

# **An introduction to pathway modelling with Excel**

Dr Sean Manzi  
PenCLARHC/PenCHORD  
University of Exeter

# What is pathway modelling and why is it useful?

- What happens when you arrive as a patient to an Emergency Department?

# What is pathway modelling and why is it useful?

- What you have just described is a care pathway

# What is pathway modelling and why is it useful?

- What if we want to change something in that pathway?
- How can we know what the impact of that change might be?

What is pathway modelling and why is it useful?

We can model it!

# What is pathway modelling and why is it useful?

A model is a simplification of reality

**Warning!!!!**

All models are wrong but some are useful

# What is pathway modelling and why is it useful?

## Benefits of modelling

- **Emulation** : A model is a version of reality that can be altered without risk or consequence
- **Speed** : Typically, models can be designed and built much more quickly than real world changes can be effected.
- **Communication** : A model can help people to communicate about a problem using a shared language and point of reference
- **Systems Thinking** : The process of designing the model can help people to think about their systems

# What is pathway modelling and why is it useful?

## **Benefits of pathway modelling**

- Identify queues/blockages in the system
- Determine resource requirements
- Test new pathway configurations



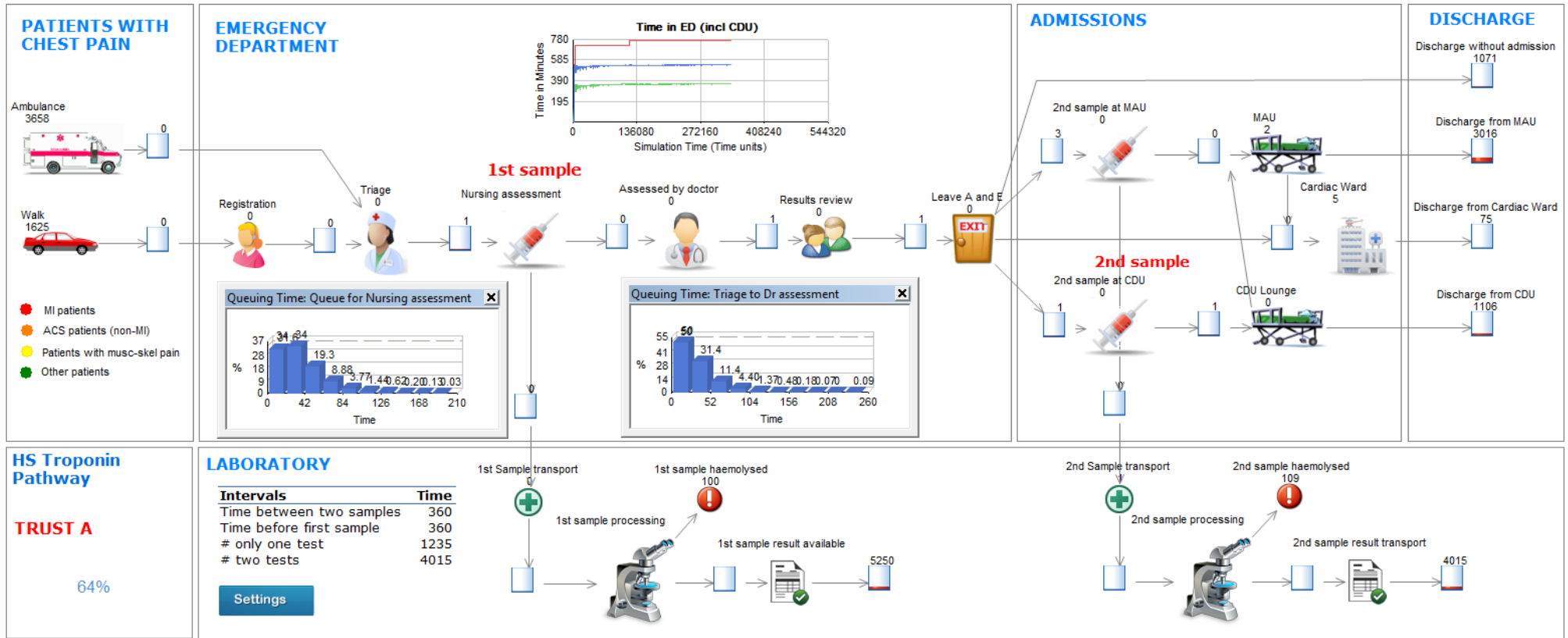
# What is pathway modelling and why is it useful?

