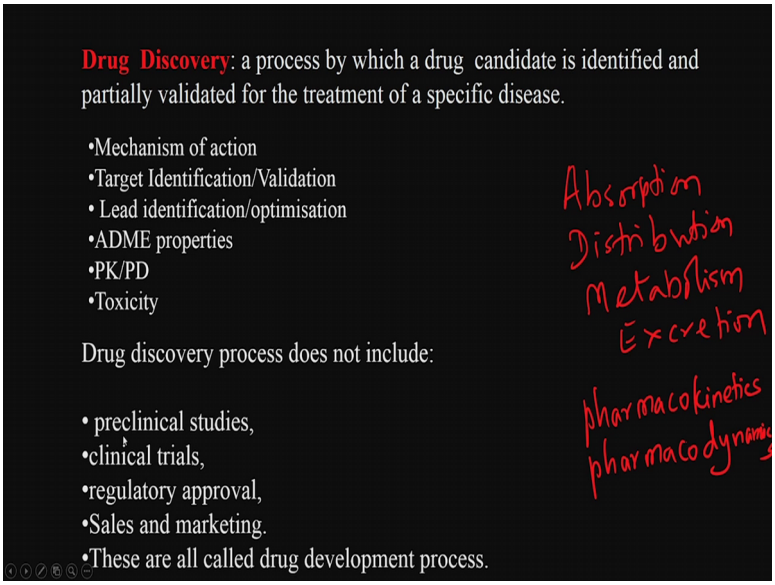
Computer Aided Drug Design
Prof. Mukesh Doble
Department of Biotechnology
Indian Institute of Technology - Madras

Lecture - 02
Drug Discovery - Issues

Hello everyone, we will continue on the course on Computer aided drug design. We will talk about some of the issues in the drug discovery process.

(Refer Slide Time: 00:30)



So let us recap drug discovery is a process by which a drug candidate is identified okay, and partially validated for the treatment of a specific disease. So as I said partially validated, because it becomes a drug only after it gets approval from FDA that is the food and drug administration of USA, so in order to get approval you need to go through all the various preclinical trials involving animals and then human volunteer trials, so until then it is called as lead molecule.

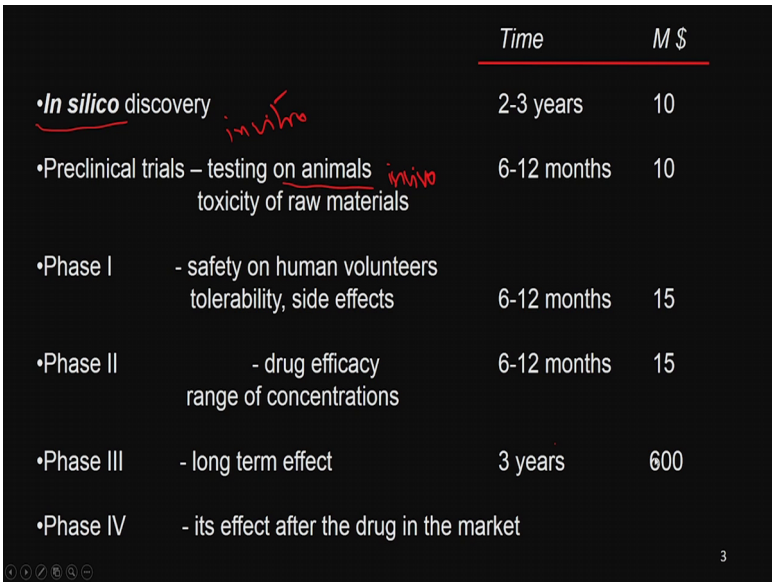
So in my lab I may get a nice candidate which looks very promising, which shows very good activity in my activity screen maybe anti-inflammatory activity or anti-cancer activity, so I call that as a hit molecule, and then I go and finally 0 in 1 candidate I call it as a lead molecule, and the lead molecule is what is being tested in the preclinical clinical trials, and these lead molecules should have not only the activity but other properties which we are going to spend a lot of time on.

So the drug discovery we need to understand the mechanism of action, we need to identify the target which target or enzyme or protein, it is going to go and bind to, and inhibit it or inactivate it, we need to validate that target, and then we need to optimize the lead, so there could be 1 possible candidate which looks very promising. So we may have to change some of its properties, so that it increases solubility maybe the toxicity is reduced side effects is reduced, that is what is called lead optimization.

And it should satisfy all the ADME properties okay, yesterday I mentioned A is absorption, D is distribution, M is metabolism and E is excretion. So it should get absorbed very nicely into the human system okay absorb as we can say, and then comes a distribution into the bloodstream okay maybe it gets distributed into the various tissues okay, then it should get it may be getting metabolized, and then finally once it has done its job it should get excreted that is also very important.

