FALSE DIAGNOSIS +/-

Do you care about a false diagnosis?

01 PROJECT DESCRIPTION

INITIAL GOAL

To study the incidence rate of misdiagnosis in rare diseases.

PROBLEM EXPLANATION

"If there was a medical test with 98% reliability to detect if a person suffers a certain illness (such as cancer) and that 1 in 200 suffer that illness and we are tested positive one day, to what extent should we worry about it? How likely is it that we have that illness?"

Source: Raúl Ibáñez ("We teach the mathematicians to hunt dragons, what are maths and what are they use for?")

	Sick	Healthy	TOTAL
Positive test	49	199	248
Negative test	1	9751	9752
TOTAL	50	9950	10000

Chart explanation:

- The relation 50 ill to 9950 healthy comes from applying the incidence of 1 in each 200 to the 10000 people total.
- The relation 49 ill to 1 healthy comes from applying 98% reliability to the 50 who *"will develop cancer"*.
- The relation 199 ill to 9751 healthy comes from applying 98% reliability to the 9950 who *"won't develop cancer"*.

RESULTS

After analyzing the data from the contingency chart above, we can reach the following conclusions:

Data / **Conclusions**

Test (-)

ILL	HEALTHY	TOTAL
1	9751	9752

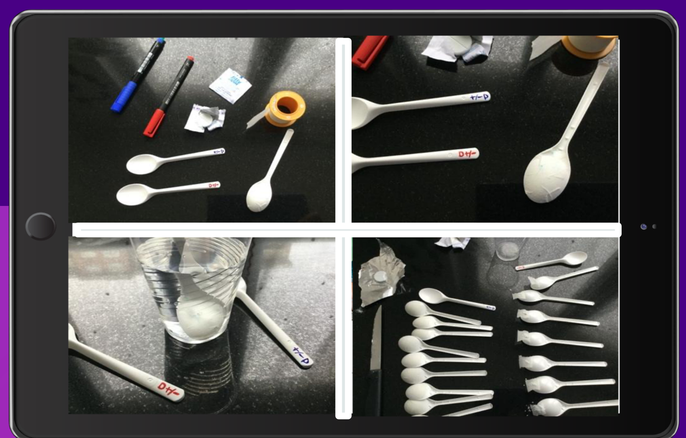
$P(false\ negative) = \frac{1}{9752} = 0{,}0001$

Test (+)

ILL	HEALTHY	TOTAL
49	199	248

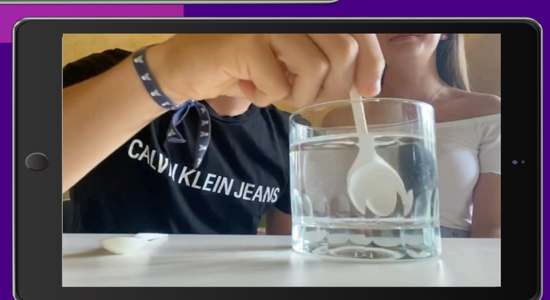
$P(illness \mid true\ positive) == \frac{49}{248} = 0{,}1975$

EXPERIMENT DESIGN



In order to get a device, which gave many fake results we created it ourselves.

We designed some plastic spoons which contain powder that changes colour in contact with water, depending on what we wanted to get. If we wanted a positive (fake) the colour changes and if we wanted a negative result (fake) the colour doesn't change.



Video:

BAYES' THEOREM

- ***Thomas Bayes (s XVIII)***: *If $A_1, A_2, \dots, A_n$ is a system of events, where S is an event of anytype for which we know the $P(S|A_i)$, then:*

$$P(A_i|S) = \frac{P(A_i)\cdot P(S|A_i)}{P(A_1)\cdot P(S|A_1) + \dots + P(A_n)\cdot P(S|A_n)}$$

*$P(A_i)$ are called **"Prior probability"** because they are known.*
*$P(S|A_i)$ are called **"Likelihood"** as they are easily understood.*
*$P(A_i|S)$ are called **"Posterior probability"**, they must be calculated.*

__Application:__ medical diagnosis; establish the diagnosis of a patient from a series of symptoms; the symptoms and the illness don't have a biunivocal correspondence.

02 DATA and GRAPHS

RAREDISEASEDAY.ORG

ABOUT RARE DISEASES (R.D.)

What is considered a rare disease?
A disease that affects 1 out of 2000 people.

How many rare diseases are there?
Up to now 6038 rare diseases are known.

Rare diseases day
It is the last day of February since 2008.

