Taylor & Francis
Taylor & Francis Group

The Journal of Maternal-Fetal & Neonatal Medicine

ISSN: 1476-7058 (Print) 1476-4954 (Online) Journal homepage: informahealthcare.com/journals/ijmf20

Why 99% may not be as good as you think it is: limitations of screening for rare diseases

Monica A. Lutgendorf & Katie A. Stoll

To cite this article: Monica A. Lutgendorf & Katie A. Stoll (2016) Why 99% may not be as good as you think it is: limitations of screening for rare diseases, The Journal of Maternal-Fetal & Neonatal Medicine, 29:7, 1187-1189, DOI: 10.3109/14767058.2015.1039977

To link to this article: https://doi.org/10.3109/14767058.2015.1039977

Published online: 16 Jul 2015.

THE JOURNAL OF MATERNAL-FETAL & NEONATAL MEDICINE

http://informahealthcare.com/jmf
ISSN: 1476-7058 (print), 1476-4954 (electronic)

J Matern Fetal Neonatal Med, 2016; 29(7): 1187–1189
DOI: 10.3109/14767058.2015.1039977

informa healthcare

EDITORIAL

Why 99% may not be as good as you think it is: limitations of screening for rare diseases

Monica A. Lutgendorf[1] and Katie A. Stoll[2]

[1]Department of Obstetrics and Gynecology, Division of Maternal Fetal Medicine, Naval Medical Center San Diego, CA, USA and [2]Genetic Support Foundation, Olympia, WA, USA

Screening recommendations

Currently all pregnant women should be offered the option of invasive diagnostic testing or screening for aneuploidy early in their pregnancy [1]. Available screening options include serum screening for aneuploidy, and non-invasive prenatal testing (NIPT) which can be used to screen high risk women. NIPT is most commonly used to screen women who are at increased risk for fetal aneuploidy (either based on maternal age, ultrasound findings or other risk factors); however, the use of NIPT in the lower-risk population is expected to increase.

Screening involves the testing of asymptomatic, apparently well individuals (in this case apparently well fetuses). Specific criteria must be met prior to implementation of a screening program: the disease must be clinically important, the latent period must be long enough to allow for intervention, a treatment or intervention must be available, the test must be reliable and valid, acceptable to the population being screened and the disease must be relatively prevalent in the population [2].

Serum screening for aneuploidy has a sensitivity of 69–96% depending on the testing regimen chosen [1,3,4]. For sequential screening, sensitivity is 90% for Down syndrome and specificity is 95% when the false positive rate is set at 5% [3]. In comparison, NIPT has reported sensitivity of >98% for Down syndrome and specificity >99% for Down syndrome [4] (Table 1).

Test performance characteristics

How well does a particular test perform? The following test characteristics are used to define test performance. Sensitivity is the ability to identify those with the disease, or the probability that an individual who has the condition will have a positive test. Specificity is the ability to identify those without disease, or the probability that an individual without the condition will have a negative test. These values are related to intrinsic test characteristics.

Predictive values are what clinicians need to know. The positive predictive value (PPV) is the probability that those who have a positive test have the condition, and the negative predictive value (NPV) is the probability that those who have a negative test are without condition. Predictive values are influenced by disease prevalence. Even an excellent test when used in a low-prevalence population will have a poor PPV. In contrast, with a good screening test the NPV will be high when the incidence of a disease is low.

To illustrate this point, consider the familiar 2×2 table of test performance (Table 2), where true positive results are represented in cell ''a'', false positive results are represented in cell ''b'', false negative results are represented in cell ''c'' and true negative results are represented in cell ''d''. Good screening tests will have most results in the true positive and true negative cells.

To calculate test sensitivity (or ability to identify those with disease), the true positives (a) are divided by all those with the disease (a + c). Test specificity (or ability to identify those without disease) is calculated by dividing the true negatives (d) by all those without the disease (b + d). Notice that these calculations are carried out vertically using the information in Table 2. In comparison, calculation of the PPV (the probability that those with a positive test have the disease) is carried out horizontally by dividing the true positives (a) by all those with a positive test result (a + b). The NPV (the probability that those with a negative test are without disease) is calculated by dividing the true negatives (d) by all those with a negative test (c + d). These calculations are carried out horizontally using the same information in Table 2.