So it should satisfy these ADME properties, and then it should have good pharmacokinetics and pharmacodynamics property okay. yesterday I did talk about pharmacokinetics and pharmacodynamics, so all these and then of course it should not have any toxicity, all these are very important. So the drug discovery we are not going to talk about preclinical studies, clinical trials, regulatory approvals, sales marketing and so on actually okay.

(Refer Slide Time: 03:45)



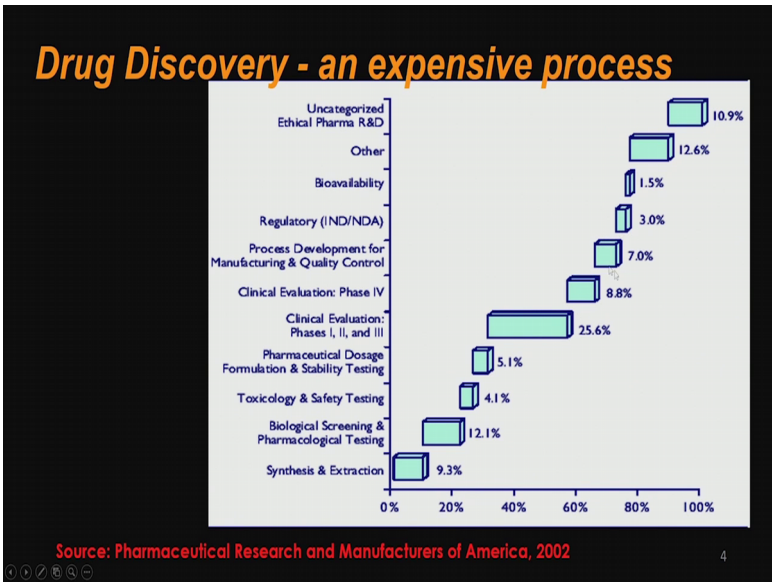
These are not part of the drug discovery process okay, so it is a very long process it should involves a lot of money as you can see here in silico, when we say in silico we use computational tools okay that is called in silico. In vitro means we use laboratory

biochemical assays, proteomic assays that is called in vitro. When we say in vivo it means using animals okay, so there are 3 terms in silico that is the term here which talks using computational tools okay.

In vitro is laboratory studies, and then in animal studies is always called in vivo okay. So there are 3 types of terminologies which we use. So the in silico takes 2 to 3 years may cost about 10 million, then we go to testing on animals that is the in vivo, in between you may you will also have the in vitro also that may take another one year or more it may cost 10 million dollars, then we go through the various human voluntary trials okay phase 1, phase 2, phase 3.

Again you can see 1 year, 1 year maybe 3 years, and these are the cost factors okay, so finally it may add up to almost a billion US dollar that is the current cost of manufacturing sorry current cost of introducing a new molecule into the world.

(Refer Slide Time: 05:16)



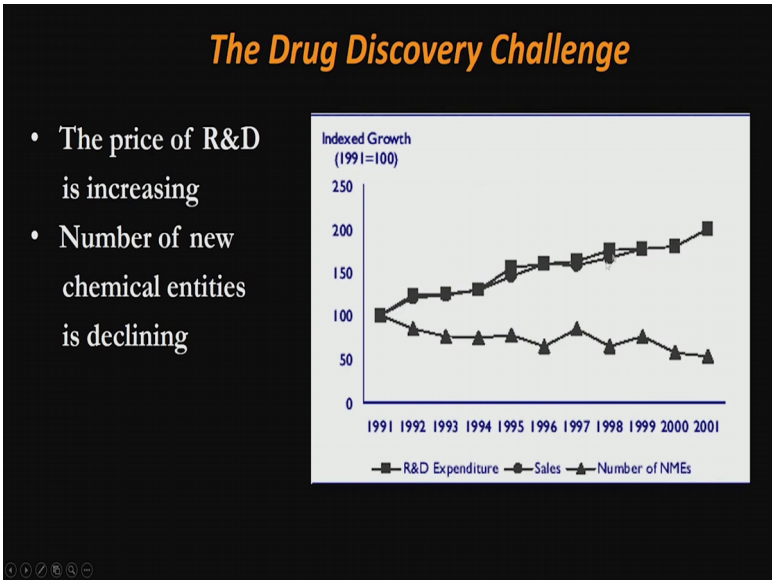
So it is a very expensive process and this is taken from this particular reference here, so as you can see synthesis, extraction if you are a chemist maybe synthesizing a new molecule. If you are using natural products maybe extracting, and then we are going test it against maybe anti-cancer drug or anti-inflammatory drug or diabetic related. So biological testing, then you have toxicological studies, pharmaceutical dosage formulations.