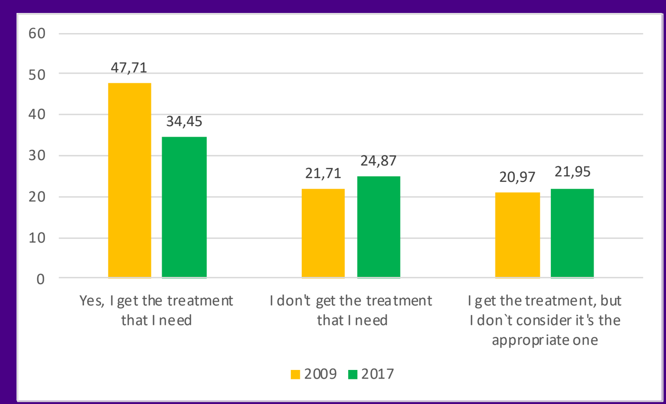
Medical and social consequences
Most of them cannot be cured, which makes the patients vulnerable.

Treatment and diagnosis improvements
There has been progress but these are orphan diseases, only 20% of them are studied.

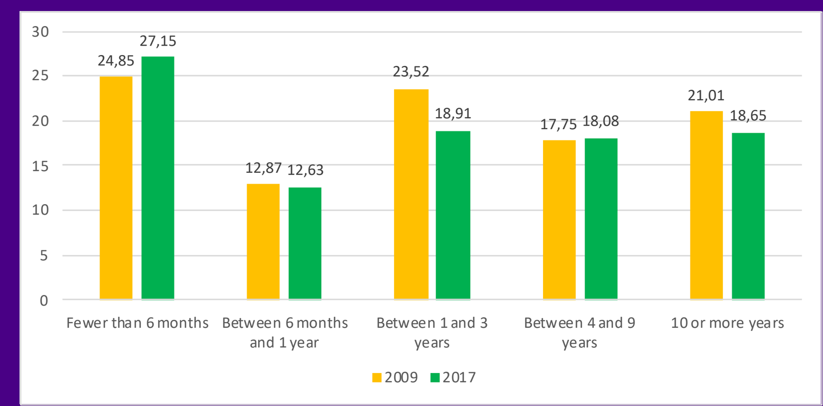
Origin and characteristics
Most of them are genetic, but not all of them. They are chronic and progressive.

MEDICAL ATTENTION IN R.D.

As we can observe:

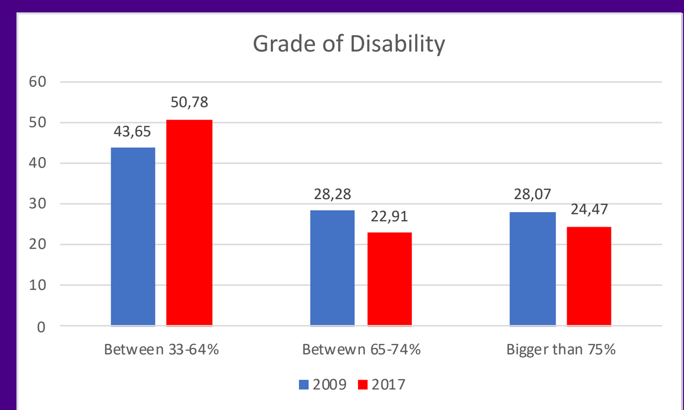


Only 34.5% get the treatment that they need.

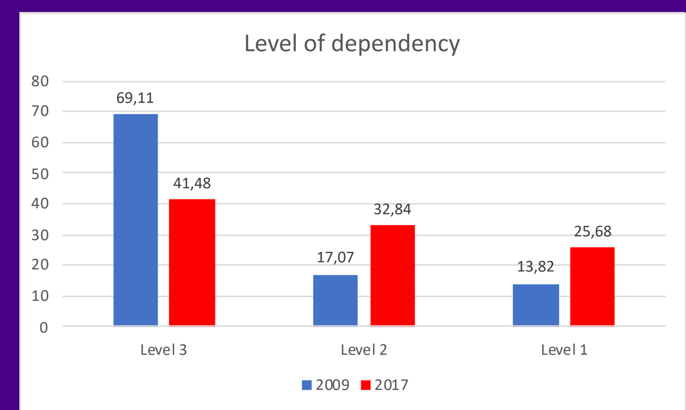


Almost 20% of the patients have had to wait 10 or more years for their diagnosis.

DISABILITY AND DEPENDENCY WITH R.D.