## **Pathway modelling and economic modelling**

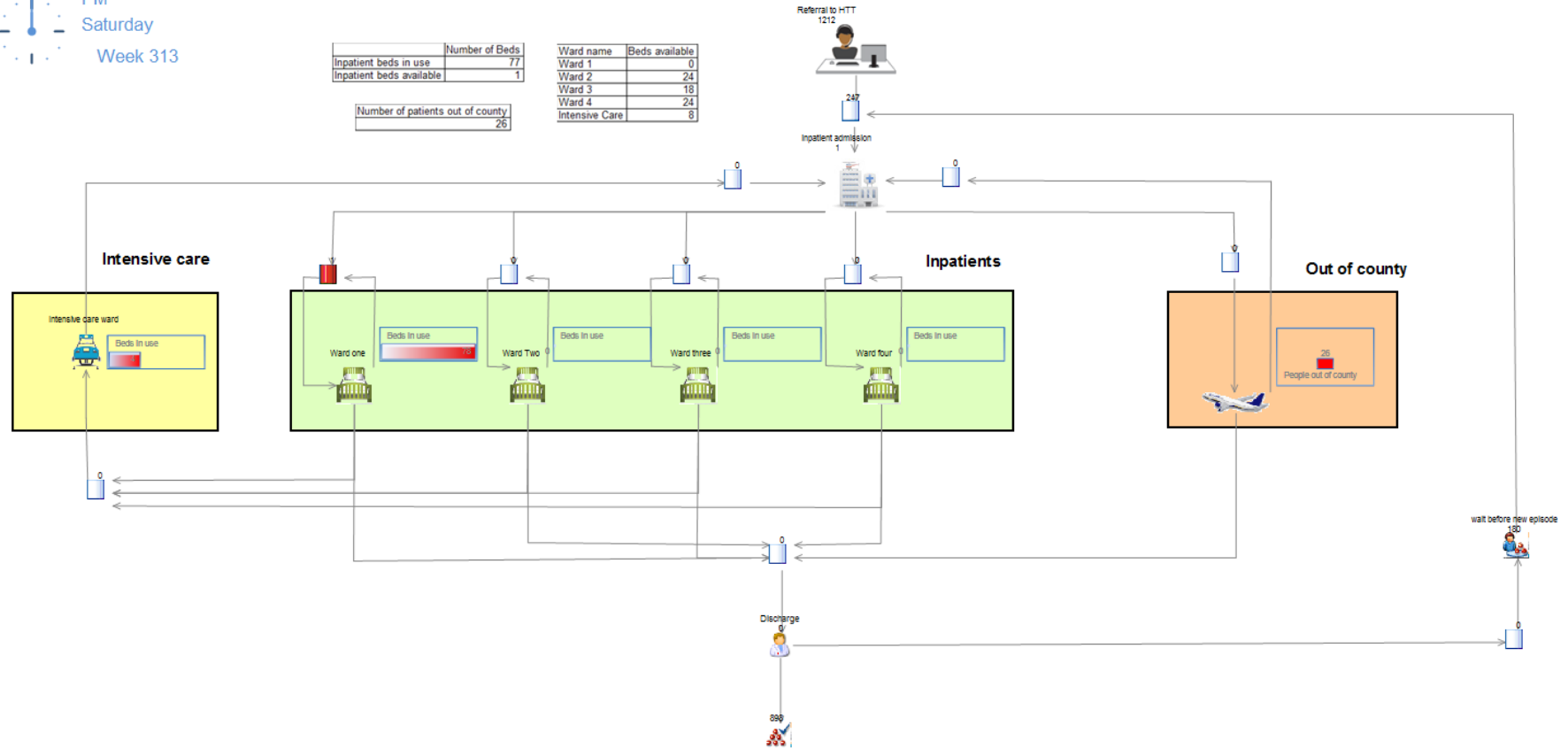
- Understanding the scale of change to service provision
- Estimation of efficiency savings
- Supporting a financial case for capital investment

