Published in final edited form as:

JAMA. 2011 August 10; 306(6): 627–636. doi:10.1001/jama.2011.1114.

Noninvasive Fetal Sex Determination Using Cell-Free Fetal DNA:

A Systematic Review and Meta-analysis

Stephanie A. Devaney, PhD, Glenn E. Palomaki, PhD, Joan A. Scott, MS, CGC, and Diana W. Bianchi, MD

Genetics and Public Policy Center, Johns Hopkins University, Washington, DC (Dr Devaney and Ms Scott); Department of Health and Human Services, National Institutes of Health, Bethesda, Maryland (Dr Devaney); Women & Infants Hospital, Alpert Medical School of Brown University, Providence, Rhode Island (Dr Palomaki); National Coalition for Health Professional Education in Genetics, Lutherville, Maryland (Ms Scott); and Mother Infant Research Institute at Tufts Medical Center and Division of Genetics, Floating Hospital for Children, Boston, Massachusetts (Dr Bianchi)

Abstract

Context—Noninvasive prenatal determination of fetal sex using cell-free fetal DNA provides an alternative to invasive techniques for some heritable disorders. In some countries this testing has transitioned to clinical care, despite the absence of a formal assessment of performance.

Objective—To document overall test performance of noninvasive fetal sex determination using cell-free fetal DNA and to identify variables that affect performance.

Data Sources—Systematic review and meta-analysis with search of PubMed (January 1, 1997–April 17, 2011) to identify English-language human studies reporting primary data. References from review articles were also searched.

Study Selection and Data Extraction—Abstracts were read independently to identify studies reporting primary data suitable for analysis. Covariates included publication year, sample type, DNA amplification methodology, Y chromosome sequence, and gestational age. Data were independently extracted by 2 reviewers.

Corresponding Author: Stephanie A. Devaney, PhD, Department of Health and Human Services, National Institutes of Health, One Center Dr, Room 103, Bethesda, MD 20892 (stephanie.devaney@nih.gov).

Author Contributions: Dr Devaney had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Palomaki, Scott, Bianchi.

Acquisition of data: Devaney, Bianchi.

Analysis and interpretation of data: Devaney, Palomaki, Bianchi.

Drafting of the manuscript: Devaney.

Critical revision of the manuscript for important intellectual content: Devaney, Palomaki, Scott, Bianchi.

Statistical analysis: Devaney, Palomaki.

Obtained funding: Scott.

Study supervision: Scott, Bianchi.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. No other authors reported disclosures.

Disclaimer: The research that is the basis of this study was conducted prior to Dr Devaney starting at the National Institutes of Health (NIH), and the views expressed in the article do not necessarily represent the views of the NIH or the United States Government.

Online-Only Material: eTables 1 and 2 are available at <http://www.jama.com>.

Results—From 57 selected studies, 80 data sets (representing 3524 male-bearing pregnancies and 3017 female-bearing pregnancies) were analyzed. Overall performance of the test to detect Y chromosome sequences had the following characteristics: sensitivity, 95.4% (95% confidence interval [CI], 94.7%–96.1%) and specificity, 98.6% (95% CI, 98.1%–99.0%); diagnostic odds ratio (OR), 885; positive predictive value, 98.8%; negative predictive value, 94.8%; area under curve (AUC), 0.993 (95% CI, 0.989–0.995), with significant interstudy heterogeneity. DNA methodology and gestational age had the largest effects on test performance. Methodology test characteristics were AUC, 0.988 (95% CI, 0.979–0.993) for polymerase chain reaction (PCR) and AUC, 0.996 (95% CI, 0.993–0.998) for real-time quantitative PCR (RTQ-PCR) ($P=.02$). Gestational age test characteristics were AUC, 0.989 (95% CI, 0.965–0.998) (<7 weeks); AUC, 0.994 (95% CI, 0.987–0.997) (7–12 weeks); AUC, 0.992 (95% CI, 0.983–0.996) (13–20 weeks); and AUC, 0.998 (95% CI, 0.990–0.999) (>20 weeks) ($P=.02$ for comparison of diagnostic ORs across age ranges). RTQ-PCR (sensitivity, 96.0%; specificity, 99.0%) outperformed conventional PCR (sensitivity, 94.0%; specificity, 97.3%). Testing after 20 weeks (sensitivity, 99.0%; specificity, 99.6%) outperformed testing prior to 7 weeks (sensitivity, 74.5%; specificity, 99.1%), testing at 7 through 12 weeks (sensitivity, 94.8%; specificity, 98.9%), and 13 through 20 weeks (sensitivity, 95.5%; specificity, 99.1%).

Conclusions—Despite interstudy variability, performance was high using maternal blood. Sensitivity and specificity for detection of Y chromosome sequences was greatest using RTQ-PCR after 20 weeks' gestation. Tests using urine and tests performed before 7 weeks' gestation were unreliable.

Noninvasive prenatal determination of fetal sex could provide an important alternative to invasive cytogenetic determination, which is currently the gold standard for ambiguous genitalia, X-linked conditions, and single-gene disorders such as congenital adrenal hyperplasia. Chorionic villus sampling and amniocentesis have small but measurable rates of procedure-related pregnancy loss.^{1–3} The availability of a reliable noninvasive alternative to determine fetal sex would reduce unintended fetal losses and would presumably be welcomed by pregnant women carrying fetuses at risk for disorders. A much broader potential application for fetal sex detection is family balancing, which poses ethical concerns.

Fetal sex determination can be performed by sonography at as early as 11 weeks' gestation, although not reliably. Test performance across published studies varies significantly. According to a review by Odeh et al,⁴ fetal sex cannot be determined by ultrasound examination in 7.5% to 50% of pregnancies at 11 weeks' gestation, and this decreases to 3% to 24% at 13 weeks. When reported, the sex determination is incorrect as often as 40% of the time at 11 weeks, although by 13 weeks, accuracy (when reported) is close to 100%. A subsequent prospective study reported that fetal sex could be determined by ultrasound 90% of the time, with 86% accuracy from 11 to 14 weeks' gestation.⁵

The presence of cell-free circulating Y chromosome DNA sequences in the plasma of pregnant women was first described in 1997.⁶ Since that report, many groups worldwide have validated the initial finding that Y chromosome sequences can be amplified and used to identify male fetuses. The research has been extended using a variety of methodologies, sex-

specific markers, and sample types across all gestational ages. In countries such as the Netherlands, the United Kingdom, France, and Spain, this testing has already transitioned to routine clinical care despite the absence of a formal assessment of its performance.⁷ More recently, companies have begun offering this technology directly to the consumer over the Internet.⁸ The tests are marketed for nonmedical use to curious parents-to-be with promises in some cases of accuracy as high as 95% to 99% at as early as 5 to 7 weeks' gestation.^{8,9}

To document the performance of noninvasive fetal sex determination using cell-free fetal DNA, identify the variables that affect performance, and identify sources of heterogeneity between studies, we performed a systematic review and meta-analysis. We sought to determine the analytic validity of cell-free fetal DNA testing, which describes the test's ability to detect Y chromosome sequences within maternal samples, as well as the clinical validity of the test, as indicated by its ability to correctly identify fetal sex. Primarily, the studies focused on the overarching question, "How reliably can fetal sex be predicted by sex-specific markers in cell-free fetal DNA from a maternal blood or urine sample?"

METHODS

Literature Search

We searched the PubMed database for English-language studies containing primary data involving humans. The search included articles published between January 1, 1997, and April 17, 2011. The first report of the presence of fetal DNA in maternal blood was published in 1997.⁶ The following search terms were used: *cell-free fetal DNA*, *fetal sex DNA*, *prenatal sex DNA*, *fetal sex detection*, *cell-free prenatal DNA*, *non-invasive fetal sex*, and *non-invasive prenatal*. The alternate spelling *foetal* was also used. Reference lists from review articles were searched for additional relevant publications.