Then comes the clinical trial okay clinical evaluation that is on the human. And then phase 4 trials process development, because you need to manufacture in large quantity, then we need

to get approval from the regulatory authority looking at bioavailability and so on actually. So if you look at this cost factors clinical trials contributes to 25% and so on actually okay, then we are going to regulatory, authorities and that also going to cost a lot of money, biological screening 12%.

So it is an expensive process, that is why when a drug is introduced a new drug is introduced into the market it will be very expensive, because cost of introducing a new drug from the lab is a very expensive process okay.

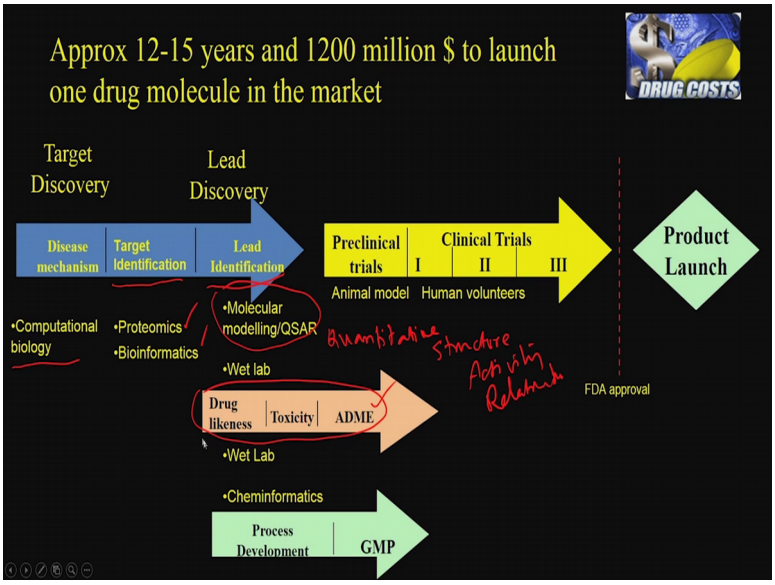
(Refer Slide Time: 06:37)



So if you look over a period of time the cost keeps increasing the R and D cost keeps increasing, but the new chemical entities are sort of decreasing that is because the FDA has become more stringent, they want more information they want to know what are the side effects and they want to know that long term toxic effects of the chemical the toxic effect, the metabolized products.

So that is why the cost is going up and many compounds do not succeed and across the FDA barrier okay. And that is a big problem nowadays in drug discovery that is a big challenge.

(Refer Slide Time: 07:11)



So this is the very interesting slide, so it may cost about 1 billion or even 1.2 billion US dollars to launch one drug molecule in the market that is a new drug I am talking about, if it is already old and you want to use that particular drug for some other disease that is called repurposing. For example, aspirin, aspirin was originally introduced for fever then pain later on a aspirin is being used quite a lot for blood thinning purpose okay, that is called repurposing that will not cost you much.

So it takes about 12, 15 years and for 1 new drug, so it is a very long process as you can see very long process. We start with some computational studies here okay, that is the computational biology studies we decide on the disease, what type of disease you are looking? Are you looking at inflammation? And are you looking at cancer of a certain colorectal cancer or breast cancer? So which target I am going to look at okay, so I need to identify which target.

Because there could be many enzymes, proteins, molecules are involved, and your drug maybe working only one particular target okay, so I try to identify which target. Once you have identified the target you start designing molecules and that is what is called a lead identification okay, once you have decided go for lead identification. So there are a lot of techniques involved if you want to look at the disease mechanism.

How does say an inflammation progress starting from the site of inflammation right down to various leukotrienes or prostaglandins, so what are the various enzymes involved, how is the metabolic pathway, so I can use computational biology type of approach, draw a big ne2rk of

various pathways, and then I can see how the flux flows, so that is called a computational biology.

If I want to identify the protein okay if I want to identify the protein, then I may use proteomics tools, we will talk little bit on proteomics, but I will talk too much on that trying to identify the protein structure, 3 dimensional structure its function and its active site that is all called proteomics, you may have to use some bioinformatics tools, because we may have to compare the protein which you have isolated for your disease, are there similar proteins available the databases.

What are the properties of those proteins, so I can connect with the new unknown protein okay that is called bioinformatics, then the actual discovery of the lead which involves molecular modeling, quantitative structure activity relationship, QSAR means quantitative structure activity relationship. We are going to talk quite a lot about this particular portion okay, and then we need to do the wet lab experiments that is experiment in your lab okay.

That is where you check the biological activity of your compound, you may use bacteria, you may use virus, you may use animal cells whatever is your biological assessment. Simultaneously, I need to understand the properties of the compound the lead drug likeness does it have the drug likeness properties, that means after all it is going to be consumed by human, so it should not have certain problems.

