Introduction

While non-invasive prenatal screening (NIPIs) for fetal aneuploidy has high sensitivity and specificity, prevalence varies significantly by maternal and gestational age. Variable prevalence affects the probability that a positive test indicates an affected fetus (positive predictive value, PPV). While ACMG directs laboratories to report PPV individualized to the particular patient, previous work has not addressed how uncertain PPV calculations are or to what precision PPV can be estimated.

Methods

Positive predictive value is a function of condition prevalence, test sensitivity, and test specificity. Uncertainty in any of these parameters will affect the size of the confidence interval for PPV:

$$PPV = \frac{TP}{TP + FP} = \frac{SENS \times PREV}{SENS \times PREV + (1 - SPEC) \times (1 - PREV)}$$

For all three, uncertainty arises from the size of the dataset used to estimate the parameter (e.g., the number of tests used to estimate sensitivity and specificity, or the size and length of population surveillance/data collection for condition prevalence), with larger datasets providing more certainty about the value of a parameter.

Note that uncertainty and accuracy are two different concepts: it is possible to have a very accurate test and be uncertain about the precise value of its accuracy, or conversely to have an inaccurate test and be confident in the estimate of the value of its accuracy (Fig 1).

![Figure 1](image)

Performing a smaller study (blue) makes us less confident in the exact value of the accuracy of a test than if we had performed a larger study (red), regardless of whether the test is accurate (left) or inaccurate (right).

Sampling uncertain values

Test sensitivity and specificity are reported as point estimates with 95% confidence intervals and can be modeled by fitting a beta distribution to these three parameters.

Population studies of common aneuploidies (trisomies 13, 18, and 21) typically analyze population health records and report the number of total pregnancies and the number of trisomy-positive pregnancies at each maternal age. These data are then fit by regression to generate a modeled curve estimating population prevalence by maternal age. These counts can be used as parameters to a beta distribution (Fig 1), from which samples may be repeatedly drawn and regression repeated in order to estimate uncertainty in the final estimate of prevalence (Fig 2).

![Figure 2](image)

Population data on birth prevalence is used to parameterize a beta distribution at each maternal age (gray). Samples can be repeatedly drawn from these distributions; regression on each sample produces a family of estimates for the modeled prevalence function (red).

Prevalence by gestational age is typically modeled as a probability of fetal demise at each gestational age and may be modeled using the beta-regression method above or by bootstrap resampling of the Kaplan-Meier fetal demise curve depending on the available data (Fig 3).

![Figure 3](image)

Samples of gestational prevalence for T13, T18, and T21

Study results

Even incorporating uncertainty, PPV CIs are well-separated between maternal ages, suggesting that patient-specific PPVs are statistically significant and add value to counseling (Fig 4).

![Figure 4](image)

Distribution of positive predictive values for T13, T18, and T21 at a fixed gestational age of 16 weeks, at maternal ages of 20, 25, 30, 35, and 40 years

Conclusion

- Simulation from published data enables computation of the confidence interval (CI) over NIPIs PPV.
- NIPI PPV CIs are well-separated by maternal age, indicating the value of patient-specific PPV computation.
- Additional study of population prevalence could further narrow CIs.

References: