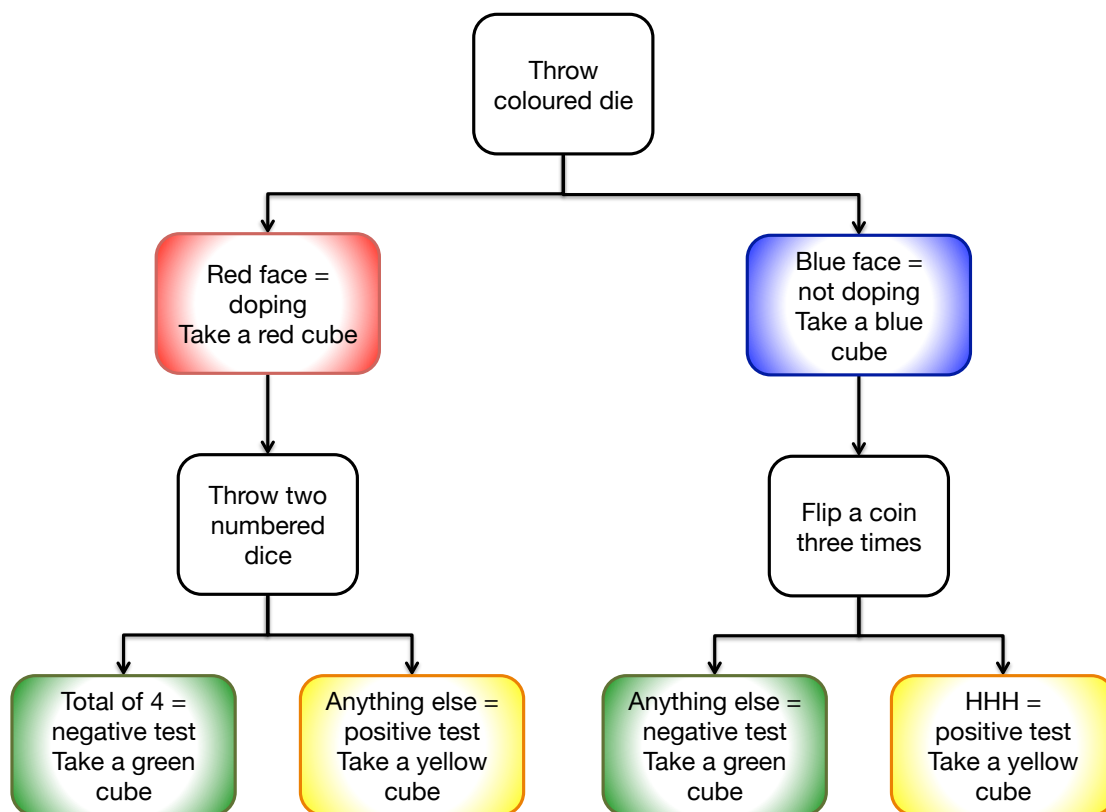# Who is cheating?

A new test is in development to try to identify athletes who use a certain banned substance to enhance their performance.

The test shows quite good results in detecting an athlete who has used this substance, but the false positive rate, ie. athletes who have not used it testing positive, is more worrying.

You are going to investigate the probability that an athlete who is not taking the substance is wrongly accused, and the probability that an athlete who is taking the substance is missed.

Carry out a series of 36 experiments, using the rules given in the flow chart - you should have 36 pairs of multi-link cubes when you finish.
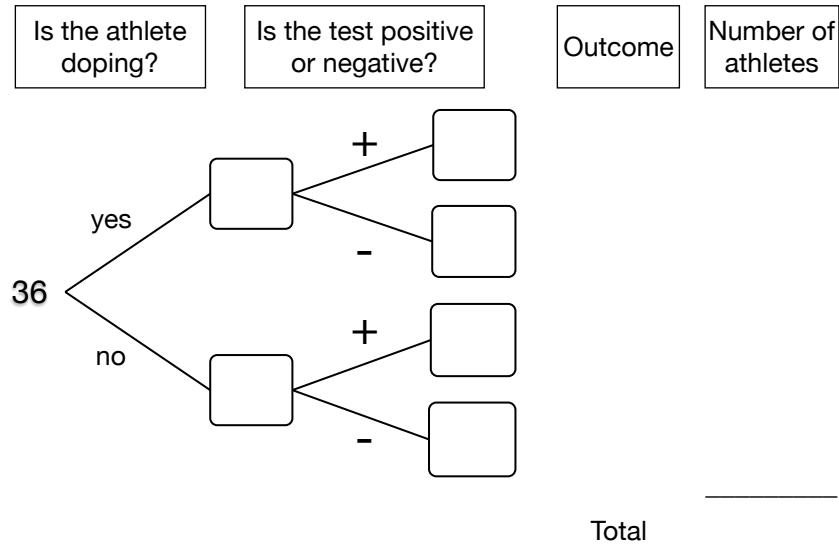
```
                        Throw
                     coloured die
                  /                  \
     Red face =                        Blue face =
       doping                          not doping
   Take a red cube                     Take a blue
                                           cube
         |                                  |
     Throw two                          Flip a coin
     numbered                           three times
       dice
      /      \                          /          \
Total of 4 =   Anything else =   Anything else =    HHH =
negative test  positive test     negative test   positive test
Take a green   Take a yellow     Take a green    Take a yellow
   cube            cube              cube             cube
```