Or it should have certain properties like good solubility, good absorption okay that is called drug likeness property. Does it have toxicity short term toxicity long term, how is it ADME we keep on introducing this term ADME it is going to come quite often now, does it have good ADME properties? So simultaneously, when we are looking at possible candidates and testing it out in the wet lab biological assays I need to understand these things also, this is very, very important.

Because many drugs may have good activity but they may be very toxic, many leads may have very good activity but in the stomach it gets degraded, because as you know the stomach pH is extremely acidic pH of 2, it may have good activity maybe it is toxic, but may have good activity or maybe it gets metabolized inside by liver and the other enzymes involved, so the active drug concentration may be very, very less at the target site.

So the ADME could be very poor, the absorption could be poor, the distribution inside could be very poor, and it gets metabolized so concentration is very poor, or it does not get excreted from the body, so it keeps staying inside it gets accumulated and over long period of time the concentration maybe very high which may be toxic. For example, if I am using nanoparticles metal nanoparticles they may stay inside the body get absorbed by the tissues, because they are nano size.

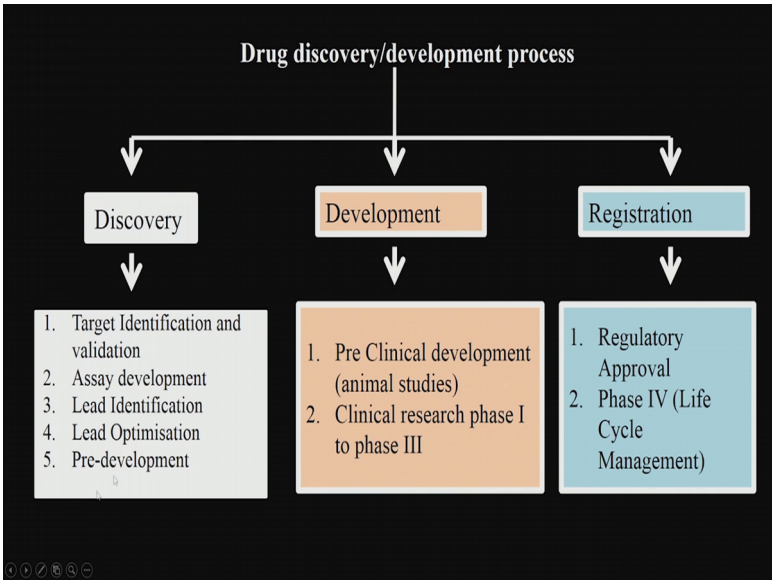
And they have metal toxicity that is why now use of metal nanoparticles in the body there is a lot of worry, because of this particular problem. So simultaneously when we look at lead an active molecule, we may have to simultaneously look at all these I may have to perform experiments to determine all these parameters okay, or I can use computational tools to predict some of the drug likeness property toxicity, ADME and so on actually.

So the course is predominantly going to cover this molecular modeling QSAR, course is going to cover prominently some of these computational approaches for determining ADME and drug likeness property okay. So of course simultaneously one has to see how to manufacture the drug in large quantities that is called process development bio-process, and it should have the good manufacturing practices approved by FDA and so on.

Once a lead is identified it goes through the animal preclinical trials then clinical trial 1, 2, 3 and then FDA approval and finally gets launched okay. So this is the pipeline for new drug discovery, so here we are talking quite a lot about computational approaches computational biology, proteomics, bioinformatics here we have the molecular modeling, and then here we have the drug likeness, ADME prediction okay.

So the course he is going to cover these 2 circles how computers can be used for measuring not measuring calculating some of the parameters that is what we are going to talk about okay.

(Refer Slide Time: 15:00)



So the drug discovery development process we can call it the discovery stage, the development stage, the registration stage okay. Discovery like I said target identification, validation we need to develop the assay, for example if I am going to study how my drug is going to bind to a protein or enzyme and inhibit I should have a biochemical assay okay, maybe a fluorescence assay or calorimetry assay okay.

So I need to or a radioactive assay or I may use an animal cell, I may be looking at some metabolites that are produced. So assay development is also very, very important, then once I identify my lead I need to optimize the lead, that means it is the balance between activity vis-a-vis its properties okay, so that is called lead optimization. Then the pre development. Of course the development stage we have the animal studies and clinical studies.

And then of course registration regulatory approval looking at the life cycle, sometimes drugs after being introduced into the market may be withdrawn, because certain things which have not been thought of it may have been having problems actually, that is why phase 4 is like getting feedback from the people who have been using the drug globally okay or in particular continent, and then see if there are any problems which has not been thought of okay.

