Calculation of Positive Predictive Value

The positive predictive value (PPV) is the probability that an individual with a positive screening result (denoted +) has the disease (denoted D). The sensitivity is the probability that an individual with the disease is screened positive and the specificity is the probability that an individual without the disease (denoted ~D) is screened negative. The prevalence of the disease can be interpreted as the probability that a randomly chosen member of the population being screened has the disease.

The relationship between PPV, prevalence, sensitivity and specificity is most succinctly derived using the language of conditional probability. For those familiar with this, we see that the prevalence can be written Pr(D), the sensitivity is Pr(+|D) and specificity is Pr(-|~D). The PPV is Pr(D|+) and can be related to the preceding by an application of Bayes theorem, i.e. Pr(A|B)=Pr(A & B)/Pr(B), so

$$\Pr(\mathrm{D}|+) = \frac{\Pr(+|\mathrm{D})\Pr(\mathrm{D})}{\Pr(+)} = \frac{\Pr(+|\mathrm{D})\Pr(\mathrm{D})}{\Pr(+|\mathrm{D})\Pr(\mathrm{D}) + \Pr(+|\sim\mathrm{D})\Pr(\sim\mathrm{D})}$$
$$= \frac{\text{sensitivity}\times\text{prevalence}}{\text{sensitivity}\times\text{prevalence} + (1-\text{specificity})\times(1-\text{prevalence})}$$

The derivation in words is much less precise. The PPV is the proportion of those with positive screening tests who have the disease. This can be found as the ratio of:

proportion of *all* those screened who have the disease and a positive screening test
to
proportion of *all* those screened who have a positive screening test.

The former is simply the product of the sensitivity and the prevalence, essentially because this is:

$$\text{sens.}\times\text{prev.} = \frac{\text{Number who have D \& screened+}}{\text{Number who have D}} \times \frac{\text{Number who have D}}{\text{Total number screened}}$$

Those with a positive screening result are made up of those who have D and have been correctly screened + PLUS those who do not have the disease and have been incorrectly screened +, so the latter proportion is:

$$\frac{\text{(Number with D and screened +) + (Number without D and screened +)}}{\text{Total number screened}} = \frac{\text{Number with D and screened +}}{\text{Total number screened}} + \frac{\text{Number without D and screened +}}{\text{Total number screened}}$$

The first term on the right is the same as above, namely sensitivity $\times$ prevalence. The second term can be broken down in a similar way to the above, namely

$$\frac{\text{Number without D \& screened+}}{\text{Number without D}} \times \frac{\text{Number without D}}{\text{Total number screened}} = (1-\text{specificity})\times(1-\text{prevalence})$$

Puting these together gives the formula for PPV derived above.
If the sensitivity is 93%, 1-specificity = 6.5% and the prevalence is 0.6%, the formula gives:

$$PPV = \frac{0.93\times 0.006}{0.93\times 0.006 + 0.065\times 0.994} = 0.079 \approx 8\%$$

Further Reading & Sources:

Baum, M. Lancet, 1995, **346**, 436-7.
Forrest, P. *Breast Cancer: the decision to screen*, 1990, Nuffield Provincial Hospitals Trust.
Morrison, A.S. *Screening in Chronic Disease*, 1985, OUP.