasts	A = 15	B = 5	A/(A+B) = 75% Positive predictive value
Either not round or no breasts	C = 10	D = 70	D/(C+D) = 88% Negative predictive value
	A/(A + C) = 60% Sensitivity	D/(B + D) = 93% Specificity	

Captain: So the test is only moderately sensitive, and a negative result won't really rule it out. A human could easily be pregnant and not be round with breasts.

Org: But the test is fairly specific, so a positive result will rule it in reasonably well. And with a positive predictive value of 75%, three-quarters of the humans who are both round and have breasts will be pregnant. This way we will be abducting fewer non-pregnant humans and wasting less food.

Captain: That is better, but it still seems to be wasting a lot of food.

Zoltan: The positive predictive value will vary a great deal depending on how common the condition is in the population. It might be more useful to look at *likelihood ratios*—LRs. Then we can take the pregnancy rate in any given population, calculate the odds of being pregnant before our test, multiply that by the LR of our test, and determine how likely it is that a particular human is actually pregnant.

Captain: Show me how that would work in the island population.

Zoltan: We can calculate the LR for a positive test result on our "round and breasts" test as LR+ = sensitivity/(1-specificity)=0.6/(1-0.93)=9. A good test will have an LR+ of 10 or more, so this is mediocre.

Captain: Because we know that the odds of any given islander being pregnant were 1:3 before the test—

Org: I thought that the pregnancy rate was 25%!

Captain: Yes, Org. That is the *probability* of pregnancy in any given human in this population. Odds compares group 1 with group 2, while probability compares group 1 with the total. So odds of 1:3 equal a probability of 1/(3+1) or 25%. Expressed as a fraction, this is an *odds ratio* of 1:3.

Org: I wish I had paid more attention to mathematics as a juvenile.

Zoltan: So we multiply our pretest odds ratio of 1:3 times an LR+ of 9 and we get posttest odds ratio of 9:3 … or 3:1. In other words, an islander who is round and breasted is 3 times as likely to be pregnant as not … or has a 75% posttest probability of being pregnant.

Captain: So, is that the same as the positive predictive value of 75% that we calculated before?

Zoltan: Yes. But now we can use the LR+ to apply to other populations. If we can find a population that is usually made up of 80% pregnant humans …

Captain: Like some hospital wards.

Zoltan: Exactly. Then the pretest odds will be 4:1. Multiply by an LR+ of 9, and you get posttest odds of 36:1. And the probability that a human on the hypothesized hospital ward who is round and breasted will, in fact, be pregnant is 36:37 or 97%.

Org: Is there a negative LR?

Zoltan: Yes. LR- = (1-sensitivity)/specificity. You again multiply the LR by the pretest odds to get posttest odds. A useful negative test for ruling out a condition would have an LR of 0.1 or less.

Captain: So we either need a better test or we need to find a population with a very high pregnancy rate.

Org: Did you know that the world's smallest country has no army? It is conveniently surrounded by a wall that could prevent escape if, for example, aliens attacked one night.

Captain: It sounds promising. What is it called?

Org: Vatican City.

Will Org convince the Captain to abduct a lot of obese nuns? Will Zoltan convince the nuns to abduct Org? How much gynecomastia does it take for obese monks to be considered pregnant? The answers to these questions are all out there in … the twilight zone.

Dr Suss *is an Assistant Professor in the Department of Family Medicine at the University of Manitoba in Winnipeg.*

Competing interests
None declared

Correspondence to: **Dr Roger Suss**, *400 Tache Ave, 5th Floor, Winnipeg, MB R2H 3E1; telephone 204 237-2054; e-mail* **RSUSS@sbgh.mb.ca**

Hypothesis is a new quarterly series in *Canadian Family Physician,* coordinated by the Section of Researchers of the College of Family Physicians of Canada. The goal is to explore clinically relevant research concepts for all *CFP* readers. Submissions are invited from researchers and nonresearchers. Please send any ideas or submissions to Dr Bo Miedema at **bo.miedema@rvh.nb.ca**.