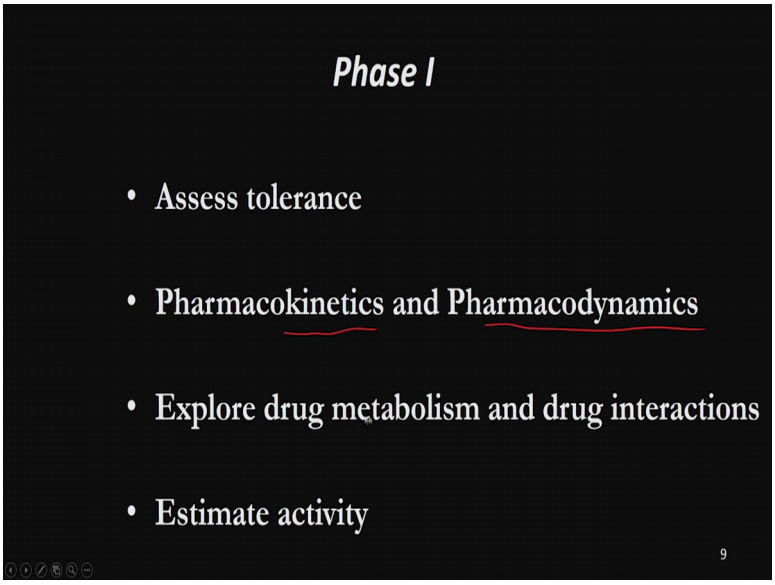
For example, there are many drugs, we have the anti-inflammatory, selective Cox 2 inhibitor which was introduced into the market, and then it was withdrawn because it had cardiovascular issues on some of those patients who took that okay. So phase 4 is also very important which is like a follow up okay, but as I said our focus is more on the side of it and nothing to do on that side okay.

(Refer Slide Time: 17:04)

So but still we need to understand little bit on phase 1, phase 2, phase 3 and phase 4. So there are 4 different phases of operation phase 1 involves human pharmacology, phase 2 involves therapeutic exploratory, phase 3 involves therapeutic confirmatory, phase 4 like I said post marketing after it has been marketed. The phase 1 we are looking at any side effects that may be caused by the drug when it is given to healthy volunteers.

And phase 2 is looking at what is the dose response okay, if I give 1 milligram what is the response. For example, it is bacterial like an antibiotic, if I give 1 milligram how much bacteria is killed, if I give 2 milligram how much bacteria is killed, so that is called therapeutic exploratory. Then in phase 3 we are looking at long term side effects that is why it is done for 3 years and so on actually okay.

(Refer Slide Time: 18:04)

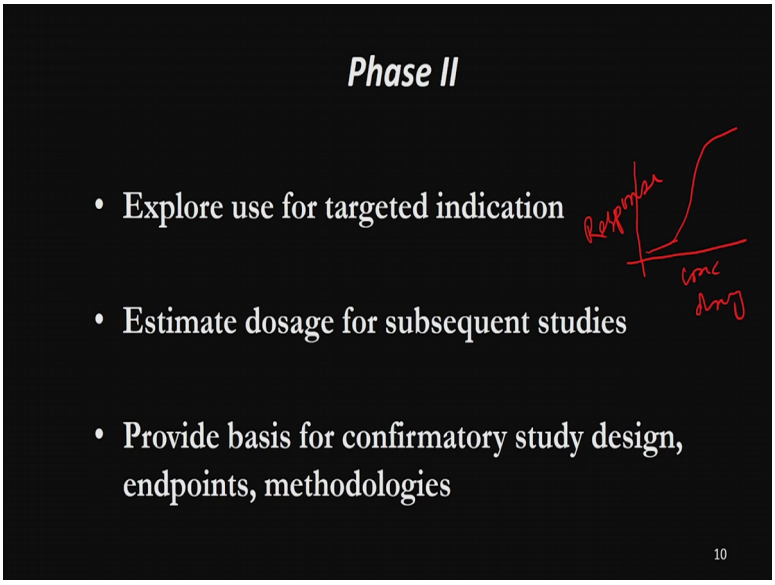


So phase 1 we are looking at tolerance, we are looking at pharmacokinetics and pharmacodynamics okay that is a very important thing that happens. Pharmacokinetics is when the drug is given to a patient gets a little bit absorbed then it gets excreted so in how many hours does it get excreted. What is the maximum concentration of drug in the body okay because of the metabolism and absorption?

So that sort of parameters is determined in pharmacokinetics okay, so you may take a sample either from the blood or target site and see what is the concentration of the drug and so on. In pharmacodynamics what the drug does to the patient or the target? So if I the drug is there does the bacteria go down 10%, does it go down 20% or if you are looking at a tumour the tumour size goes down as a function of drug concentration.