Consider the following for a 39-year-old woman, with a risk trisomy 21 of 1:100 at 16 weeks gestation [5], with a total population of 100 000 women, test sensitivity of 99.4%, and

Address for correspondence: Monica A. Lutgendorf, Department of Obstetrics and Gynecology, Division of Maternal Fetal Medicine, Naval Medical Center San Diego, 34800 Bob Wilson Drive, San Diego, CA 92134, USA. E-mail: malutgendorf@gmail.com

test specificity of 99.9% (Table 3). Both sensitivity and specificity are high (>99%). The NPV is also >99%, and the PPV is 91% (95%CI 89–93%), also high, but not 99%. Now consider a 25-year-old woman, with a risk of Down syndrome of ~1:1,000 [5] (Table 4). Note the PPV is only 50% (95%CI 43–57%). This is equivalent to flipping a coin.

Although the NPV remains high in both cases, the PPV varies tremendously between the 25 and 39 year old populations, even with high sensitivity and specificity in both cases. This variation in the PPV is explained by the differences in the prevalence of Down syndrome based on maternal age (Figure 1).

Table 1. Detection rates and false positive rates for trisomy 21 screening tests.

Test	Detection rate (%) (sensitivity)	False positive rate
First trimester screen [1]		
NT alone	64–70%	5%
NT + serum screen	82–87%	5%
Second trimester screen [1]		
Triple screen	69%	5%
Quad screen	81%	5%
First and second trimester screen [1]		
Integrated (with NT)	94–96%	5%
Stepwise sequential	95%	5%
Contingent sequential	88–94%	5%
NIPT [4]	>99%	1–2%

Table 2. Test performance characteristics.

	Truth		
	Disease (aneuploid)	No disease (euploid)	
Test result			
POSITIVE	True positive (a)	False positive (b)	a + b
NEGATIVE	False negative (c)	True negative (d)	c + d
	a + c	b + d	

Sensitivity = a/(a + c), specificity = d/(b + d), PPV = a/(a + b), NPV = d/(c + d).

Accuracy?

Companies also claim ''high accuracy'' when describing NIPT, but the term *accuracy* is also often misunderstood by providers and patients. Accuracy describes the proportion of all tests that was correct. Given that the vast majority of pregnancies are not affected with aneuploidy and will correctly be ''screen negative'', NIPT can be described as highly accurate. However, accuracy of NIPT should not be used to explain the probability that a positive result is a true positive. In fact, if screening is applied to a sufficiently rare condition, the PPV may be low even when accuracy is >99%.

With an increasing number of publications demonstrating the use of NIPT in lower risk women, and for conditions which have a far lower prevalence than Down syndrome, clinicians need to understand these principles. Although the calculations are straightforward, physicians have been noted to have difficulty understanding diagnostics and PPV [6]. Additionally, NIPT is aggressively marketed to patients and physicians alike. This has led to increased patient demand, and a recent study of 356 high-risk patients showed that 22 (6.2%) had abortions without confirmatory karyotyping [7], suggesting that patients may fail to recognize the possibility that the NIPT test may be a false positive.

NIPT is a screening test, which may be most useful for its NPV. Because aneuploidy is uncommon, the NPV will be high and because the sensitivity of the testing is high, false negative results are expected to be rare events. However, in low-prevalence populations, the PPV will be unacceptably low, and warrant additional testing. Clinicians and patients must not make clinical decisions regarding a pregnancy based

Table 3. Test performance characteristics in a 39-year-old woman.

	Trisomy 21		
	Present ($n = 1000$)	Absent ($n = 99\,000$)	
Test result			
POSITIVE	True positive 994	False positive 99	PPV = 994/1093 = 0.91
NEGATIVE	False negative 6	True negative 98 901	NPV = 98 901/98 907 = 1.00
	Sensitivity = 994/1000 = 0.99	Specificity = 98 010/99 000 = 0.99	

Prevalence of trisomy 21 in a 39-year-old woman at 16 weeks gestation = 1:100 [5]. PPV, positive predictive value; NPV, negative predictive value.

Table 4. Test performance characteristics in a 25-year-old woman.

	Trisomy 21		
	Present ($n = 100$)	Absent ($n = 99\,900$)	
Test result			
POSITIVE	True positive 99	False positive 100	PPV = 99/199 = 0.50
NEGATIVE	False negative 1	True negative 99 800	NPV = 99 800/99 801 = 1.00
	Sensitivity = 99/100 = 0.99	Specificity = 99 800/99 900 = 1.00	

Prevalence of trisomy 21 in a 25-year-old woman at 16 weeks gestation = 1:1000 [5]. PPV, positive predictive value, NPV, negative predictive value.