# What is pathway modelling and why is it useful?

**Lets have a look at some real examples of pathway modelling!**

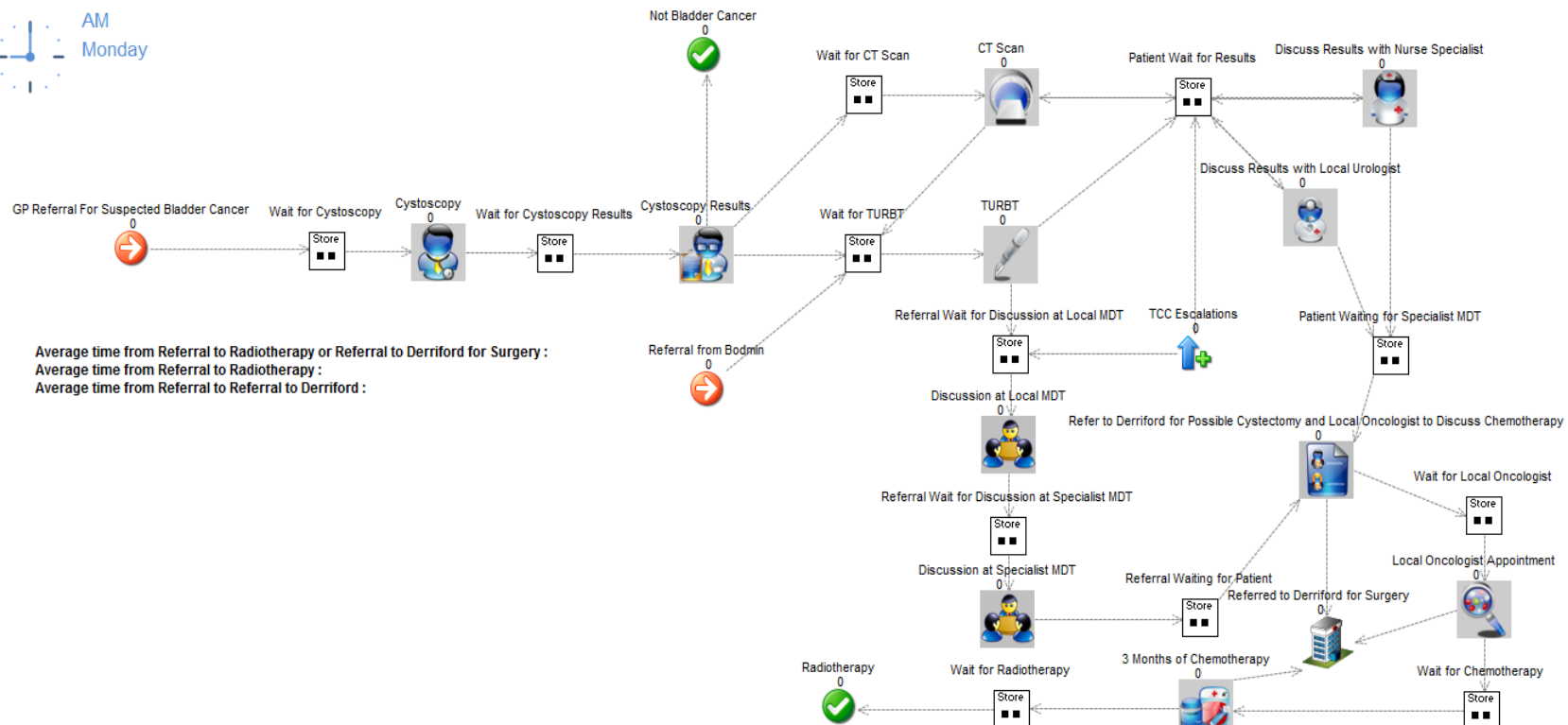


# High-sensitivity troponin for the diagnosis of acute myocardial infarction in A&E

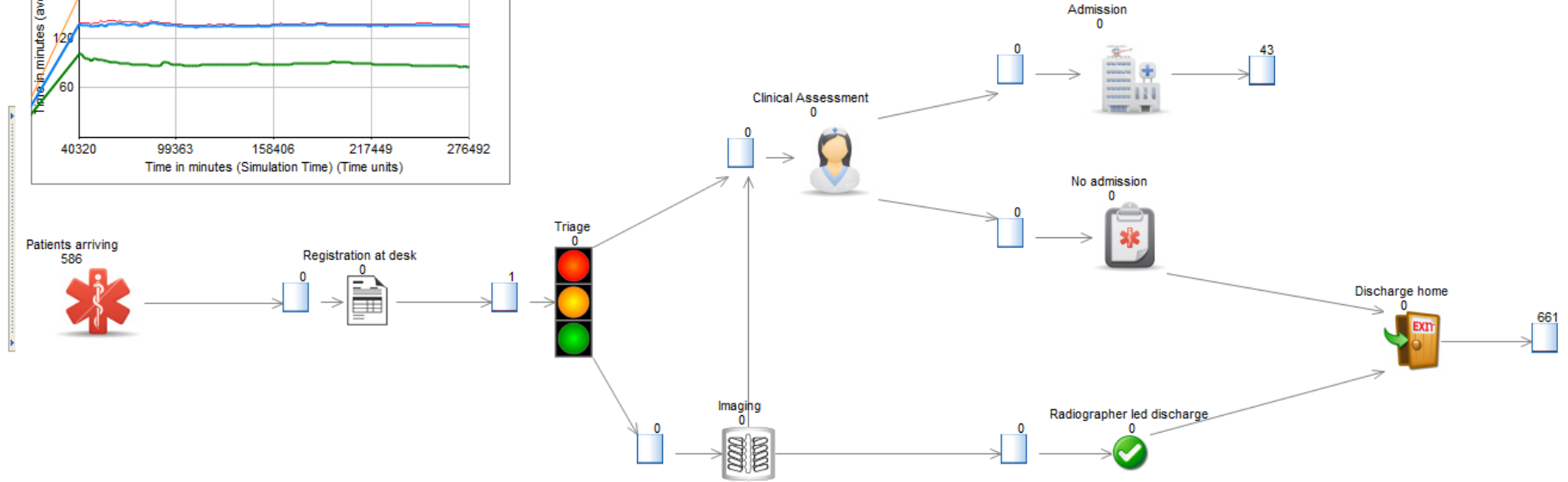
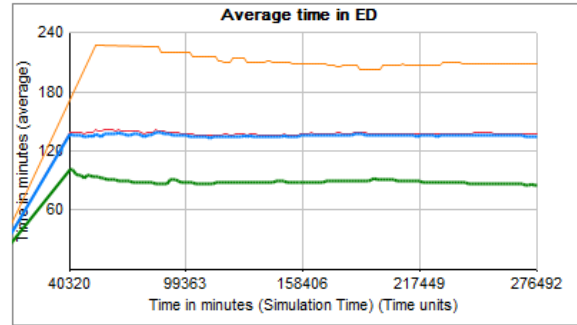