Study Selection

Two authors (S.A.D., D.W.B.) read the abstracts of each study. Review articles and others that did not contain primary data were excluded. Each abstract was screened for specific key words representing inclusion criteria: *DNA*, *fetal sex detection*, *Y chromosome sequences*, *pregnancy*, *blood*, *plasma*, *serum*, *amelogenin*, and *quantitative PCR*. Amelogenin is a protein encoded by a gene present on both sex chromosomes but with a 6-base pair deletion on the X chromosome version (the genes are *AMELX* and *AMELY* on the X and Y chromosomes, respectively). Because the focus of our study was on blood, urine, DNA, and pregnancy, each identified abstract was then screened for the following key words representing exclusion criteria: *RNA*, *amniotic fluid*, *peritoneal fluid*, *Rhesus D*, *single gene disorder*, *formaldehyde*, *beta thalassemia*, *beta globin*, *fetal cells*, *leptin*, *postpartum*, and *spontaneous abortion*. The covariates chosen were publication year, sample type, DNA amplification methodology, Y chromosome sequence, and gestational age.

Data Extraction

The primary reviewer (D.W.B.) performed the preliminary extraction of data from each selected study using a standard form. The review form was designed to capture primary data, including the number of samples from women with singleton male and female fetuses,

the number of each postulated to be male, the number of indeterminate cases, and the cause of false-negative or false-positive results. Additional data included sample type, methodology (sample handling, polymerase chain reaction [PCR] technique, amplified gene sequence, housekeeping gene[s], type[s] of control[s] used, the number of replicates, and the referent test [ie, the method used to confirm fetal sex]), and any other clinical factors described in the study, such as preeclampsia or preterm labor. The primary reviewer then made a judgment about whether the study contained data useful for the subsequent analysis of clinical validity. Using the same form, a second reviewer (S.A.D.) also extracted the data to be used in the meta-analysis. Interreviewer discrepancies were resolved by discussion.

Assessment of Study Quality

The primary reviewer assessed study quality with a standardized form. The reviewer rated aspects of study design across 23 relevant criteria that have implications for analytic validity and clinical validity using a published methodology for grading evidence from genetic tests.¹⁰ We then applied numeric values to the grades (unsure=0, inadequate= 0, adequate=1, and convincing= 2), for a potential total score of 46. We divided the total possible score into tertiles: 0 through 15 (low quality), 16 through 30 (average quality), and 31 through 46 (high quality).

Data Analysis

The data were analyzed using Meta-DiSc version 1.4 (available at http://www.hrc.es/investigacion/metadisc_en.htm).¹¹ We calculated the sensitivity (proportion of male fetuses with a positive test result) and specificity (proportion of female fetuses with a negative test result) for each included data set. Indeterminate test results were considered to be false-positive or false-negative depending on the true fetal sex. The threshold for calling a result indeterminate may differ between studies. However, we did not account for this variable, because indeterminate results were uncommon and the threshold used for making a determination was not often reported.

Given the relationship between sensitivity and specificity introduced by a potential threshold effect, the diagnostic odds ratio (OR) was used as the primary method of calculating the summary test performance. We decided a priori that we would not rely solely on summary sensitivities and specificities, because groups are likely to have used different cutoffs to identify a positive test result (threshold effect). To test for this possibility, we plotted study-specific sensitivity and specificity on a summary receiver operator characteristic (ROC) curve and looked for a curvilinear shape that is characteristic of the threshold effect. The curve was fit to the data using the Moses-Shapiro-Littenberg model.¹² Any difference in the threshold used to call a result indeterminate is not accounted for by the summary ROC curve. The Spearman correlation coefficient was used to assess the existence of an inverse relationship. Area under curve (AUC) values and asymmetric confidence intervals (CIs) were calculated as previously described.¹³

To deal with zero observations in 2×2 contingency tables, ½ was added to each cell, reducing performance in the small studies. Six outliers outside the 99% CI of the summary ROC curve were identified and removed from further analyses. In 3 of the outlier data sets,

urine was used as the sample type, which is unreliable for detecting cell-free fetal DNA.^{14,15} Two data sets had high rates of indeterminate results (24% and 7.4%),^{16,17} and this translated into high rates of false-positive and false-negative results. The remaining data set reported a number of false-positive and false-negative results; in this study, whole-blood samples were collected between 7 and 11 weeks' gestation, and nested PCR used for the amelogenin-encoding gene was performed.¹⁸ We removed the outlier data sets so they would not have undue influence on the summary ROC curve. The 99% CI was chosen to ensure that only the few most aberrant results would be excluded from the analysis. To reduce the potential for publication bias, we set an a priori inclusion criterion of at least 10 male samples and 10 female samples per study.

Heterogeneity

Heterogeneity was evaluated using I^2 , the proportion of heterogeneity not explained by random chance. I^2 is calculated from the H statistic, which is the square root of the commonly used χ^2 statistic (also known as Q test of heterogeneity).¹⁹ I^2 is the interpretation of the χ^2 value, accounting for the df .

Positive predictive value (PPV) (true positives/[true positives + false positives]) and negative predictive value (NPV) (true negatives/[true negatives + false negatives]) were calculated based on the observed rate for each test. The Q* index, derived from the fitted curve, is the point at which sensitivity equals specificity and was calculated for each test. Because of the threshold effect, we wanted a summary measure of test performance based on the fitted ROC curve.^{20,21} Because the test for fetal sex determination can have high performance, the Q* index usually falls in a clinically relevant region (high sensitivity and specificity). The alternative would be to choose a fixed sensitivity (or specificity), but the choice of appropriate values would be dependent on the setting in which the test is used. Independently pooling sensitivity and specificity across multiple studies may be misleading, because studies used varying cutoff levels to define a positive test result (threshold effect). Fitting the paired results to a summary ROC curve minimizes this issue and allows for more reliable performance estimates. The Q* index allows for a single number to summarize test performance, facilitating comparisons. Level of significance was set at $P < .05$. Testing was 2-sided.

Sensitivity Analysis

To determine whether any single data set was incurring undue weight in the analysis, we systematically removed 1 data set at a time and computed Q* index for the remaining group. This was conducted for each data set.

Univariate Analyses

Prior to analysis, the following potential covariates were chosen: publication date (2003 or earlier, 2004 or later), based on an observed change in study size (Figure 1); sample type (plasma, serum, whole blood, or urine); amplification technique (conventional PCR or real-time quantitative PCR [RTQ-PCR]); Y chromosome sequence (*SRY* [GenBank NG_011751.1], *DYS14* [GenBank NG_027958.1], *DYS1/DAZ* [GenBank NG_004755.2], *DYZ3* [GenBank NG_004755.2], and *AMELY* [GenBank NG_008011.1]); and gestational

age in completed weeks. We chose the following gestational age ranges: early gestation (<7 weeks), 2 clinically relevant ranges (prechorionic villus sampling [7–12 weeks] and preamniocentesis [13–20 weeks]), and late gestation (>20 weeks). Each data set was coded for these covariates. The diagnostic OR was compared between covariate groups using a random-effects model.

Meta-regression

To investigate whether the statistically significant univariate effects were independent, we performed a meta-regression using the Meta-Disc software.¹¹ Meta-Disc uses the Moses-Shapiro-Littenberg method.¹² Studies with missing data were excluded from the meta-regression.

RESULTS

Literature Search

The primary literature search yielded 330 published studies. An additional 529 were identified by searching the reference lists of review articles, for a total of 859 studies. We excluded 713 based on abstract content alone. The remaining 146 studies were read in their entirety (eTable 1, available at <http://www.jama.com>). A total of 74 studies^{6,7,14–18,22–88} were suitable for inclusion in the analyses (Figure 2). Interreviewer agreement was high. There were 1504 data points collected from the 74 studies, and only 15 discrepancies (1.0%) occurred. All were resolved through discussion.

Analytic Validity

Sufficient information was not available to calculate analytic validity. Although a handful of studies did address this issue, most of the studies focused on clinical validity (ie, how reliably fetal sex could be predicted). Imperfect analytical validity will be apparent when assessing clinical validity, because incorrect fetal sex assignments will nearly always be due to analytic error. Therefore, a separate assessment of analytic validity is not necessary in this setting.

