How to Judge Better Screening Tests

Jacob Canick, Ph.D.

Director, Division of Prenatal and Special Testing
Department of Pathology
Women and Infants Hospital
Professor, Department of Pathology
Brown Medical School
Providence, RI, USA

Definition of Prenatal Screening

“The identification, among apparently normal pregnancies, of those at sufficient risk of a specific fetal disorder to justify subsequent invasive and/or costly prenatal diagnostic tests or procedures.”

Cover of the Journal of Medical Screening
### Requirements of a Worthwhile Screening Program

- **Disorder**: Well defined
- **Prevalence**: Known
- **Natural history**: Medically important, for which there is an effective remedy available
- **Financial**: Cost effective
- **Facilities**: Available or easily installed
- **Ethical**: Procedures following a positive result generally agreed and accepted both to the screening authority and to the subjects
- **Test**: Simple and safe
- **Test performance**: Distribution of test values in affected and unaffected known; small overlap; cut-off defined


---

**John Langdon Down (1828-1896)**

**Patient photographed by Langdon Down in 1865**

Features of Down Syndrome (Trisomy 21)

Incidence:
- Overall, 1 in 700 livebirths (23% higher in 2nd trimester)
- Increases with advancing maternal age

Clinical features:
- Mental retardation (mild to severe)
- Heart malformations (40%) and medical complications
- Presenile dementia after age 40

Test Performance:
The challenge in screening is to have a test that has a high detection rate and low false positive rate.

- detection rate: percentage of affecteds called screen positive by the test
  - The higher the better!

- false positive rate: percentage of unaffecteds called screen positive by the test
  - The lower the better!
Determining the Performance of a Screening Test

Need to know:

Detection Rate  percentage of affected pregnancies called positive by the test
False Positive Rate  percentage of unaffected pregnancies called positive by the test

Don’t need to know, but is important in implementation:

Prevalence  how often is the affected pregnancy found in the population being tested?
OAPR (PPV)  odds of affected given a positive result
≡ average risk amongst positives
≡ equivalent to the positive predictive value of the test (percentage of positives that are affected)

EXAMPLE

There are 195 pregnancies to be screened

5 are affected

190 are unaffected

Therefore, the prevalence is 5 in 195 or 1 in 39.
190 unaffected

5 affected

Detection Rate
If 3 of the 5 cases are screen positive.
DR = 60%
False Positive Rate

If 6 of the 190 unaffected pregnancies are screen positive
FPR = 3%

OAPR or PPV
9 pregnancies are called screen positive by the test.
3 of the 9 are affected.
OAPR = 3:6
PPV: 3/9 (33%)
Concept of the Median and MoM

**Background:**

Described by N. Wald in 1976 *

**Rationale:**

- maternal serum (and NT) levels are continually changing during gestation
- for any point in gestation, marker levels are usually log distributed (i.e., skewed to higher values)
- lab to lab variation in measurement markers can be large


**Solution:**

- the median, rather than mean, was chosen as a better estimate of the gestation-specific reference level, to account for a skewed distribution and for high outliers
- the multiple of the median (MoM) was chosen to normalize for the changing AFP values with gestation and between labs

**Result:**

- a simple, easy to remember number, 1 MoM, becomes the most common value for an unaffected, singleton pregnancy.
- the MoM has become the ‘currency’ used in prenatal screening throughout the world
Gestational Dating and MoM

MoM = 2.08  MoM = 1.17
MoM = 0.79

GJ Knight & GE Palomaki, personal communication

J. Canick, 2003
Example: MSAFP MoM Values on a Linear Scale

Example: MSAFP MoM Values on a Log Scale
Calculation of Patient-Specific Risk

Patient-specific risk = patient's \textit{a priori} risk \times \text{likelihood ratio}

The \textit{a priori} risk is given by the population risk, which is empirically derived from epidemiological studies.

\textit{For example}:

- for open spina bifida, the \textit{a priori} risk is often a regional risk and a racial risk.
- for Down syndrome, the \textit{a priori} risk is the risk based on maternal age.

The \textit{likelihood ratio} is the ratio of the heights of the gaussian curves at a specific analyte value.

\[
\text{LR} = \frac{\text{height}_{\text{aff}}}{\text{height}_{\text{unaff}}}
\]

J. Canick, 2003

Maternal Serum AFP (MoM)
Maternal Serum AFP (MoM)

at 3.0 MoM

LR = \frac{\text{height}_{\text{aff}}}{\text{height}_{\text{unaff}}}

LR = \frac{6}{1} = 6

1:1000 \times 6 = 1:167

Maternal Serum AFP (MoM)

at 1.0 MoM

LR = \frac{\text{height}_{\text{aff}}}{\text{height}_{\text{unaff}}}

LR = \frac{1}{8} = 0.125

1:1000 \times 0.125 = 1:8000
At what MoM will the risk not change?

Unaffected

Open Spina Bifida

\[ LR = \frac{\text{height}_{\text{aff}}}{\text{height}_{\text{unaff}}} \]

Maternal Serum AFP (MoM)

\[ LR = \frac{3}{3} = 1 \]

1:1000 x 1 = 1:1000

At what MoM will the risk not change?

Where the two curves cross!
Calculation of Patient-Specific Risk Using Multiple Markers

Prenatal Screening:
*Calculation of Risk Using Multiple Markers*

- Each marker must be provide information on risk that is not provided by another marker used in that test (degree of independence).

- Each marker generates a likelihood ratio:
  
  \[
  \begin{align*}
  &>1 \quad \text{the risk increases} \\
  &<1 \quad \text{the risk decreases}
  \end{align*}
  \]

- The individual likelihood ratios are multiplied to generate the overall likelihood ratio.

- Risk after testing = likelihood ratio\text{overall} \times a\ priori\ risk

- The risk estimate is changed somewhat to account for small correlations between pairs of markers.
Example:
Triple marker screening

\[ \text{Example:} \]
\[ \begin{align*}
\text{AFP} & = 0.5 \text{ MoM} \\
\text{uE3} & = 0.5 \text{ MoM} \\
\text{hCG} & = 2.0 \text{ MoM}
\end{align*} \]
LR at 0.5 MoM = 7/1 = 8

LR at 2.0 MoM = 2/1 = 2

LRAFP x LRuE3 x LRhCG = LRtriple
2.5 x 8 x 2 = 40

40 x a priori risk = risk after test
40 x 1:1000 = 40:1000 = 1:25

Example:

AFP = 1.0 MoM
uE3 = 1.0 MoM
hCG = 1.0 MoM
How can we visualize improvements in screening performance?
A Scale of Risks

The pregnancy is either affected or unaffected

J. Canick, 2003
Maternal Age as a Screening Test:

**Range of Risks: 150 fold**

Detection Rate = 30%
False Positive Rate = 5%

1:1500 to 1:10
Second Trimester Triple Test:
*Range of Risks: 100,000 fold*

Detection Rate = 69%
False Positive Rate = 5%

J. Canick, 2003
The First Trimester Combined Test:
Range of Risks: 1,000,000 fold

Detection Rate = 85%
False Positive Rate = 5%

1:100,000 to 10:1

J. Canick, 2003
The Integrated Test:

Range of Risks: $1,000,000,000,000 \times 10^{12}$ fold

Detection Rate = 85%

False Positive Rate = 1%

J. Canick, 2003
The future?

unaffected

Down syndrome

Reported risk

J. Canick, 2003