What outcome does each of the following indicate:

• red/green

• red/yellow

• blue/green

• blue/yellow

Which outcomes indicate the test is not doing its job appropriately?
Which is the false positive? Which is the false negative?

Complete the tree diagram and 2-way table below for **your results for 36 athletes**.

Is the athlete doping?	Is the test positive or negative?	Outcome	Number of athletes

```
                                  +  ☐
                    ☐
        yes               -  ☐
36
        no                +  ☐
                    ☐
                          -  ☐
```

_____
Total

		Is the test positive or negative?		
		+	-	Total
Is the athlete doping?	Y			
	N			
Total				

1. Are there any surprises in your results? ....
   What makes them surprising? ....
   ....

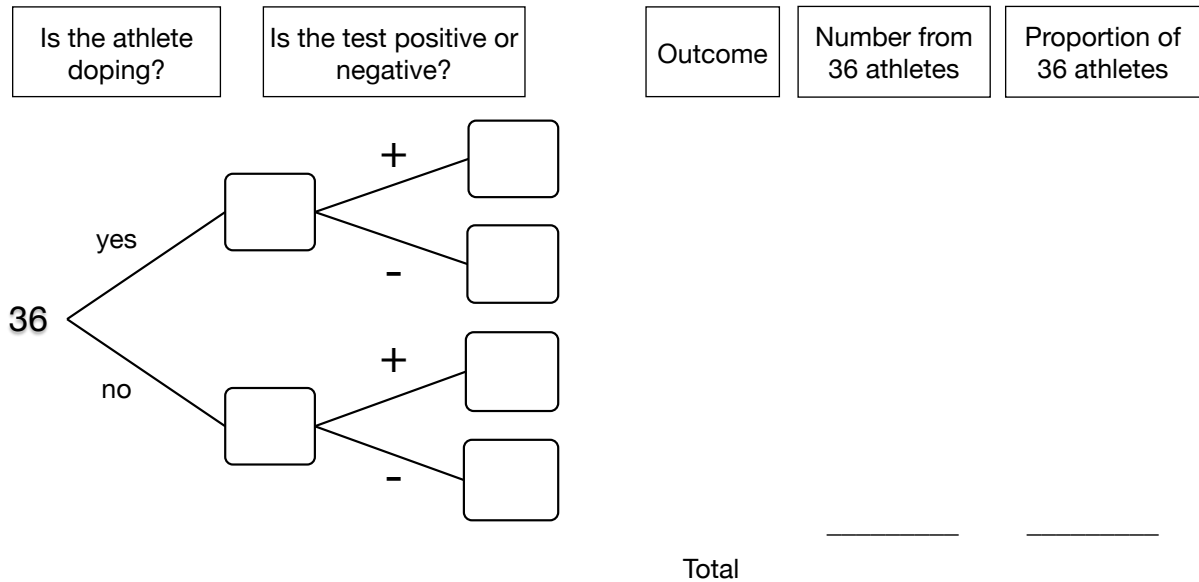For the following questions, give your answers as a fraction of 36.

For questions 4 and 5, you may find it helpful to complete the tables on the Hint Worksheet.

2. What proportion of the 36 trials involved:

   a) athletes who were doping, but tested negative ....

   b) athletes who were not doping but tested positive? ....

3. What proportion of the 36 athletes would you **expect** to be doping? Why? ....
   ....

4. What proportion of athletes **who are doping** would you **expect** to test **negative**? Why? ....
   ....

5. What proportion of athletes **who are not doping** would you **expect** to test positive? Why? ....
   ....

http://nrich.maths.org/9840

How do the observed results compare with what we might expect?