Study Selection

Of the 74 studies, 5 were excluded for having insufficient data: 4 had data on sensitivity only (male fetuses only)^{49,73,85,87} and 1 had data on specificity only (female fetuses only).⁶⁶ Two additional studies^{74,82} were excluded because the design was of poor quality; analysis was not blinded in one, and no controls were used in the other. This left 67 studies. Ten small studies* were removed because they included fewer than 10 male fetuses and 10 female fetuses. These also tended to be pilot trials or studies that included only male fetuses or only female fetuses. This left 57 studies for the analyses. Fifteen[†] of these 57 studies could be stratified into multiple data sets. For example, a study that tested the performance of 2 different Y chromosome sequences on the same samples was divided into 2 data sets. Therefore, from the 57 selected studies used in the meta-analysis, we extracted 80 data sets

*References 33–35, 37, 56, 57, 60, 61, 67, 77.

†References 6, 14–17, 23, 25, 27, 55, 59, 62, 70, 75, 78, 88.

(eTable 2). We treated all data sets as independent for the univariate analyses. Among these data sets were discrete samples from pregnant women bearing 3524 singleton male and 3017 singleton female fetuses.

Description of Studies and Data Sets

The sensitivity and specificity of fetal sex determination in each of the 80 data sets was calculated and plotted on a summary ROC curve (Figure 3). The Q^* index is represented by the off-diagonal line. Six outlier data sets (outside the 99% CI) were removed,^{14–18} 3 of which were derived from studies using maternal urine as the sample type, leaving 74 data sets. There were 7 data sets for which urine samples were tested, including the 3 outliers; 2 others that had been excluded for having fewer than 10 male samples, fewer than 10 female samples, or both^{35,60}; 1 that had been removed previously for having data on sensitivity only⁴⁹; and 1 additional data set.⁷⁹ We analyzed urine separately; thus, the additional urine data set was removed,⁷⁹ leaving 73 data sets reporting test results from blood components for analysis. Five data sets^{16,25,62} (described below) were removed for using a method other than conventional PCR or RTQ-PCR, leaving 68 data sets from 52 studies for the univariate analyses (eTable 2).

Heterogeneity

The summary ROC curve (Figure 3) illustrates the significant between-study heterogeneity in test performance (for overall diagnostic OR, $\chi^2=108$, $P<.001$, and $I^2=38\%$). Sensitivities ranged from 65% to 100%, and specificities from 73% to 100%. The AUC was 0.993 (95% CI, 0.989–0.995). Both χ^2 and I^2 showed more heterogeneity than expected by chance alone.

Sensitivity Analysis

We examined whether any one data set had undue influence on the analysis by recomputing Q^* index repeatedly, excluding a different data set on each pass. The data set with the largest influence increased Q^* index from 96.84% to 97.06% (+0.22%)⁵⁵; the next largest reduced Q^* index to 96.63% (–0.21%).⁷⁰ This indicates that no one study had a large influence on overall test performance.

Study Quality

We assessed the 68 data sets included in the univariate analyses for quality of study design. The studies were stratified by overall score into 3 groups: 0 through 15 (poor quality [n=43 data sets]), 16 through 30 (average quality [n=22 data sets]), and 31 through 46 (high quality [n=3 data sets]). There was no significant difference between the performance of data sets that scored 0 through 15 (diagnostic OR, 903; PPV, 98.9%; NPV, 96.7%), those that scored 16 through 30 (diagnostic OR, 863; PPV, 99.2%; NPV, 93.5%), and those that scored 31 through 46 (diagnostic OR, 657; PPV, 94.9%; NPV, 87.0%) ($P=.96$ for comparison of diagnostic ORs) (Table 1).

Univariate Analyses

To examine potential sources of heterogeneity not accounted for by random chance, the 68 data sets were stratified by the following 5 covariates: year of publication, sample type,

amplification technique, Y chromosome sequence, and gestational age. The sensitivity, specificity, diagnostic OR, PPV, NPV, I^2 , and Q* index were stratified by each potential covariate (Table 1). The diagnostic OR and the Q* index were used because they take into account both sensitivity and specificity, which on their own can be less informative as summary statistics because of the threshold effect. The diagnostic OR can be interpreted as the increase in the odds that a fetus determined by the test to be male is actually male.¹¹

Of the 57 studies (80 data sets) plotted on the summary ROC curve, only 7 reported indeterminate results.^{16,17,28,42,43,80,84} The average rate of indeterminate results across these 7 studies was 6.8%. The threshold chosen to make a determination of fetal sex will affect the test failure rate; more stringent thresholds will increase the likelihood of indeterminate results. Two of the outlying studies excluded had high failure rates (7.4% and 24%) attributable to indeterminate results.^{16,17}

Publication Date

Studies were grouped based on their publication date (2003 or earlier, 2004 or later). The year 2003 corresponded to a reduction in the number of studies published that year, as well as a general change in study sizes and methodologies. With one exception, most of the large studies ($n > 150$) were conducted in 2004 or later (Figure 1). In addition, the year of publication may be an indirect measure of improving technology, methodology, or both. Twenty-one of the data sets were published before 2004, and 47 were published in 2004 or later. There was no significant difference between the performance of studies published in 2004 and later (diagnostic OR, 1055; PPV, 98.7%; NPV, 95.5%) compared with those published in 2003 and earlier (diagnostic OR, 532; PPV, 99.1%; NPV, 92.4%) ($P=.15$ for comparison of diagnostic ORs) (Table 1).

Sample Type

In studies for which the samples were derived from maternal blood, for 11 data sets, maternal serum was tested; for 55 data sets, plasma was tested; and for 2 data sets, whole blood was tested. For 7 data sets, urine samples were tested. Plasma (diagnostic OR, 997; PPV, 98.8%; NPV, 95.3%) and serum (diagnostic OR, 776; PPV, 98.9%; NPV, 94.6%) performed similarly across studies ($P=.66$ for comparison of diagnostic ORs). Tests using whole blood were only available for 2 data sets.^{22,68} The CIs for summary sensitivity (85.7% [95% CI, 78.1%–91.5%]) and specificity (92.4% [95% CI, 83.2%–97.5%]) for whole blood were too large to allow drawing any conclusions on the true test performance, so we excluded the 2 data sets prior to performing the meta-regression (Table 1).

Of the 7 data sets for which maternal urine was used, 2 were excluded for having fewer than 10 male samples, fewer than 10 female samples, or both, and 1 data set was excluded for testing males only (ie, only sensitivity data were available).^{35,49,60} The remaining 4 data sets^{14,15,79} showed poor and inconsistent performance (sensitivity: 6.3%, 32.3%, 38.2%, and 95.1%; specificity: 98.2%, 98.2%, 96%, and 87.9%; and PPV, 90.7% and NPV, 52.4% [across the 4 data sets]). The high specificity is likely an artifact; if no fetal DNA can be amplified, the test will not produce false-positive results.

Amplification Technique

Conventional PCR amplification was reported used in 25 data sets and RTQ-PCR in 43 data sets. Four studies (5 data sets) used alternate methods: probe microplate hybridization,²⁵ matrix-assisted laser desorption/ionization time-of-flight mass spectrometry,^{62,84} and pyrophosphorolysis-activated polymerization.¹⁶ We excluded these 5 data sets because of insufficient data for analysis. RTQ-PCR (diagnostic OR, 1284; PPV, 99.1%; NPV, 95.5%; AUC, 0.996 [95% CI, 0.993–0.998]) performed significantly better than conventional PCR (diagnostic OR, 424; PPV, 97.8%; NPV, 92.6%; AUC, 0.988 [95% CI, 0.979–0.993]) for predicting fetal sex ($P=.02$ for comparison of AUCs; $P=.01$ for comparison of diagnostic ORs) (Table 1).