So pharmacokinetics is what the body does to the drug, pharmacodynamics is what the drug does to the body okay. You also look at drug metabolism and drug interactions, and also estimate the activity.

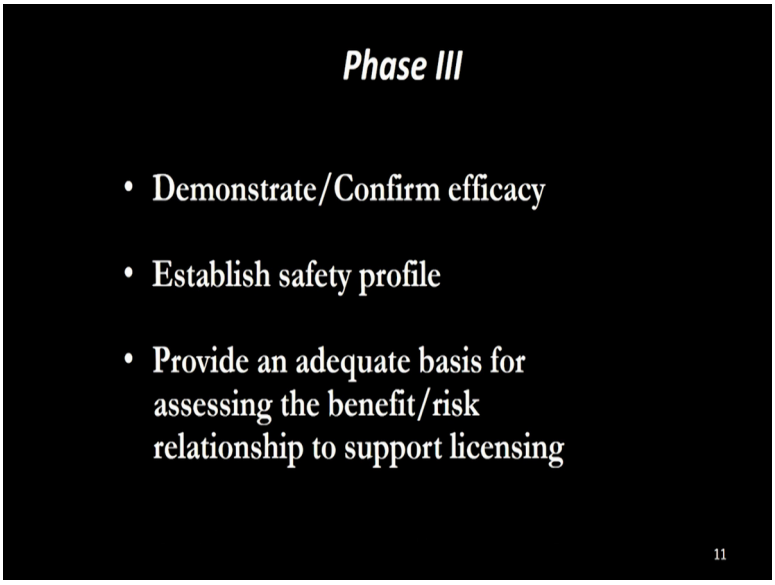
(Refer Slide Time: 19:24)



In a phase 2, you look at where the drug goes on acts, so you try to look at whether there is up-regulation and down-regulation of the target, what is the dosage require to kill a particular amount of bacteria or reduce the activity of some proteins? Okay. So we developed something called dose response curve, so if this is concentration of drug what is the response? So initially for some concentration they do not be much in drug, the fever goes down by 1 degree Fahrenheit.

If I give 2 milligram fever goes down by 5 degree Fahrenheit, so that is called a dose response curve okay, that is what you measure in the phase 2. This also helps you to identify the endpoints okay, when do you stop giving the drug okay, so do I stop and when the temperature comes down to somewhat certain values, do I stop for an infection when the amount of infection or bacteria is less than 5000 like that is called the endpoints okay. So all these are measured in phase 2.

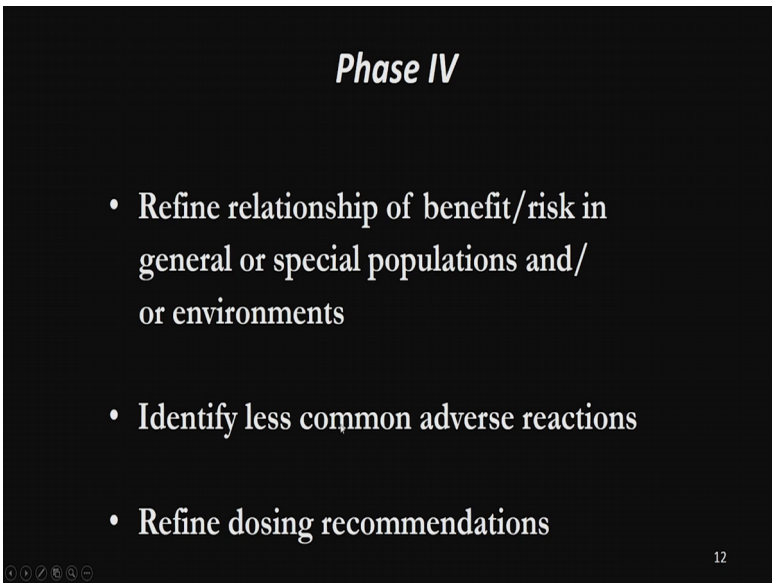
(Refer Slide Time: 20:37)



Then the phase 3 we are looking at safety profiles that is a very, very important point in phase 3 the safety profiles, look at the benefit of the versus risk. Because there are cancer anti-cancer drugs which could be very toxic to the healthy cells, but at the same time they may be killing the cancerous cells, so we look at the benefits versus risk okay. If there is terminally ill patient and the patient will not survive more than six months.

So by giving the drug which may be toxic patient may survive for 2 years, so it is worth it, so that sort of studies we do actually okay, we do all these things in phase 3.