# Cornwall mental health inpatient pathway modelling

AM  
Monday



# Reducing delays in the bladder cancer pathway at Royal Cornwall Hospital



# Radiographer-led discharge at Yeovil District Hospital

# Modelling healthcare pathways in Excel

Today we are going to:

- Understand some important concepts for pathway modelling in Excel
- Understand and process map our problem
- Build a base case model
- Test a change
- If there is time: think about even more important concepts for pathway modelling

# Modelling healthcare pathways in Excel

- Important concepts for today
  - Functions in Excel
  - Conditional logic
  - Variability and distributions



# Important concepts - functions

- **Functions take zero to many inputs (“arguments”), do something with them, and generate an output.**
- **In Excel, functions are denoted as follows :**
- **=FUNCTIONNAME(argument1, argument2, [argument3])**

# Important concepts - functions

- **The arguments in a function can be either :**
  - **Absolute values (such as 2, “hello”, 37)**
  - **Cell references (such as A2, B7:B28, C:C)**
  - **Functions - these are called nested functions**

# Important concepts - functions

- When a function is used as an argument to another function, the internal function is performed to generate an output. This output is then used as the argument.
- For example :

**=SUM(MIN(A:A),AVERAGE(B:B))**

- This SUM function has two nested functions - MIN() and AVERAGE(). Here, the minimum of column A is found, then the average of column B, then these two numbers are used as arguments for the SUM function, and so are added together.

# Important concepts - functions

- **=RAND()** - generates a pseudo-random number
- **=ROUND(Number to round, number of decimal places)** - rounds up a number to a specified number of decimal places
- **=COUNT()** - which counts the number of non-empty cells in a range
- **=SUM()** - which adds up a series of values
- **=AVERAGE()** - which takes the mean of a series of values

# Important concepts – conditional logic

- **Modelling often requires different things to happen depending on whether or not something is true / has happened.**
- **For example, a model may check a patient's condition, and allocate them to a doctor accordingly. Or a model might check the time of day to determine how frequently patients arrive into the model.**
- **This kind of logic is called “Conditional Logic”**

# Important concepts – conditional logic

- If we were writing a computer program, we would specify conditional logic using IF-THEN-ELSE statements :  
IF this is true THEN  
Do this  
ELSE  
Do this
- In Excel, we can use the IF Function to write such statements :  
=IF(logical test, value if true, value if false)
- Which is another way of saying :  
=IF(check if this is true, put this value in the cell if it is, put this value in the cell if it isn't)

# Important concepts – conditional logic

- **Let's consider an example.**

**=IF(B7>=65, "Over 65", "Under 65")**

- **This function checks the value in cell B7 to see if it contains a number that is equal to or greater than 65. If it does, the value in the cell in which you wrote the function will be "Over 65". If it doesn't, it will display "Under 65" instead.**

# Important concepts – conditional logic

- **COUNTIFS()** counts the number of values that meet a certain criteria. It has the following format :

**=COUNTIFS(criteria range 1, criteria 1, [criteria range 2], [criteria 2]...)**

- **E.g. =COUNTIFS(A:A, “Female”)**



# Important concepts – conditional logic

- **SUMIFS()** only adds up numbers that meet specified criteria. It has the following format :

**=SUMIFS(sum range, criteria range 1, criteria 1, [criteria range 2], [criteria 2]...)**

- **E.g. =SUMIFS(B:B, A:A, “Male”)**

# Important concepts – conditional logic

- **AVERAGEIFS()** takes an average but only includes values where a corresponding field meets a certain criteria. It has the following format :

**=SUMIFS(average range, criteria range 1, criteria 1, [criteria range 2], [criteria 2]...)**

- E.g. **=AVERAGEIFS(B:B, C:C, “Elderly male”)**

# Important concepts – conditional logic

- If we want to look for values greater than (>) or less than (<) a given value in these conditional functions we have to use a certain notation
- For example :  
=COUNTIFS(A:A, ">="&5)  
=SUMIFS(B:B, A:A, "<"&B6)
- The operand is in inverted commas : ">="
- The ampersand (&) denotes the concatenation (attaching the operand to the value)
- An IF() function doesn't need this, as it only checks a single value.