Y Chromosome Sequence

Five different Y chromosome sequences were used: *SRY* ($n=31$), *DYS14* ($n=21$), *DYS1/DAZ* ($n=9$), *DYZ3* ($n=4$), and *AMELY* ($n=3$). Among the 146 studies reviewed in full, only 5 used a Y chromosome sequence other than the 5 listed above: *ZFX*, *ZFY*⁷⁴; short tandem repeat *DYS390*⁸⁹; Y sequence primer pairs *Y1.5/Y1.6*⁴⁶; *CYorf15B*, *NLGN4Y*, *XKRY*⁸²; and *DBY*, *TTY2*.⁸⁴ There was no significant difference between the test performance of *SRY* (diagnostic OR, 996; PPV, 99.3%; NPV, 94.5%), *DYS14* (diagnostic OR, 701; PPV, 97.9%; NPV, 94.6%), and *DYS1/DAZ* (diagnostic OR, 1387; PPV, 99.2%; NPV, 97.5%) ($P=.61$ for comparison of diagnostic ORs). The other 2 genes, *DYZ3* and *AMELY*, performed well (*DYZ3* [$n=4$]: diagnostic OR, 646; PPV, 98.4%, NPV, 95.0%; *AMELY* [$n=3$]: diagnostic OR, 802; PPV, 98.3%, NPV, 95.2%). However, the CIs were too large to be informative (*DYZ3*: sensitivity, 95.4% [95% CI, 90.2%–98.3%]; specificity, 98.3% [95% CI, 93.9%–99.8%]; *AMELY*: sensitivity, 95.9% [95% CI, 90.8%–98.7%]; specificity, 98.0% [95% CI, 93.0%–99.8%]). We removed these 7 data sets^{27,53,59,68,78} prior to meta-regression (Table 1).

Gestational Age

Gestational age was divided into 4 different ranges: early gestation (<7 weeks); 2 clinically relevant ranges (prechorionic villus sampling [7–12 weeks] and preamniocentesis [13–20 weeks]); and late gestation (>20 weeks). In the 30 studies included, the researchers had either tested pregnant women within a gestational age range that fit our categories, or they provided the individual data to be properly categorized. A total of 49 data sets were analyzed: 4 from less than 7 weeks, 21 from 7 through 12 weeks, 16 from 13 through 20 weeks, and 8 from more than 20 weeks.

The performance of the test increased with gestational age. The 7 data sets that reported results prior to 7 weeks' gestation had poor performance. The 4 studies included in our univariate analysis^{63,71,72,88} (the other 3 were excluded for insufficient sample numbers^{36,65,69}) showed low sensitivity: 74.5% (53.9%, 72.7%, 72.7%, and 88.9%, respectively). The specificity was 99.1%, which is probably artificially high as a result of no fetal DNA amplification at such early gestation (Table 1). For gestational age less than 7 weeks ($n=4$), the AUC was 0.989 (95% CI, 0.965–0.998). In the late first trimester (7 through 12 weeks, $n=21$) (diagnostic OR, 846; PPV, 98.9%; NPV, 95.1%; AUC, 0.994 [95% CI, 0.987–0.997]) and early second trimester (13 through 20 weeks, $n=16$) (diagnostic

OR, 604; PPV, 99.4%; NPV, 93.3%; AUC, 0.992 [95% CI, 0.983–0.996]), the tests performed equally well ($P=.57$ for comparison of diagnostic ORs). The best test performance was demonstrated in women at 20 weeks' gestation and later ($n=8$) (sensitivity, 99%; specificity, 99.6%; diagnostic OR, 3196; PPV, 99.6%; NPV, 98.9%; AUC, 0.998 [95% CI, 0.990–0.999]) ($P=.02$ for comparison of diagnostic ORs across 4 gestational age ranges) (Table 1).

Multivariate Analysis

Meta-regression was performed on the 46 data sets that had available data for publication year, sample type, DNA amplification methodology, gene, and gestational age (eTable 2). This represents 60% of the original 6541 pregnancies identified in the literature search. We accounted for duplicate data sets, in which the same samples were used for different tests, by stratification. Based on a small number of studies and limited predictiveness of gestational age and method, we chose not to include interaction terms. Only amplification technique and gestational age were found to be significant predictors of test performance. Year of publication (diagnostic OR, 0.87 [95% CI, 0.25–3.01; $P=.83$]), sample type (diagnostic OR, 1.19 [95% CI, 0.27–5.15; $P=.82$]), and gene (diagnostic OR, 0.81 [95% CI, 0.34–1.92; $P=.62$]) were not. Based on these results, we restricted the multivariate model to amplification technique (diagnostic OR, 3.39 [95% CI, 1.08–10.66; $P=.04$]) and gestational age (diagnostic OR, 6.06 [95% CI, 2.24–16.38; $P<.001$]). We then modeled the diagnostic OR for various combinations (Table 2). The highest performance is achieved after 20 weeks' gestation using RTQ-PCR, with a diagnostic OR of 5388.

COMMENT

This systematic review identified 146 publications that could be used to determine the clinical validity of noninvasive prenatal sex determination using cell-free fetal DNA in maternal blood and urine. Possible limitations of our search and potential sources of bias include the use of only 1 medical database, which may result in omission of studies indexed elsewhere^{90,91}; exclusion of unpublished reports; lack of contact with experts for additional studies not found in the electronic search; exclusion of publications not written in English; and a limited exploration of potential publication bias. In addition, the 2 reviewers were not blinded to the article citation during review.

Despite the variability between studies, the overall sensitivity (95.4%) and specificity (98.6%) is high (Table 1). Testing performed prior to 7 weeks' gestation using blood, and all tests using urine, are unreliable.

Univariate analyses showed that amplification technique and gestational age had the strongest effects on test performance, that RTQ-PCR outperformed conventional PCR, and that performance of the test was much lower prior to 7 weeks. Regression analysis and modeling results found that these 2 variables in testing are responsible for some of the heterogeneity observed across studies. The best performance is predicted with use of RTQ-PCR after 20 weeks' gestation. The performance difference between amplification methods is likely attributable to the greater dynamic range of detection of RTQ-PCR vs most traditional PCR methods. The improved performance with later gestation is likely

Author Manuscript

attributable to the increased concentration of cell-free fetal DNA within maternal blood as the fetus and placenta develop. This would explain the poor performance of the test prior to 7 weeks' gestation and the near-perfect performance in the third trimester. Date of publication, sample type (serum vs plasma), and gene target did not have significant effects on test performance.

Author Manuscript

One variable unaccounted for is interlaboratory variability in the criteria for positive and negative determinations. This can be seen in the discordant reporting of indeterminate results. Within our review of the studies, we noted how the researchers handled indeterminate results. This included reading the study for any hints that samples may have been removed from the reported results. We reviewed the studies with this in mind and captured all indeterminate results, whether explicitly stated or made apparent through the data. Apart from the 7 studies with indeterminate results, none of the others reported or made reference to indeterminate results or otherwise indicated samples for which results were not reported. Because the majority of studies do not report the threshold used to make a determination or to declare a result indeterminate, it is difficult to discern if a determination is made for every sample or if indeterminate results might not be reported. If the latter is true, then our understanding of the true failure rate is limited. This will be important to know in the event that this testing transitions to clinical care in the United States.

Author Manuscript

One study that reported a 5.4% failure rate used a strict threshold to make a positive determination: amplification of both *DYS14* and *SRY*.⁸⁰ For tests in which there was no amplification, paternal and maternal biallelic polymorphisms were used to validate the presence of fetal DNA. If fetal DNA could not be validated, the sample was called indeterminate. This example illustrates how the likelihood of obtaining indeterminate results increases as the stringency for making a positive determination increases. The opposite of this is evidenced through poorly performing tests, such as tests conducted before 7 weeks' gestation. For these tests, sensitivity was low (74.5%) but specificity was high (99.1%). It is probable that the concentration of cell-free fetal DNA in some samples was undetectable, but this is not reflected in test specificity. Use of this technology for clinical diagnosis should require stringent criteria that provide validation of negative results. Before calling a sample female, the laboratory should show that fetal DNA is present and can be amplified.