(Refer Slide Time: 21:18)

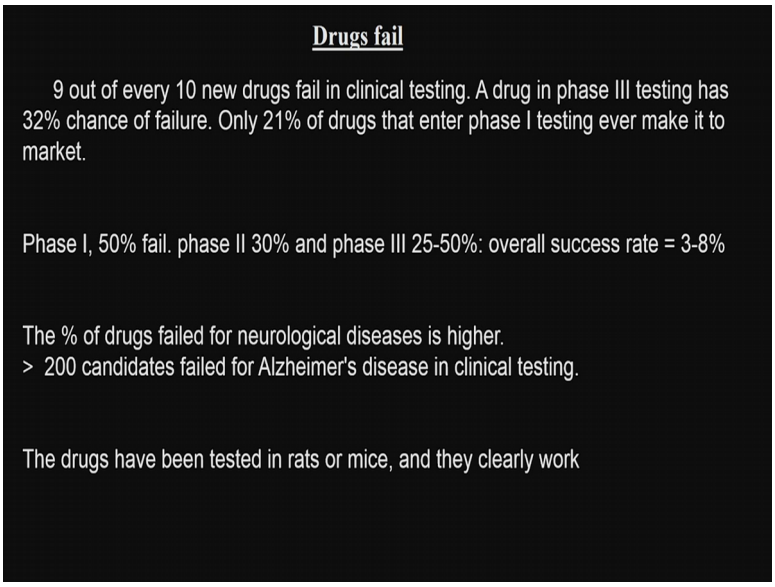


Then of course phase 4 is after it is being marketed, is there are any other adverse reactions do I have to again give some warnings to people who are taking the drug after all in the phase 1, phase 2, phase 3 trials the drug is tested only for about 1000 volunteers, but then when it is

given to public at large there would be some new reactions which has not been thought of okay, so that is done in phase 4 trials okay.

So drugs fail, lot of drugs fail very little lead compounds get converted as a new drug enters market lot of drugs fail okay, why do they fail? Okay.

(Refer Slide Time: 22:04)



As you can see 9 out of every 10 new drugs failed in clinical testing okay, 9 out of 10 that means the success is 10% or even less. The drug in phase 3 testing has 32% chance of failure, that means it is crossed phase 1, it crossed phase 2, but still it can fail because it has crossed phase 1 and phase 2, the percentage failure rate is much lower when compared to this right, here the success is only 10% or 90% failure.

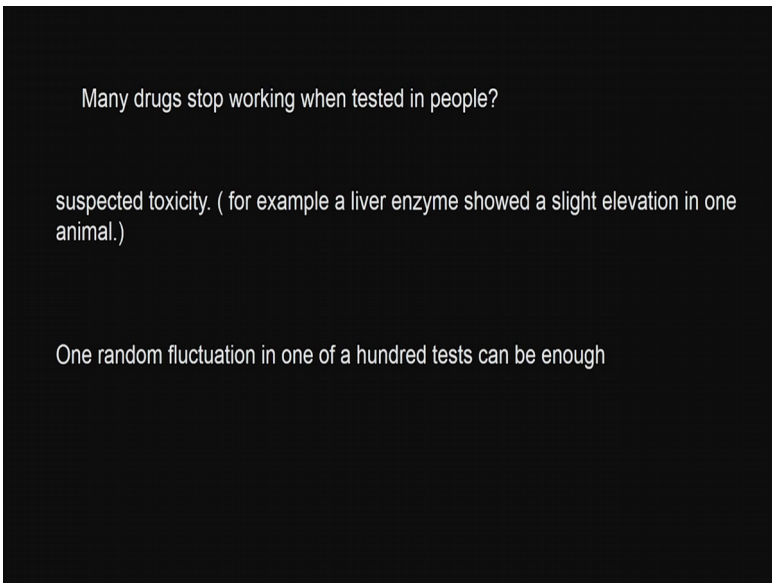
Once it has crossed phase 3 success the failure percentages is only 30%, because at each phase the pharma companies spend millions of dollars. So if the drug fails and it has to be withdrawn they have lost a lot of money. So only 20% of drugs that enter phase 1 make it to the market, that means remaining 80% failed, so whatever money they have spent goes down the drain. So that is why the computation tools are being widely used.

So that we try to find out failure possible failure compounds and do not take it further. Phase 1 50% fail, phase 2 30% and phase 3 25 to 50%, so overall success rate is 3 to 8%, so like I said here 10% okay. And the percentage of drugs failed for neurological disease is higher okay, because neurological diseases are much complicated, the drug has to enter the brain region and then do its job.

Whereas remaining drugs like inflammation or cancer or stomach pain or fever they do not have to go to the brain region okay, so the requirements are very different so that is why percentage of drugs failed for neurological disease is higher. 200 candidates failed for Alzheimer disease in clinical testing okay, Alzheimer's disease is more to do with the loss of memory and so on actually okay.