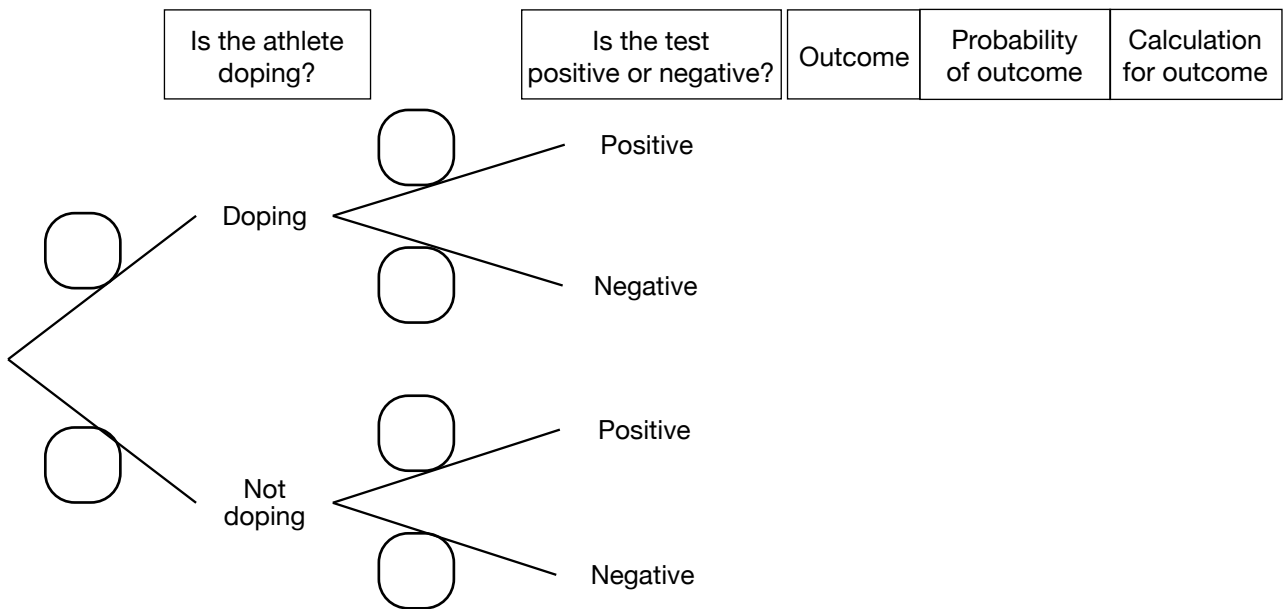Complete the tree diagram below to show what we would **predict for 36 athletes**.

How do your predicted results compare with your experimental data?

| Is the athlete doping? | Is the test positive or negative? | | Outcome | Number from 36 athletes | Proportion of 36 athletes |

```
                              +   [ ]
             [ ]
      yes   /    \
           /      -   [ ]
   36  <
           \      +   [ ]
      no    \    /
             [ ]
                   -   [ ]
```

_____   _____

                                        Total

		Is the test positive or negative?		
		+	-	Total
Is the athlete doping?	Y			
	N			
Total				

6. How many of the 36 athletes would you **expect** to test **positive**, given that they had been taking the banned substance?   ....

7. How many of the 36 athletes would you **expect** to test **positive** in total?   ....

8. Why are these two answers different?   ....
....

9. How many of the 36 athletes would you **expect** to test **negative**, given that they had **not** been taking the banned substance?   ....

10. How many of the 36 athletes would you **expect** to test **negative** in total?   ....

11. Why are these two answers different?   ....
....

On this tree diagram, put the probability (expected proportion) for each event on the appropriate branch. Then complete the 'Outcome' and 'Probability' columns **using the previous tree diagram or 2-way table**.

Is the athlete doping?		Is the test positive or negative?	Outcome	Probability of outcome	Calculation for outcome

Doping

Positive

Negative

Not doping

Positive

Negative

We can use the expected proportions for each outcome to find a general rule for calculating probabilities from a tree diagram:

First add up the four probabilities - is the result you would expect?

Look at the numbers on the two branches for each outcome. How might you combine these to give the required probability?

Does your rule work for all four outcomes?

Can you explain your rule?

Extension questions

You should answer these questions from the tree diagram/2-way table showing expected results (p3). Give answers to probability questions as fractions of the appropriate whole number.

12. How many athletes test positive? ....

13. What is the probability that an athlete who tests positive is taking the banned substance? ....

14. How many athletes are taking the banned substance? ....

15. What is the probability that an athlete who is taking the banned substance tests negative? ....

16. How many athletes are not taking the banned substance? ....

17. What is the probability that an athlete who is not taking the banned substance tests positive? ....

18. Why are these two probabilities not the same? ....
    a) an athlete who is taking the banned substance tests negative ....
    b) an athlete who tests negative is taking the banned substance. ....

http://nrich.maths.org/9840