# Important concepts – variability

## Scenario A

Treatment  
duration  
Inter-arrival  
time

5	5	5	5	5	<b>Average</b> 5
5	5	5	5		5

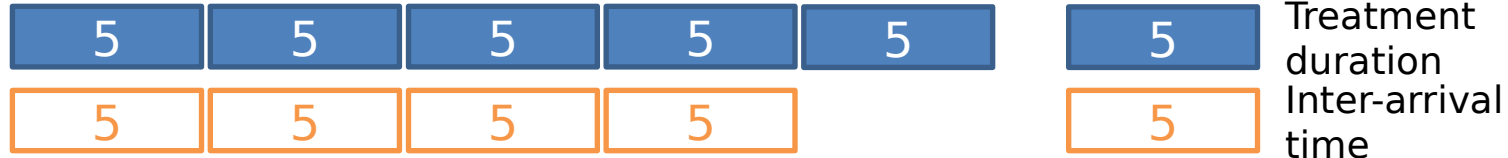
## Scenario B

Treatment  
duration  
Inter-arrival  
time

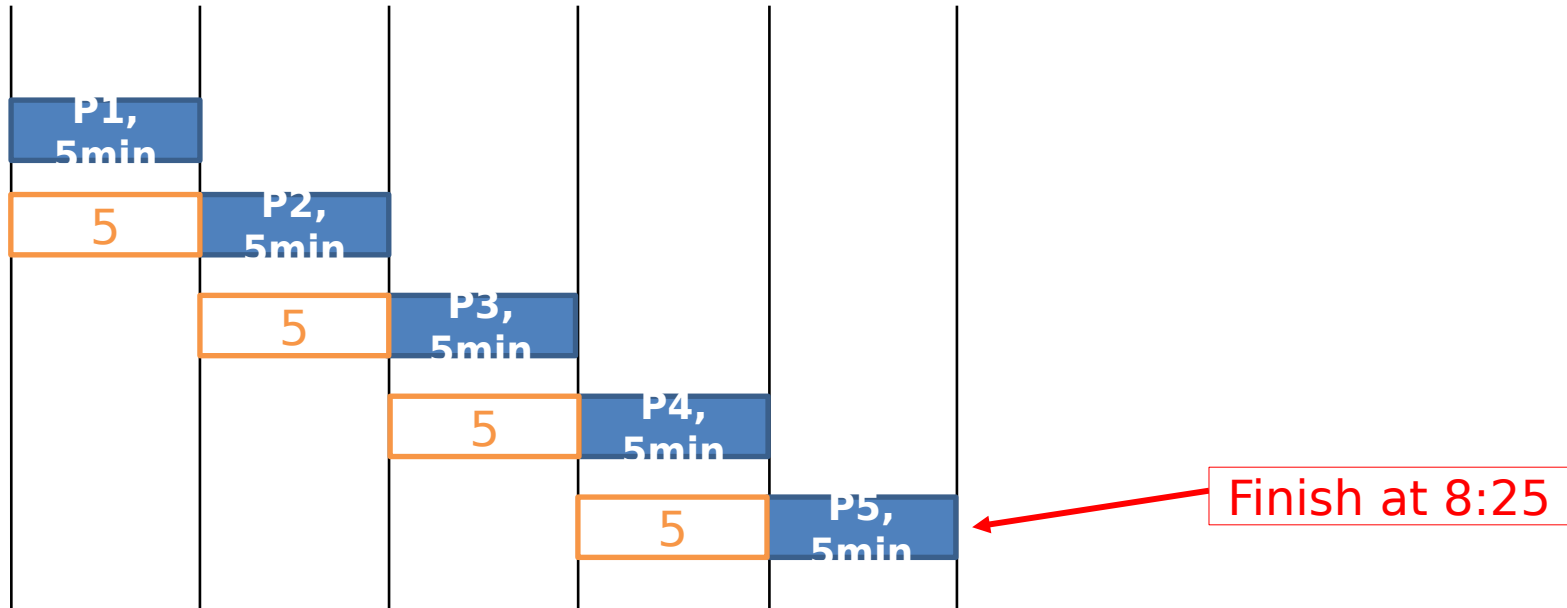
5	5	5	4	6	<b>Average</b> 5
5	6	4	5		5

# Important concepts – variability

## Scenario A



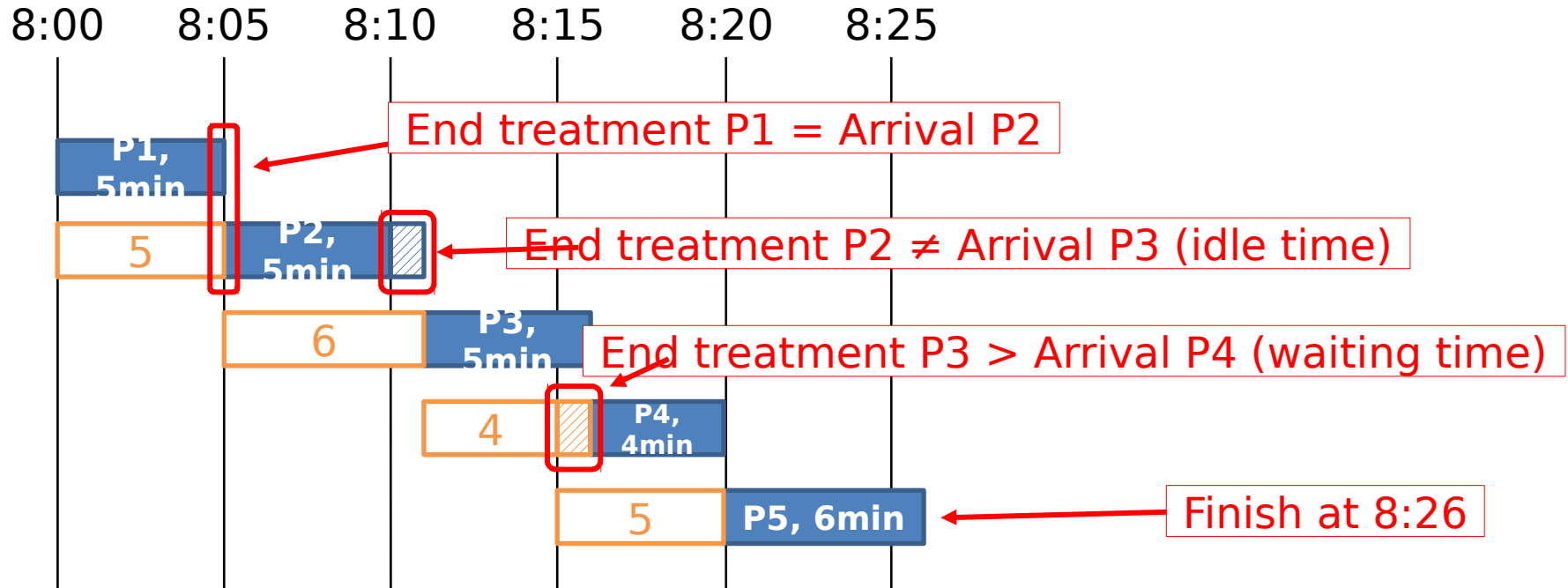
8:00      8:05      8:10      8:15      8:20      8:25