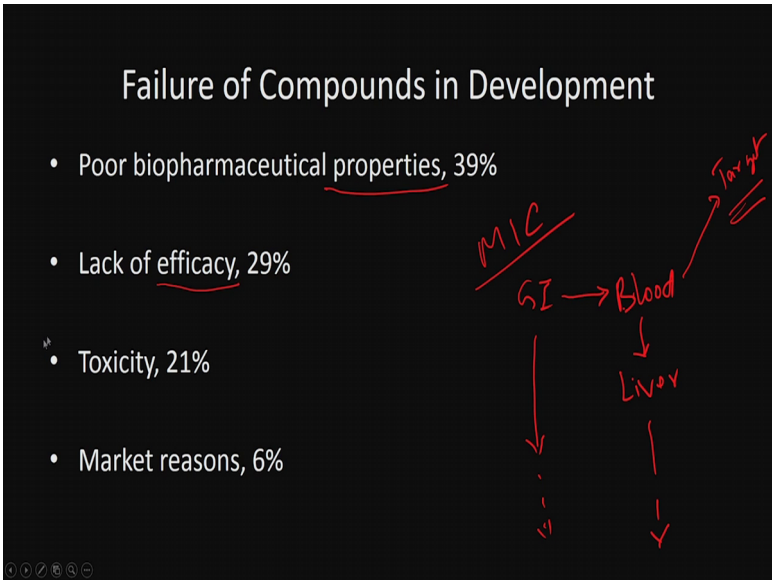
But they are being tested in rats or mice that means preclinical trials they work, but when it comes to human volunteer trial it fails okay, of course there is a difference between the system of rats, mice versus human. But still rats and mice is used widely in preclinical trials, because they are the closest to human, and they are much cheaper okay.

(Refer Slide Time: 24:51)



So many drugs stopped working and tested in people okay, why is that? Maybe toxicity, maybe when it was tested in animal slight elevation of some liver enzyme, which was not taken very seriously, but when it was given to human the enzyme level may be getting elevated too much. Even if there is one random fluctuation in 100s of test, then they decide that drug is not very safe and it has to be withdrawn okay, 1 in 100 okay, that is very, very tough business.

(Refer Slide Time: 25:28)



Poor biopharmaceutical properties like I mentioned solubility is very poor, maybe absorption is poor, metabolism is happening either in the stomach or in the liver region, because of various enzymes present okay, it is not getting excreted properly or it gets excreted very fast, it is not it is getting distributed too much that means the concentration of the drug in the blood is so low, all these 39%.

Lack of efficacy, that means the concentration at the target site is not sufficient okay, imagine I have a pain in my finger and I take a drug called Ibuprofen, I am sure all of us would have taken Ibuprofen okay. So the drug is taken orally, so we have the GI that is the gastrointestinal, then it gets absorbed into the blood okay, then you also have these liver which keeps degrading this okay, then it which is wasted, and then finally it reaches the target.

If it does not get absorbed properly from GI again it gets wasted through the fichus, if the liver degrades this it comes out through the urine, again it is waste so finally it reaches the target okay. So the concentration of the drug in the target is much, much less than the concentration of the drug taken by us orally okay. So it might not be sufficient to perform its duty. For example, if I am looking at an antibacterial drug, the concentration of the drug should be sufficiently larger than the minimum inhibitory concentration of the bacteria it is called MIC.

So the concentration of the drug in the target site is very, very low, that drug is of no use that is called lack of efficacy okay understand. Then toxicity the drug has toxicity or the

metabolites, the drug gets degraded in the liver maybe those metabolites are toxic, short term toxicity long term toxicity, so 21%. So you see if you add all these 39, 29, 21 it is a huge number okay, and so the properties of the drug are very, very important okay.

The ADME properties, the efficacy as we can see which I explained here, because of lack of efficacy, then of course toxicity many, many drugs fail during the development process either preclinical clinical trials actually. So one need to put in a lot of focus in this area not only just looking at the activity in my lab, anti-cancer activity is showing very good activity against certain cell lines, but I need to also understand the physicochemical properties of the molecule.

And also so that I can design a molecule which satisfies all these condition in addition to good activity, so that is very, very important. So a drug is not only doing very well in the lab because it is showing a very good activity, but it should also have all the properties okay the biopharmaceutical properties, and also this good efficacy, so that it passes the clinical trials and it gets approved okay.

That is why we are going to spend a lot of time on these aspects as well, and not only just looking at a candidate which shows very good activity or which has very good inhibitory power against an enzyme or a protein and so on. And of course toxicity is another big issue, one needs to do lot of experimental studies on fish, on animals and so on to identify toxicity, and there are different computational tools also which can help you to predict toxicity of compounds, the metabolites and so on okay.

So in the next few weeks, we are going to talk more about this, before we actually jump into the predicting the activity of various lead molecules okay. So we will talk more in the next class, thank you very much for your time.