Author Manuscript

Clinical laboratories also should consider the reason for the test when setting the threshold for positive determinations. For X-linked disease, it is more important to eliminate false-negative results (eg, calling a male fetus female), because all male-bearing pregnancies will need follow-up invasive testing. Therefore, the threshold should be set to favor sensitivity at the deficit of specificity. In the case of congenital adrenal hyperplasia, which affects female genitalia in utero, specificity should be favored.

Author Manuscript

Currently, the standard non-invasive method of determining fetal sex is by ultrasound examination. Typically, fetal sonographic studies are performed as part of routine maternal care with the intention of identifying fetal anomalies. Fetal sex is often provided as part of the analysis. Advantages of ultrasound examination are that it is noninvasive, widely available, and accurate after approximately 13 weeks' gestation. However, some studies have shown that diagnosis of fetal sex is not always possible, even at 13 weeks.⁴ One clear

disadvantage is that ultrasound is unreliable before 11 weeks' gestation. For women who are carriers of X-linked diseases and require invasive testing, sex determination earlier than 13 weeks is critical to reduce invasive testing. Our review of the data regarding amplification of Y chromosome markers from maternal blood indicates that this can be performed reliably between 7 and 12 weeks' gestation, earlier than with sonography. A disadvantage of maternal blood testing is the need to verify the presence of fetal DNA to validate determinations of female sex. Other potential disadvantages include that DNA testing is not available at point of care, not presently approved by the Clinical Laboratory Improvement Amendments, and not currently reimbursed by insurers.

Many of the studies included here were relatively small. A large, prospective, randomized, blinded clinical trial would be beneficial to help validate test performance under highly controlled testing conditions. In addition, research to develop evidence-based methods for validating the presence of cell-free fetal DNA would strengthen the test by ensuring that negative test results reflect the presence of a true female fetus and not the absence of fetal DNA. A standard set of biallelic maternal and paternal markers—locations across the genome where defined variation allows for comparison of different genome sequences—could provide this validation. Last, this technology is currently advertised directly to consumers by Internet-based companies, some of which have claimed test accuracies of 95% to 99% at as early as 5 to 7 weeks' gestation.^{8,9} These companies could use the results of such reports to ensure that their claims are accurate. In addition, it would be useful to directly query consumers of these tests to better understand their motivations for undergoing testing as well as their intended or actual actions following receipt of test results.

In summary, the overall performance of noninvasive fetal sex determination using maternal blood can be high, if performed using RTQ-PCR on a blood sample taken at a time during pregnancy when sufficient cell-free fetal DNA is present (7 weeks' gestation or later). This technology can be useful in clinical settings for early detection of fetuses at risk for sex-linked disorders requiring follow-up testing.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Dr Palomaki reported having a sponsored research agreement with Sequenom Inc to study circulating cell-free fetal DNA. Dr Bianchi reported serving as a member of the scientific and clinical advisory boards of Verinata Health Inc, an early-stage biotechnology company; she receives honoraria and holds stock options in this company.

Funding/Support: This study was supported by the National Human Genome Research Institute (1R21HG004865-02).

Role of the Sponsors: The National Human Genome Research Institute had no involvement in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation, review, or approval of the manuscript.

Additional Contributions: We thank the following Technical Expert Panel members for their expert guidance throughout the project planning: Linda Bradley, PhD (Department of Pathology and Laboratory Medicine, Women & Infants Hospital of Rhode Island, Alpert Medical School of Brown University), and Katherine Klinger, PhD (Genzyme Corp). Dr Klinger provided input on the analytic framework and key questions, the search terms for the

literature search, the study inclusion and exclusion criteria, and the information to extract from each study for the evidence tables and was not involved with any other aspect of the study. We also thank Sara Katsanis, MS (Duke Institute for Genome Sciences & Policy at Duke University), for her help in creating the data forms and the list of articles for review. None of the individuals acknowledged were compensated for their contributions.

REFERENCES

1. Eddleman KA, Malone FD, Sullivan L, et al. Pregnancy loss rates after midtrimester amniocentesis. *Obstet Gynecol.* 2006; 108(5):1067–1072. [PubMed: 17077226]
2. Mujezinovic F, Alfirevic Z. Procedure-related complications of amniocentesis and chorionic villous sampling. *Obstet Gynecol.* 2007; 110(3):687–694. [PubMed: 17766619]
3. Tabor A, Vestergaard CH, Lidegaard O. Fetal loss rate after chorionic villus sampling and amniocentesis. *Ultrasound Obstet Gynecol.* 2009; 34(1):19–24. [PubMed: 19504504]
4. Odeh M, Granin V, Kais M, Ophir E, Bornstein J. Sonographic fetal sex determination. *Obstet Gynecol Surv.* 2009; 64(1):50–57. [PubMed: 19099612]
5. Chelli D, Methni A, Dimassi K, et al. Fetal sex assignment by first trimester ultrasound: a Tunisian experience. *Prenat Diagn.* 2009; 29(12):1145–1148. [PubMed: 19813214]
6. Lo YM, Corbetta N, Chamberlain PF, et al. Presence of fetal DNA in maternal plasma and serum. *Lancet.* 1997; 350(9076):485–487. [PubMed: 9274585]
7. Hill M, Finning K, Martin P, et al. Non-invasive prenatal determination of fetal sex. *Clin Genet.* 2011; 80(1):68–75. [PubMed: 20840219]
8. Bianchi DW. At-home fetal DNA gender testing. *Obstet Gynecol.* 2006; 107(2, pt 1):216–218. [PubMed: 16449102]
9. Javitt GH. Pink or blue? the need for regulation is black and white. *Fertil Steril.* 2006; 86(1):13–15. [PubMed: 16818030]
10. Teutsch SM, Bradley LA, Palomaki GE, et al. EGAPP Working Group. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative. *Genet Med.* 2009; 11(1):3–14. [PubMed: 18813139]
11. Zamora J, Abaira V, Muriel A, et al. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol.* 2006; 6:31. [PubMed: 16836745]
12. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve. *Stat Med.* 1993; 12(14):1293–1316. [PubMed: 8210827]
13. Agresti, A. *An Introduction to Categorical Data Analysis.* 2nd ed.. Hoboken, NJ: John Wiley & Sons Inc.; 2007.
14. Al-Yatama MK, Mustafa AS, Ali S, Abraham S, Khan Z, Khaja N. Detection of Y chromosome-specific DNA in the plasma and urine of pregnant women using nested polymerase chain reaction. *Prenat Diagn.* 2001; 21(5):399–402. [PubMed: 11360283]
15. Majer S, Bauer M, Magnet E, et al. Maternal urine for prenatal diagnosis—an analysis of cell-free fetal DNA in maternal urine and plasma in the third trimester. *Prenat Diagn.* 2007; 27(13):1219–1223. [PubMed: 17968856]
16. Boon EM, Schlecht HB, Martin P, et al. Y chromosome detection by Real Time PCR and pyrophosphorolysis-activated polymerisation using free fetal DNA isolated from maternal plasma. *Prenat Diagn.* 2007; 27(10):932–937. [PubMed: 17600849]
17. Smid M, Lagona F, de Benassuti L, et al. Evaluation of different approaches for fetal DNA analysis from maternal plasma and nucleated blood cells. *Clin Chem.* 1999; 45(9):1570–1572. [PubMed: 10471665]
18. Falcinelli C, Battafarano S, Neri C, et al. First-trimester fetal sex prediction by deoxyribonucleic acid analysis of maternal peripheral blood. *Am J Obstet Gynecol.* 1999; 181(3):675–680. [PubMed: 10486483]
19. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003; 327(7414):557–560. [PubMed: 12958120]
20. Rosman AS, Korsten MA. Application of summary receiver operating characteristics (sROC) analysis to diagnostic clinical testing. *Adv Med Sci.* 2007; 52:76–82. [PubMed: 18217394]