# Important concepts – variability

## Scenario B

## Average



# Important concepts – variability

**Real life is full of variability. And if our models are to be as accurate as possible, they should account for this variability as much as they can.**

**To do this, we can use **distributions****

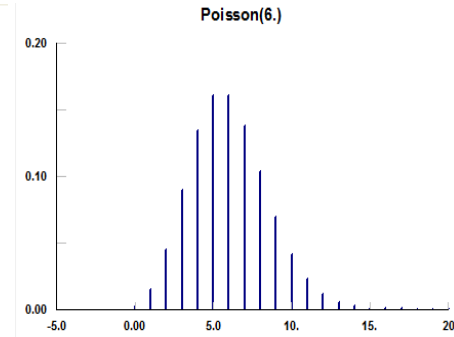
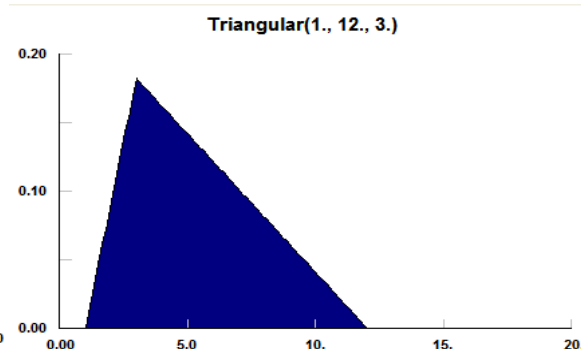
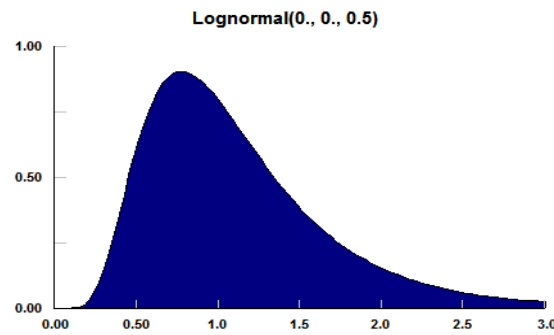
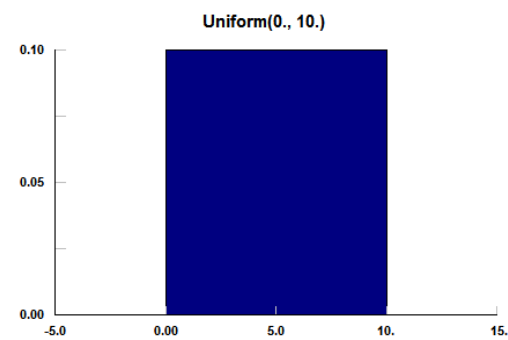
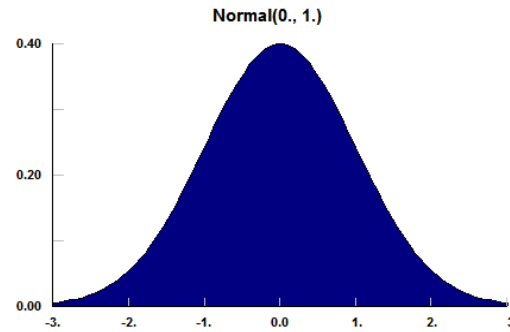
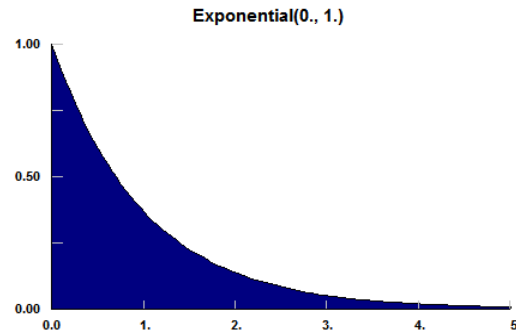
# Important concepts – distributions

