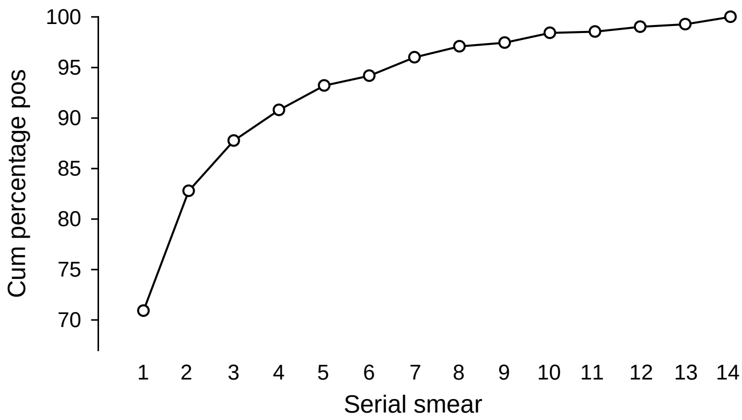
Exercise 3: Incremental yield from serial smears

At the end of this exercise you should be able to:

a. Create a subset of ‘suspects’ from the working dataset
b. Create a string variable that combines the three results for each examinee
c. Make calculations using a spreadsheet
d. Test the given hypothesis on the incremental yield from the third smear
e. Reject or accept a study hypothesis for each country

The diminishing return of serial smears is known from studies that have examined multiple serial specimens, as for example the following study from the 1930s:

Cumulative Yield of Sputum Smear Examination Among 825 Smear-Positive Patients, United Kingdom, 1930s



Hunter RA. Tubercle 1940;21:341-59

This study suggests that each serial smear adds an additional increment in case yield, but the incremental yield gets smaller with each additional examination. Program managers must thus arrive at some optimum that requires the least amount of work (number of smear examinations) to yield a large proportion of cases. The “three smear policy” is such a compromise that has been reached internationally and became reflected in the above mentioned guidelines.

The Union and WHO recommended in the past that each suspect should have three sputum smear examinations before being declared to be “sputum smear-negative”. Some countries recommended only two examinations. The reason for this difference is that The Union and the WHO thought that making a third examination after two smears are negative would offer a sufficiently rewarding incremental yield (how much is rewarding – has anybody ever defined it?) from this third smear as to justify the additional work load for laboratories. Some microscopy laboratories are, however, so burdened with work (particularly in Africa) that a reduction in the required number of examinations would come as a great relief. It is also very possible that over-burdened laboratories may become less meticulous in the examination of a third smear after a first and second smear have been negative, which may reduce the potential incremental gain. Most of the studies determining the incremental yield from the third examination were done under relatively controlled conditions, but there was not much information around on the yield under routine conditions in low- and middle-income countries. The primary hypothesis for the operations research study of the course cohorts of 2003 and 2004 was precisely to test the effectiveness under routine conditions from a representative sample of laboratories in four countries. As these were the only studies of this extent and representativeness, the published findings of these studies greatly contributed to the change in policy of WHO in June 2008 to recommend that routine screening of tuberculosis suspects should be limited to two serial examinations to exclude sputum smear-positive tuberculosis. This demonstrates how powerful a relatively simple study design can be in public health policy shaping, if carried out in a representative manner with diligent adherence to quality assurance.

In this exercise, the approach to this issue will be reproduced.

With our case definition that a suspect becomes a case once acid-fast bacilli are found implies that an additional third examination has to be done only if the two preceding examinations have been negative. *It is not the same as asking how many smears have to be done to find an additional case with the third smear among all suspects*.

Our case is much simpler:

NNP / (NNP + NNN)

This is the incremental gain from a third examination, given that the two preceding examinations have been negative.

This is a fraction, but the hypothesis was about the number of smears. In analogy, you may consider the situation where you know that 20 out of 100 people have a characteristic and you now ask how many you have to examine to find the characteristic.

Confidence intervals

Our fraction might be very small despite the large number of suspects in the database. As the implication of refuting the hypothesis has serious programmatic consequences it is advisable to calculate confidence intervals around the number of smears and decide only to refute if the lower interval is in excess of the hypothesis number X.

The classic approach to estimating 95% confidence intervals is used when the population from which the cases arise is defined (observable) and a subset of this population is examined.

We define **P** as the proportion of cases found on the third smear only among those with three examinations:

P = (NNP/(NNN+NNP)

The standard error of P [(SE(P)] is calculated from the square root of a function derived from P:

SE(P) = SQRT(P*(1-P)/(NNN+NNP))

And the 95% confidence intervals are:

$95\%_{low}$ = P – 1.96*SE(P)

$95\%_{upper}$ = P + 1.96*SE(P)

However, for the number of slides we will need the reciprocals of these values.

Tasks:

Exercise hypothesis:	
H_0:	Not more than 125 third smear examinations have to be made to find one additional case of tuberculosis in each of the four study countries

- ***Determine with a program C_EX03.PGM the number of suspects with the patterns listed above***
- ***Create a table in spreadsheet by country as follows:***

	Moldova	Mongolia	Uganda	Zimbabwe	Total
Total					
Pattern					
N99					
NN9					
NNN					
NNP					
Npx					
Px					
Prop positive					
Yield					
First					
Second					
Third					
X					
P					
SE(P)					
95% low					
95% high					
Smears					
95% low					
95% high					
Hypothesis:					

- ***Interpret the findings***