Putting guidelines into action: Accurate computation of individualized positive predictive value for aneuploidy screening in cell-free DNA

Counsyl

Christine Lo PhD, Eric A Evans PhD, Colleen L Schmitt MGC, Carrie Haverty MS, Imran S Haque PhD, James Goldberg MD

South San Francisco, CA

Objective

Despite the ACOG and SMFM recommendation for patient specific PPV on results of cell-free DNA screening for fetal aneuploidy[1], commercial laboratories have not reported risk estimates on reports, leaving providers to compute PPV using independent calculators[2,3] or rely on a published PPV for the test as a whole[7].

We describe the statistical integration procedures required to compute accurate PPV for noninvasive prenatal screening, taking into account commonly ignored confounding factors such as data resolution and validation confidence intervals.

Methods

PPV is the probability that a patient is positive given a positive screen result and is a function of prevalence, sensitivity and specificity.

$$\mathrm{PPV} = \frac{TP}{TP+FP} = \frac{\text{sensitivity} \times \text{prevalence}}{\text{sensitivity} \times \text{prevalence} + (1.0 - \text{specificity}) \times (1.0 - \text{prevalence})}$$

Individualized prevalence based on maternal and gestational age was bilinearly interpolated based on published prevalence data, specific to maternal age (years and months) and gestational age (weeks and days). Gestational age was calculated using the collection date and the expected due date (EDD).

$$\text{gestational age (weeks)} = \frac{\text{collection date} - \text{EDD} + 280}{7}$$

We further compute the range of PPV within reasonable bounds of sensitivity and specificity (specifically, within the 95% Clopper-Pearson confidence interval) based on published cfDNA screening tests. For studies with small sample sizes reporting a zero error rate, a pseudo-count correction (*i.e.* n/(n+1)) was used to approximate sensitivity and specificity.

Results

We calculate patient specific risks on cell-free DNA screening utilizing a greater number of data points than any other known available risk estimate. Individualized prevalence (IP) and pseudo-count approximation lead to the most specific post-test risk available.

Conclusion

PPV is a key issue in patient counseling. Because patient-specific risk estimates can greatly affect counseling practice and patients' perception of risk, it is important to provide estimates that are as accurate as possible. We show that accurate PPV computation requires data interpolation and confidence propagation, and provide the methods to use these procedures accurately even as test error rate trends towards zero.

Patient	Maternal age	Gestational age	Overall quoted PPV[7]	PPV without IP[2]	PPV with IP
A	35 years, 0 months	12 weeks, 3 days	92.8%	79%	81.4%
B	35 years, 11 months	11 weeks, 6 days	92.8%	79%	84.8%
C	26 years, 5 months	18 weeks, 0 days	92.8%	53%	52.5%
D	42 years, 8 months	11 weeks, 0 days	92.8%	96%	97.4%

Table 1

Variability of patient specific PPV calculated using pooled sensitivity (99.2%) and specificity (99.91%)[5]. Note the difference in PPV with individualized prevalence (including year and month of maternal age and gestational age to date) compared to PPV calculation that does not consider all factors[2] and overall test PPV ignoring maternal and gestational age[7].

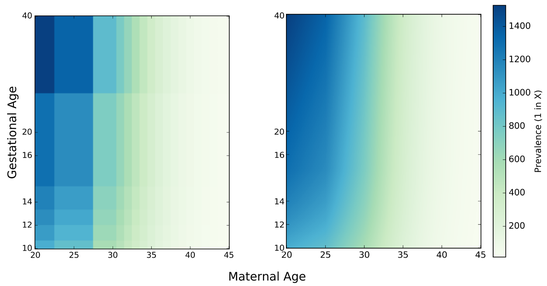

Figure 1

Interpolated prevalence of trisomy 21 using (a) nearest-neighbor and (b) bilinear interpolation of data[4]. The bilinear method displays increased subtlety of prior risk approximation when utilizing a greater number of data points.

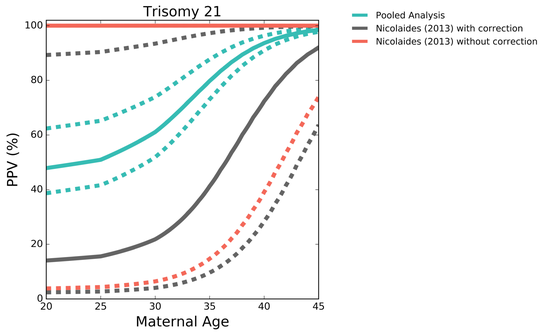

Figure 2

Sample size and related confidence intervals have a significant effect on PPV. Adding pseudo-counts improves the PPV accuracy for small data sets.

The green line shows an approximate "true" T21 PPV, computed from an N>10k meta-analysis[5]. The red line shows T21 PPV computed directly from a study with 25 positive samples, 204 negative samples and 100% sensitivity and specificity[6]. The black line shows T21 PPV computed from that study's data with a pseudo-count correction to account for missed rare events. The trace with pseudo-count correction much more closely approximates the true value of PPV, which is particularly important for events rarely observed in a study.

CITATIONS 1| Cell-free DNA screening for fetal aneuploidy. Committee Opinion No. 640. American College of Obstetricians and Gynecologists. Obstet Gynecol. 2015. 2 | https://www.perinatalquality.org/Vendors/NSGC/NIPT/ 3 | http://mombaby.org/nips_calculator.html 4 | R. Snijders et al. Fetal Diagnosis and Therapy. 1995. 5 | M. M. Gil et al. Ultrasound in Obstetrics & Gynecology. 2015. 6 | K.H Nicolaides, et al. Prenatal Diagnosis. 2013. 7 | P. Taneja et al. Prenatal Diagnosis. 2016.

All Counsyl posters available online at research.counsyl.com