21. Walter SD. Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data. *Stat Med*. 2002; 21(9):1237–1256. [PubMed: 12111876]
22. Lagona F, Smid M, Papasergio N, Ferrari A, Ferrari M, Cremonesi L. Multiple testing in fetal gender determination from maternal blood by polymerase chain reaction. *Hum Genet*. 1998; 102(6):687–690. [PubMed: 9703431]
23. Lo YM, Tein MS, Lau TK, et al. Quantitative analysis of fetal DNA in maternal plasma and serum. *Am J Hum Genet*. 1998; 62(4):768–775. [PubMed: 9529358]
24. Zhong XY, Bürk MR, Troeger C, Kang A, Holzgreve W, Hahn S. Fluctuation of maternal and fetal free extracellular circulatory DNA in maternal plasma. *Obstet Gynecol*. 2000; 96(6):991–996. [PubMed: 11084191]
25. Chen H, Wang T, He G, Zhu L, Ma T. Gene analysis of free fetal DNA in maternal plasma. *J Tongji Med Univ*. 2001; 21(4):329–331. [PubMed: 12539562]
26. Costa JM, Benachi A, Gautier E, Jouannic JM, Ernault P, Dumez Y. First-trimester fetal sex determination in maternal serum using real-time PCR. *Prenat Diagn*. 2001; 21(12):1070–1074. [PubMed: 11746166]
27. Honda H, Miharu N, Ohashi Y, Ohama K. Successful diagnosis of fetal gender using conventional PCR analysis of maternal serum. *Clin Chem*. 2001; 47(1):41–46. [PubMed: 11148175]
28. Rijnders RJ, van der Schoot CE, Bossers B, et al. Fetal sex determination from maternal plasma in pregnancies at risk for congenital adrenal hyperplasia. *Obstet Gynecol*. 2001; 98(3):374–378. [PubMed: 11530115]
29. Sekizawa A, Kondo T, Iwasaki M, et al. Accuracy of fetal gender determination by analysis of DNA in maternal plasma. *Clin Chem*. 2001; 47(10):1856–1858. [PubMed: 11568102]
30. Wei C, Saller DN, Sutherland JW. Detection and quantification by homogeneous PCR of cell-free fetal DNA in maternal plasma. *Clin Chem*. 2001; 47(2):336–338. [PubMed: 11159785]
31. Zhong XY, Holzgreve W, Hahn S. Risk free simultaneous prenatal identification of fetal Rhesus D status and sex by multiplex real-time PCR using cell free fetal DNA in maternal plasma. *Swiss Med Wkly*. 2001; 131(5–6):70–74. [PubMed: 11383228]
32. Honda H, Miharu N, Ohashi Y, et al. Fetal gender determination in early pregnancy through qualitative and quantitative analysis of fetal DNA in maternal serum. *Hum Genet*. 2002; 110(1):75–79. [PubMed: 11810300]
33. Mazza V, Falcinelli C, Percesepe A, et al. Non-invasive first trimester fetal gender assignment in pregnancies at risk for X-linked recessive diseases. *Prenat Diagn*. 2002; 22(10):919–924. [PubMed: 12378578]
34. Bischoff FZ, Dang DX, Marquez-Do D, et al. Detecting fetal DNA from dried maternal blood spots. *Reprod Biomed Online*. 2003; 6(3):349–351. [PubMed: 12735873]
35. Li Y, Zhong XY, Kang A, Troeger C, Holzgreve W, Hahn S. Inability to detect cell free fetal DNA in the urine of normal pregnant women nor in those affected by preeclampsia associated HELLP syndrome. *J Soc Gynecol Investig*. 2003; 10(8):503–508.
36. Rijnders RJ, Van Der Luijt RB, Peters ED, et al. Earliest gestational age for fetal sexing in cell-free maternal plasma. *Prenat Diagn*. 2003; 23(13):1042–1044. [PubMed: 14691988]
37. Samura O, Miharu N, Hyodo M, et al. Cell-free fetal DNA in maternal circulation after amniocentesis. *Clin Chem*. 2003; 49(7):1193–1195. [PubMed: 12816922]
38. Tungwiwat W, Fucharoen G, Ratanasiri T, et al. Non-invasive fetal sex determination using a conventional nested PCR analysis of fetal DNA in maternal plasma. *Clin Chim Acta*. 2003; 334(1–2):173–177. [PubMed: 12867289]
39. Wataganara T, LeShane ES, Farina A, et al. Maternal serum cell-free fetal DNA levels are increased in cases of trisomy 13 but not trisomy 18. *Hum Genet*. 2003; 112(2):204–208. [PubMed: 12522563]
40. Cremonesi L, Galbiati S, Foglieni B, et al. Feasibility study for a microchip-based approach for non-invasive prenatal diagnosis of genetic diseases. *Ann N Y Acad Sci*. 2004; 1022:105–112. [PubMed: 15251947]
41. Ho SS, Damayanti Z, Chua WY, et al. Non-invasive prenatal diagnosis of fetal gender using real-time polymerase chain reaction amplification of *SRY* in maternal plasma. *Ann Acad Med Singapore*. 2004; 33 suppl(5):S61–S62. [PubMed: 15651212]

42. Hwa HL, Ko TM, Yen ML, Chiang YL. Fetal gender determination using real-time quantitative polymerase chain reaction analysis of maternal plasma. *J Formos Med Assoc.* 2004; 103(5):364–368. [PubMed: 15216403]
43. Johnson KL, Dukes KA, Vidaver J, et al. Interlaboratory comparison of fetal male DNA detection from common maternal plasma samples by real-time PCR. *Clin Chem.* 2004; 50(3):516–521. [PubMed: 14718394]
44. Rijnders RJ, Christiaens GC, Bossers B, et al. Clinical applications of cell-free fetal DNA from maternal plasma. *Obstet Gynecol.* 2004; 103(1):157–164. [PubMed: 14704260]
45. Wataganara T, Chen AY, LeShane ES, et al. Cell-free fetal DNA levels in maternal plasma after elective first-trimester termination of pregnancy. *Fertil Steril.* 2004; 81(3):638–644. [PubMed: 15037414]
46. Zhao XX, Suzumori N, Ozaki Y, Sato T, Suzumori K. Examination of fetal cells and cell-free fetal DNA in maternal blood for fetal gender determination. *Gynecol Obstet Invest.* 2004; 58(1):57–60. [PubMed: 15153684]
47. Galbiati S, Smid M, Gambini D, et al. Fetal DNA detection in maternal plasma throughout gestation. *Hum Genet.* 2005; 117(2–3):243–248. [PubMed: 15906093]
48. Hyett JA, Gardener G, Stojilkovic-Mikic T, et al. Reduction in diagnostic and therapeutic interventions by non-invasive determination of fetal sex in early pregnancy. *Prenat Diagn.* 2005; 25(12):1111–1116. [PubMed: 16231296]
49. Koide K, Sekizawa A, Iwasaki M, et al. Fragmentation of cell-free fetal DNA in plasma and urine of pregnant women. *Prenat Diagn.* 2005; 25(7):604–607. [PubMed: 16032774]
50. Pan PD, Peter I, Lambert-Messerlian GM, et al. Cell-free fetal DNA levels in pregnancies conceived by IVF. *Hum Reprod.* 2005; 20(11):3152–3156. [PubMed: 16006472]
51. Wataganara T, Chen AY, LeShane ES, et al. Changes of cell-free fetal DNA in maternal plasma after elective termination of pregnancy. *Clin Chem.* 2005; 51(1):217–219. [PubMed: 15528293]
52. Wataganara T, Metzenbauer M, Peter I, et al. Placental volume, as measured by 3-dimensional sonography and levels of maternal plasma cell-free fetal DNA. *Am J Obstet Gynecol.* 2005; 193(2):496–500. [PubMed: 16098876]
53. Zhu B, Sun QW, Lu YC, et al. Prenatal fetal sex diagnosis by detecting amelogenin gene in maternal plasma. *Prenat Diagn.* 2005; 25(7):577–581. [PubMed: 16032768]
54. Zimmermann B, El-Sheikhah A, Nicolaides K, Holzgreve W, Hahn S. Optimized real-time quantitative PCR measurement of male fetal DNA in maternal plasma. *Clin Chem.* 2005; 51(9):1598–1604. [PubMed: 16020496]
55. Zolotukhina TV, Shilova NV, Voskoboeva EY. Analysis of cell-free fetal DNA in plasma and serum of pregnant women. *J Histochem Cytochem.* 2005; 53(3):297–299. [PubMed: 15750006]
56. Bauer M, Hutterer G, Eder M, et al. A prospective analysis of cell-free fetal DNA concentration in maternal plasma as an indicator for adverse pregnancy outcome. *Prenat Diagn.* 2006; 26(9):831–836. [PubMed: 16832830]
57. Chi C, Hyett JA, Finning KM, Lee CA, Kadir RA. Non-invasive first trimester determination of fetal gender. *BJOG.* 2006; 113(2):239–242. [PubMed: 16412004]
58. Davalieva K, Dimcev P, Efremov GD, Plaseska-Karanfilska D. Non-invasive fetal sex determination using real-time PCR. *J Matern Fetal Neonatal Med.* 2006; 19(6):337–342. [PubMed: 16801309]
59. Ge Q, Bai Y, Liu Z, et al. Detection of fetal DNA in maternal plasma by microarray coupled with emulsions PCR. *Clin Chim Acta.* 2006; 369(1):82–88. [PubMed: 16838403]
60. Illanes S, Denbow ML, Smith RP, et al. Detection of cell-free fetal DNA in maternal urine. *Prenat Diagn.* 2006; 26(13):1216–1218. [PubMed: 17139694]
61. Jorgez CJ, Dang DD, Simpson JL, Lewis DE, Bischoff FZ. Quantity versus quality: optimal methods for cell-free DNA isolation from plasma of pregnant women. *Genet Med.* 2006; 8(10):615–619. [PubMed: 17079877]
62. Li Y, Holzgreve W, Kiefer V, Hahn S. Maldi-tof mass spectrometry compared with real-time PCR for detection of fetal cell-free DNA in maternal plasma. *Clin Chem.* 2006; 52(12):2311–2312. [PubMed: 17138855]