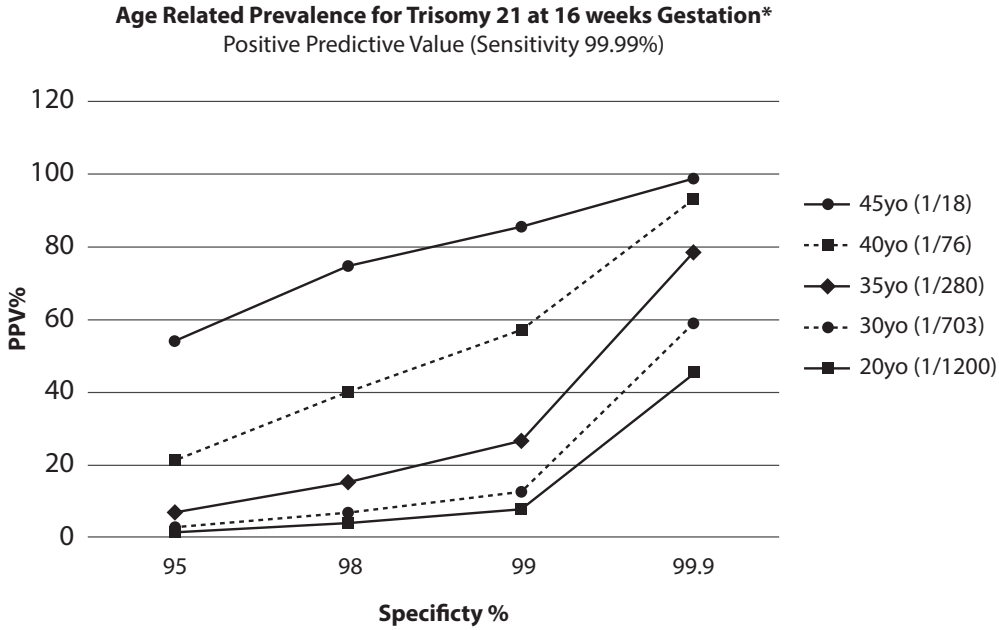

Figure 1. PPV for trisomy 21. The PPV for trisomy 21 varies based on the prevalence of the condition, and test specificity. With sensitivity set at 99.99%, at a given specificity, the PPV is higher with a higher prevalence of trisomy 21. *Data from Snijders et al. [5].

on these screening tests. Diagnostic tests, such as amniocentesis with karyotyping, are required for definitive diagnosis.

The rapid introduction of these tests into clinical use as well as direct-to-consumer marketing has resulted in increased demand without full understanding of test limitations and implications. More information is needed about how NIPT performs in clinical practice and it is imperative that providers understand the PPV of these tests. Additionally, NIPT is expensive ($800–$2000) compared to standard screening ($200). Profits are realized by private testing companies. Some would view the rapid proliferation of NIPT as contributing to the ''medical-industrial complex'' [8] without clear benefit to low risk patients, and even the potential for harm. With continued scientific advances in prenatal screening, we must fully understand testing limitations to educate and support patients to make informed screening decisions. We must first do no harm, and always strive to put ''the interests of the public before those of its stockholders'' [8].

Disclaimer

The views expressed in this article are those of the author(s) and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, or the United States Government.

One of the authors is a military service member. This work was prepared as part of their official duties. Title 17 U.S.C. 105 provides that ''Copyright protection under this title is not available for any work of the United States Government''. Title 17 U.S.C. 101 defines a United States Government work as a work prepared by a military service member or employee of the United States Government as part of that person's official duties.

Declaration of interest

The authors report no declarations of interest.

References

1. Screening for fetal chromosome abnormalities. ACOG Practice Bulletin No. 77. American College of Obstetricians and Gynecologists. Obstet Gynecol 2007;109:217–28.
2. Wilson JG, Jungner G. Principles and practice of screening for disease. Public Health Paper No. 34. Geneva: World Health Organization; 1968.
3. Malone FD, Canick JA, Ball RH, et al. First- and Second-Trimester Evaluation of Risk (FASTER) Research Consortium. First-trimester or second-trimester screening, or both, for Down's syndrome. N Engl J Med 2005;353:2001–11.
4. Lutgendorf MA, Stoll KS, Knutzen DM, Foglia LM. Noninvasive prenatal testing: limitations and unanswered questions. Genet Med 2013;16:281–5.
5. Snijders RJ, Sundberg K, Holzgreve W, et al. Maternal age- and gestation-specific risk for trisomy 21. Ultrasound Obstet Gynecol 1999;13:167–70.
6. Manri AK, Bhatia G, Strymish J, et al. Medicine's uncomfortable relationship with math: calculating positive predictive value. JAMA Intern Med 2014;174:991–3.
7. Dar P, Curnow KJ, Gross SJ, et al. Clinical experience and follow-up with large scale single-nucleotide polymorphism-based noninvasive prenatal aneuploidy testing. Am J Obstet Gynecol 2014;211:527.e1–17.
8. Relman AS. The new medical–industrial complex. NEJM 1980; 303:963–70.