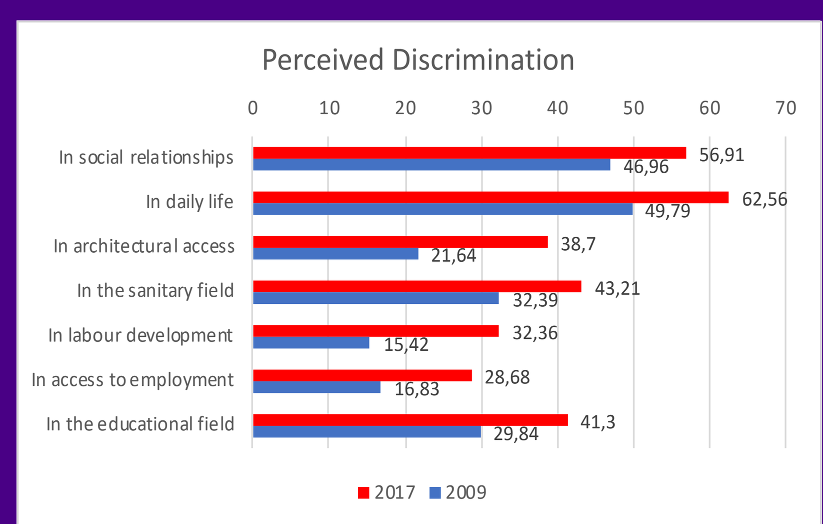


Almost a 25% of the patients have a disability bigger than 75%



More than a 40% of the patients have a severe level of dependency.

DISCRIMINATION



One of the pieces of information that stands out is that in the educational field the discrimination exceeds the 40%.

But what is most impressive is the fact that, since 2009, discrimination has risen in every field.

Source: 2018 ENSERio Study on the Social Health Needs of People with Rare Diseases in Spain

Sources:

Report:

03 STUDIED DISEASES

With the data found on prevalences and reliability of existing diagnoses, we calculated the probabilities of testing positive and of a false positive, using Bayes' Theorem:

ULLRICH CONGENITAL DYSTROPHY (UCMD)

Early appearings, muscle weakness, slow progression, multiple contractures, marked hypermobility and normal intelligence.
There are only 50 confirmed cases in the world.

Let the events (E)=" be Ill"; (S)=" be Healthy"; (+)="Positive test"
As Prevalence 9:1000000 and Test Reliability of 95%, then:
P(E)=9/1000000=0,000009 → P(+|E)=0,000009·0,95=0,00000855
P(S)=0,999991 → P(+|S)=0,999991·0,05=0,04999955

$P(+) = P(+|E) + P(+|S) = 0{,}0500081 = 5\%.$

$$\text{Applying Bayes' } Th.: P(E|+) = \frac{P(+|E)}{P(+|E)+P(+|S)} = \frac{0{,}00000855}{0{,}0500081} = 0{,}00017 = 0{,}02\ \%$$

$P(\text{false} +) = P(S|+) = 1 - P(E|+) = 100\% - 0{,}02\% = 99{,}98\%$

BETHLEM MYOPATHY

Benign autosomal dominant form of slowly progressive muscular dystrophy. Less than 100 cases throughout history.

Prevalence 1:1000000; Test Reliability of 95%, then:
P(E)=1/1000000=0,000001 → P(+|E)=0,000001·0,95=0,00000095
P(S)=0,999999 → P(+|S)=0,999999·0,05=0,04999995

$P(+) = P(+|E) + P(+|S) = 0{,}0500009 = 5\%$

$$\text{By Bayes' Th.: } P(E|+) = \frac{P(+|E)}{P(+|E)+P(+|S)} = \frac{0{,}00000095}{0{,}0500009} = 0{,}000019 = 0{,}002\ \%$$

$P(\text{false} +) = P(S|+) = 1 - P(E|+) = 100\% - 0{,}002\% = 99{,}998\%$

CUSHING'S SYNDROME (CS)

Hormonal disorders due to high and prolonged levels of exposure to glucocorticoids, can be of both endogenous and exogenous origin.

Prevalence 9:100000; Sensitivity test 100% and Specificity 93%:
P(E)=9/100000=0,00009 → P(+|E)=0,00009·1=0,00009
P(S)=0,99991 → P(+|S)=0,99991·0,07=0,0699937

$P(+) = P(+|E) + P(+|S) = 0{,}0700837 = 7\%$

$$\text{By Bayes' Th.: } P(E|+) = \frac{P(+|E)}{P(+|E)+P(+|S)} = \frac{0{,}00009}{0{,}0700837} = 0{,}001284 = 0{,}13\ \%$$

$P(\text{false} +) = P(S|+) = 1 - P(E|+) = 100\% - 0{,}13\% = 99{,}87\%$

04 ANALYSIS OF THE RESULTS:

- ❑ In cases like those studied, there is only a 5-7% chance of suffering from the disease having tested positive.
- ❑ Many shocking results stem from ***__"the false positive paradox"__***: with low prevalences, false positive tests are much more likely than true positives.
- ❑ These results are with tests whose reliability is at least 95%.
- ❑ Probability combined with Statistics is a diagnostic tool to decide if a person is sick (***__using Bayes' Theorem__***).
- ❑ Rare diseases, despite their numbers, are still truly unknown.
- ❑ People who get sick will always be a small percentage compared to what would seem to us due to the reliability of the test (95%).
- ❑ It's always better to keep calm, be patient and optimistic.

Don't worry so much but be careful, please!