63. Martinhago CD, de Oliveira RM, Tomitão Canas MdoC, et al. Accuracy of fetal gender determination in maternal plasma at 5 and 6 weeks of pregnancy. *Prenat Diagn.* 2006; 26(13): 1219–1223. [PubMed: 17089442]
64. Santacroce R, Vecchione G, Tomaiuolo M, et al. Identification of fetal gender in maternal blood is a helpful tool in the prenatal diagnosis of haemophilia. *Haemophilia.* 2006; 12(4):417–422. [PubMed: 16834744]
65. Stanghellini I, Bertorelli R, Capone L, et al. Quantitation of fetal DNA in maternal serum during the first trimester of pregnancy by the use of a DAZ repetitive probe. *Mol Hum Reprod.* 2006; 12(9):587–591. [PubMed: 16820404]
66. Zhong XY, Holzgreve W, Hahn S. Direct quantification of fetal cells in maternal blood by real-time PCR. *Prenat Diagn.* 2006; 26(9):850–854. [PubMed: 16821250]
67. Zhong XY, Holzgreve W, Tercanli S, Wenzel F, Hahn S. Cell-free foetal DNA in maternal plasma does not appear to be derived from the rich pool of cell-free foetal DNA in amniotic fluid. *Arch Gynecol Obstet.* 2006; 273(4):221–226. [PubMed: 16189693]
68. Al-Yatama MK, Mustafa AS, Al-Kandari FM, et al. Polymerase-chain-reaction-based detection of fetal rhesus D and Y-chromosome-specific DNA in the whole blood of pregnant women during different trimesters of pregnancy. *Med Princ Pract.* 2007; 16(5):327–332. [PubMed: 17709918]
69. Illanes S, Denbow M, Kailasam C, Finning K, Soothill PW. Early detection of cell-free fetal DNA in maternal plasma. *Early Hum Dev.* 2007; 83(9):563–566. [PubMed: 17234369]
70. Lapaire O, Volgmann T, Huang D, et al. Maternal smoking: effect on circulating cell-free fetal and total DNA levels in maternal plasma from the second trimester. *Obstet Gynecol.* 2007; 110(6): 1358–1363. [PubMed: 18055732]
71. Liu FM, Wang XY, Feng X, Wang W, Ye YX, Chen H. Feasibility study of using fetal DNA in maternal plasma for non-invasive prenatal diagnosis. *Acta Obstet Gynecol Scand.* 2007; 86(5): 535–541. [PubMed: 17464580]
72. Bustamante-Aragones A, Rodriguez de Alba M, Gonzalez-Gonzalez C, et al. Foetal sex determination in maternal blood from the seventh week of gestation and its role in diagnosing haemophilia in the foetuses of female carriers. *Haemophilia.* 2008; 14(3):593–598. [PubMed: 18328062]
73. Horinek A, Korabecna M, Panczak A, et al. Cell-free fetal DNA in maternal plasma during physiological single male pregnancies: methodology issues and kinetics. *Fetal Diagn Ther.* 2008; 24(1):15–21. [PubMed: 18504375]
74. Lun FM, Chiu RW, Allen Chan KC, et al. Microfluidics digital PCR reveals a higher than expected fraction of fetal DNA in maternal plasma. *Clin Chem.* 2008; 54(10):1664–1672. [PubMed: 18703764]
75. Picchiassi E, Coata G, Fanetti A, et al. The best approach for early prediction of fetal gender by using free fetal DNA from maternal plasma. *Prenat Diagn.* 2008; 28(6):525–530. [PubMed: 18509845]
76. Urato AC, Peter I, Canick J, et al. Smoking in pregnancy is associated with increased total maternal serum cell-free DNA levels. *Prenat Diagn.* 2008; 28(3):186–190. [PubMed: 18260159]
77. Vainer OB, Katokhin AV, Kustov SM, et al. A new Y chromosome marker for noninvasive fetal gender determination. *Ann N Y Acad Sci.* 2008; 1137:157–161. [PubMed: 18837941]
78. Wagner J, Dzijan S, Pavan-Juki D, Wagner J, Lauc G. Analysis of multiple loci can increase reliability of detection of fetal Y-chromosome DNA in maternal plasma. *Prenat Diagn.* 2008; 28(5):412–416. [PubMed: 18395868]
79. Shekhtman EM, Anne K, Melkonyan HS, et al. Optimization of transrenal DNA analysis: detection of fetal DNA in maternal urine. *Clin Chem.* 2009; 55(4):723–729. [PubMed: 19181739]
80. Scheffer PG, van der Schoot CE, Page-Christiaens GC, et al. Reliability of fetal sex determination using maternal plasma. *Obstet Gynecol.* 2010; 115(1):117–126. [PubMed: 20027043]
81. Sikora A, Zimmermann BG, Rusterholz C, et al. Detection of increased amounts of cell-free fetal DNA with short PCR amplicons. *Clin Chem.* 2010; 56(1):136–138. [PubMed: 19884486]
82. Treff NR, Tao X, Su J, et al. Tracking embryo implantation using cell-free fetal DNA enriched from maternal circulation at 9 weeks gestation. *Mol Hum Reprod.* 2011; 17(7):434–438. [PubMed: 21357209]