- A distribution is a way of representing the variability within data. It tells us the probability of a value occurring in the future
- The RAND function in Excel draws random numbers from a uniform distribution. This means that all values have an equal probability of occurring.
- When building a model you would plot your data to understand its shape then use distribution fitting techniques to determine the most appropriate distribution and values (not covered here see [pythonhealthcare.org](http://pythonhealthcare.org) for distribution fitting code)

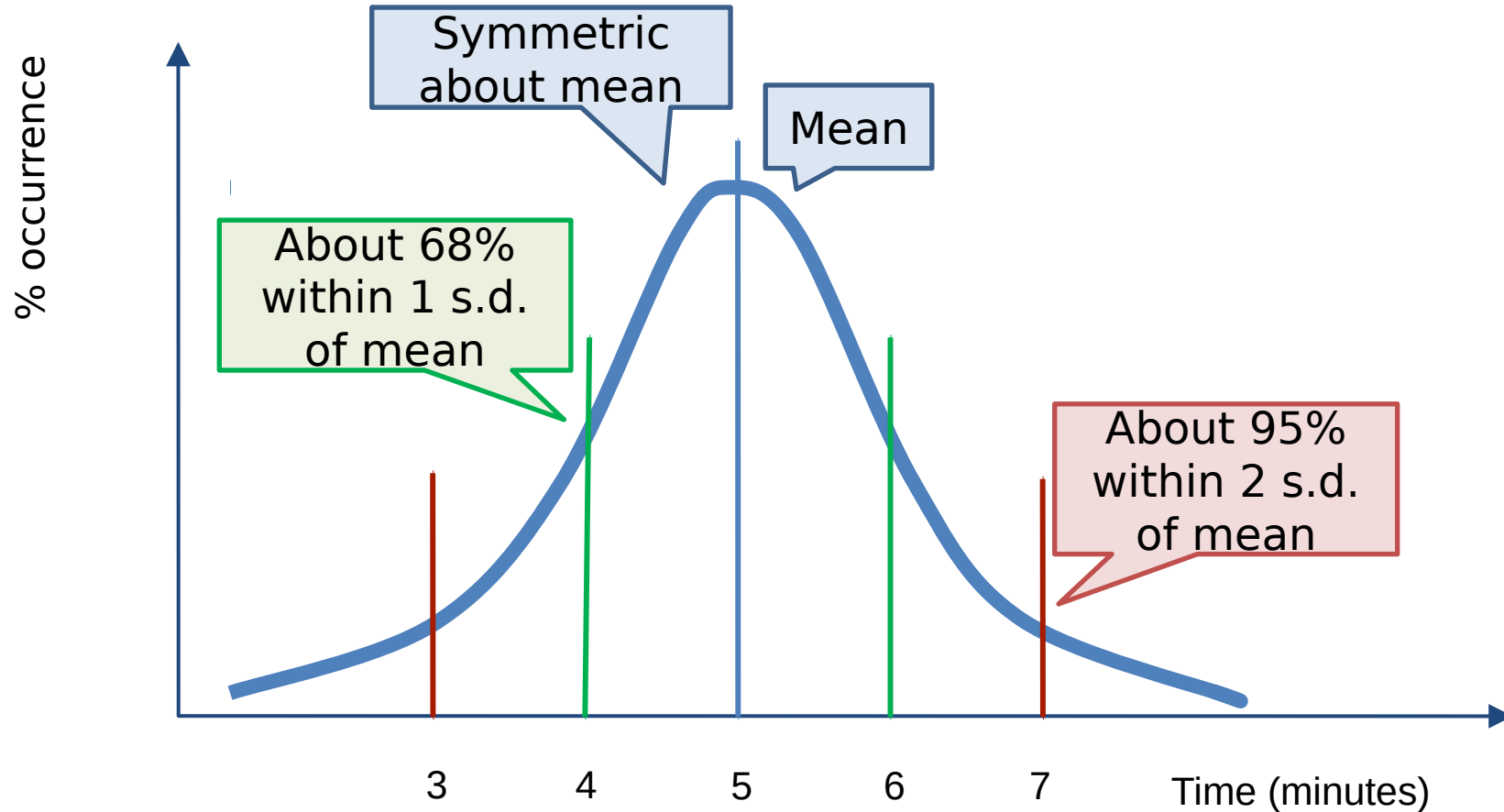


# Important concepts – distributions

- There are many different types of distributions



# Important concepts – distributions



# Important concepts – distributions

**In Excel we need an add called the Analysis ToolPak to generate random numbers from different distributions**

- Go to File then Options
- Select Add-Ins on the left-hand menu bar
- Ensure “Excel Add-ins” is selected next to “Manage” at the bottom of the window, and click “Go”
- Tick the check box next to “Analysis ToolPak”
- Click “OK”
- The ToolPak should now be installed