83. Miura K, Higashijima A, Shimada T, et al. Clinical application of fetal sex determination using cell-free fetal DNA in pregnant carriers of X-linked genetic disorders. *J Hum Genet.* 2011; 56(4): 296–299. [PubMed: 21307866]
84. Akolekar R, Farkas DH, VanAgtmael AL, et al. Fetal sex determination using circulating cell-free fetal DNA (ccffDNA) at 11 to 13 weeks of gestation. *Prenat Diagn.* 2010; 30(10):918–923. [PubMed: 20721878]
85. Vora NL, Johnson KL, Lambert-Messerlian G, et al. Relationships between cell-free DNA and serum analytes in the first and second trimesters of pregnancy. *Obstet Gynecol.* 2010; 116(3):673–678. [PubMed: 20733451]
86. Vora NL, Johnson KL, Peter I, et al. Circulating cell-free DNA levels increase variably following chorionic villus sampling. *Prenat Diagn.* 2010; 30(4):325–328. [PubMed: 20127897]
87. Hung TH, Lo LM, Chiu TH, et al. A longitudinal study of oxidative stress and antioxidant status in women with uncomplicated pregnancies throughout gestation. *Reprod Sci.* 2010; 17(4):401–409. [PubMed: 20167745]
88. Mortarino M, Garagiola I, Lotta LA, Siboni SM, Semprini AE, Peyvandi F. Non-invasive tool for foetal sex determination in early gestational age [published online ahead of print April 15, 2011]. *Haemophilia.* 2011
89. González-González C, Garcia-Hoyos M, Trujillo-Tiebas MJ, et al. Application of fetal DNA detection in maternal plasma. *J Histochem Cytochem.* 2005; 53(3):307–314. [PubMed: 15750008]
90. Egger M, Juni P, Bartlett C, Hoenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? *Health Technol Assess.* 2003; 7(1):1–76. [PubMed: 12583822]
91. Sampson M, Barrowman NJ, Moher D, et al. Should meta-analysts search Embase in addition to Medline? *J Clin Epidemiol.* 2003; 56(10):943–955. [PubMed: 14568625]

Figure 1. Study Size by Publication Year
Included studies (N=57) are listed in eTable 2.

Figure 2. Literature Search^aMost common: review study, marker not sex-specific, DNA extraction methodology.

Figure 3. Sensitivity and Specificity for 57 Included Studies (80 Data Sets) for Predicting Fetal Sex, With Summary Receiver Operator Curve

Included studies are listed in eTable 2. Diagonal dotted lines indicate the Q* index (point at which sensitivity = specificity on the curve). Indeterminate results were considered false-positive or false-negative depending on the true fetal sex. The curve shown in the right panel was fit to the data (including outliers) using the Moses-Shapiro-Littenberg model.¹² Right panel is an enlarged view of the blue tinted region shown in the left panel.

Table 1

Test Performance

Variable	Data Sets, No.	%		Measure (95% CI)			%		<i>P</i> Value ^e
		PPV ^a	NPV ^b	Sensitivity	Specificity	Diagnostic OR	<i>I</i> ^{2c}	<i>Q</i> ^{sd}	
Study quality ^f									
0–15	43	98.9	96.7	97.2 (96.5–97.9)	98.7 (98.1–99.2)	903 (527–1548)	27	96.7	.96
16–30	22	99.2	93.5	93.9 (92.6–95.1)	99.2 (98.5–99.6)	863 (355–2100)	56	97.3	
31–46	3	94.9	87.0	89.2 (84.9–92.7)	93.7 (89.5–96.6)	657 (86–5019)	53	96.3	
Year of publication									
2003	21	99.1	92.4	93.6 (91.7–95.1)	99.0 (97.9–99.6)	532 (250–1134)	23	96.3	.15
>2003	47	98.7	95.5	96.0 (95.2–96.7)	98.5 (98.0–98.9)	1055 (614–1817)	41	97.1	
Sample type									
Plasma	55	98.8	95.3	95.6 (94.8–96.3)	98.8 (98.3–99.1)	997 (610–1629)	34	97.1	<.01
Serum	11	98.9	94.6	96.6 (94.7–97.9)	98.2 (96.2–99.3)	776 (287–2097)	18	96.7	
Whole blood	2	95.3	78.2	85.7 (78.1–91.5)	92.4 (83.2–97.5)	87 (28–275)	0	50.0	
Urine	4	59.3	9.3	41.3 (36.0–46.9)	93.8 (89.8–96.6)	24 (4–133)	72	90.1	
Test method									
PCR	25	97.8	92.6	94.0 (92.4–95.4)	97.3 (95.9–98.3)	424 (224–803)	24	95.4	.01
RTQ-PCR	43	99.1	95.5	96.0 (95.2–96.6)	99.0 (98.5–99.3)	1284 (729–2261)	36	97.6	
Gene									
<i>SRY</i>	31	99.3	94.5	95.0 (93.9–95.9)	99.3 (98.7–99.6)	996 (493–2011)	41	97.2	<.01
<i>DYS14</i>	21	97.9	94.6	95.0 (93.7–96.1)	97.7 (96.7–98.5)	701 (288–1710)	57	96.5	
<i>DYS1/DAZ</i>	9	99.2	97.5	98.4 (96.9–99.3)	98.7 (96.8–99.6)	1387 (502–3831)	0	97.5	
<i>DYZ3</i>	4	98.4	95.0	95.4 (90.2–98.3)	98.3 (93.9–99.8)	646 (142–2932)	0	96.1	

Variable	Data Sets, No.	%		Measure (95% CI)			%		<i>P</i> Value ^e	
		ppv ^a	NPV ^b	Sensitivity	Specificity	Diagnostic OR	<i>I</i> ^{2c}	Q ^{sd}		
AMELY	3	98.3	95.2	95.9 (90.8–98.7)	98.0 (93.0–99.8)	802 (134–4778)	0	96.2	.02	
Gestational age, mo ^g										
<7	4	98.8	80.6	74.5 (65.1–82.5)	99.1 (95.2–99.9)	116 (31–437)	0	95.7		
7–12	21	98.9	95.1	94.8 (93.1–96.2)	98.9 (98.0–99.5)	846 (372–1925)	28	97.0		
13–20	16	99.4	93.3	95.5 (93.5–97.1)	99.1 (97.5–99.8)	604 (271–1347)	0	96.3		
>20	8	99.6	98.9	99.0 (97.7–99.6)	99.6 (98.6–99.9)	3196 (710–14397)	32	98.3		
All	68	98.8	94.8	95.4 (94.7–96.1)	98.6 (98.1–99.0)	885 (565–1386)	32	96.8		

Abbreviations: CI, confidence interval; NPV, negative predictive value; OR, odds ratio; PCR, polymerase chain reaction; PPV, positive predictive value; RTQ-PCR, real-time quantitative polymerase chain reaction.

^aTrue positives/(true positives+false positives); observed rate.

^bTrue negatives/(true negatives+false negatives); observed rate.

^cProportion of variance accounted for by study heterogeneity; based on diagnostic OR.

^dTest sensitivity equals specificity.

^eCalculated across categories of diagnostic OR using a random-effects model.

^fStudy quality was assessed across 23 relevant study criteria; total possible score, 46.

^gIncluded data sets in which gestational age range fit within the categories or in which individual data were provided to allow for stratification.

Table 2

Expected Test Performance Based on Meta-regression Results^a

Method	Gestational Age			%		
	Name	Code ^b	Range, wk	Code	Diagnostic OR ^c	Test Specificity at 98% Sensitivity
PCR	0	0	<7	0	43.3	86.8
RTQ-PCR	1	0	<7	0	146.9	92.4
RTQ-PCR	1	1	7–20	1	889.8	96.8
RTQ-PCR	1	2	>20	2	5388.4	98.7

Abbreviations: NPV, negative predictive value; OR, odds ratio; PCR, polymerase chain reaction; PPV, positive predictive value; RTQ-PCR, real-time quantitative polymerase chain reaction.

^aOnly complete data sets included; missing data were primarily attributable to studies without a usable gestational age breakdown.^bValue used in meta-regression equation (see “Methods”).^cCalculated as $e^{(3.474 + 1.417 \times \text{method} + 1.868 \times \text{gestational age})}$.^dTest sensitivity equals specificity.^eUsing test specificity at 98% sensitivity.