# Important concepts – distributions

To access the random number generator:

1. Click “Data Analysis” in the Analysis section of the Data tab on the ribbon
2. Select “Random Number Generation” and click OK

# Important concepts – distributions

The number of sets of random numbers you want

The number of random numbers you want to generate

The distribution from which you want to draw random numbers

The screenshot shows a 'Random Number Generation' dialog box with the following fields and options:

- Number of Variables:** A text input field with the value '1'.
- Number of Random Numbers:** A text input field with the value '1'.
- Distribution:** A dropdown menu set to 'Normal'.
- Parameters:**
  - Mean =** A text input field with the value '0'.
  - Standard deviation =** A text input field with the value '1'.
- Random Seed:** A text input field.
- Output options:**
  - ☐ **Output Range:** A text input field with a small icon to the right.
  - ☒ **New Worksheet Ply:** A text input field.
  - ☐ **New Workbook**

Buttons on the right: OK, Cancel, Help.

Annotations with arrows pointing to the dialog box:

- From 'The number of sets of random numbers you want' to the 'Number of Variables' field.
- From 'The number of random numbers you want to generate' to the 'Number of Random Numbers' field.
- From 'The distribution from which you want to draw random numbers' to the 'Distribution' dropdown.
- From 'The parameters needed to define your chosen distribution' to the 'Mean' and 'Standard deviation' fields.
- From 'The Random Seed' to the 'Random Seed' field.
- From 'Where do you want your random numbers to be output?' to the 'Output Range' option.

The parameters needed to define your chosen distribution

The Random Seed

Where do you want your random numbers to be output?

# Important concepts – distributions

- To define a distribution, we need to specify values that define the distribution.
- A normal distribution is defined by two values – the mean (because a Normal Distribution is centred around the mean) and the standard deviation (so we know how spread out the bell-shaped curve is).
- Often when you're modelling, the values you draw from a distribution represent times (e.g. inter-arrival times, times spent with a doctor etc).
- But in Excel, time is stored as a "Serial Number" (e.g. 0.7664) which Excel translates into a format in which we would understand time.

# Important concepts – distributions

- Why is this important? Because when you define the distribution, if your distribution represents time, you need to provide your values (e.g. mean and standard deviation) as serial numbers.
- Fortunately, the serial numbers aren't as obscure as they may first seem. In fact, they just represent the number of days that a time represents. So a serial number of 0.5 would represent 12 hours (half a day).
- Usually you'll be working in minutes when building models. So to convert your minutes to a serial number, simply use the following formula :
- **Serial Number = (time in minutes / 60) / 24**

# Modelling healthcare pathways in Excel

That is all the background information

**Let's build a model!!**



# The model – Our problem

- A local GP surgery has contacted you to undertake some modelling. The surgery is very small, and consists simply of a waiting room and a single GP who sees patients in the order in which they arrive. The GP starts seeing patients at 8am, and finishes at 5pm.
  - The surgery is looking to implement a new system that they believe will reduce consultation time with the doctor for minor injuries by an average of 20%.
- The surgery would like you to build a model of a day in the life of their surgery to help them answer the following questions :
  - How long, on average are patients waiting to be seen now? What's the longest waiting time?
  - What is the predicted impact on waiting times if the average consultation time for minor injuries was reduced by 20%?
  - What would be the impact on waiting times if a receptionist was hired to signpost patients arriving with minor injuries elsewhere, so they are not seen by the doctor?

# The model – Conceptual model

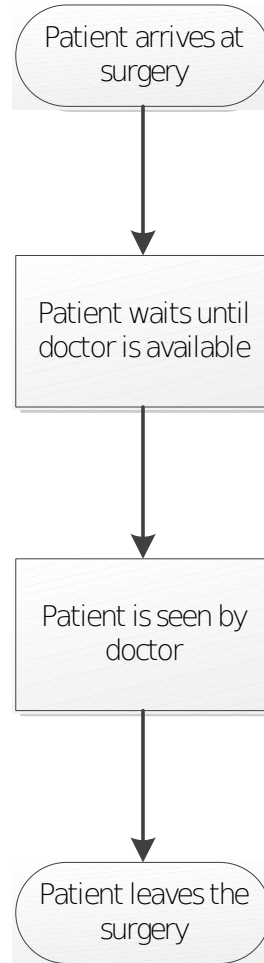
- Your first task is to develop a conceptual model. This is a design for your model that makes clear what you are going to include from the real world, as well as the assumptions and simplifications that you're making, and the data you're going to need. There is no set standard for developing this, but we'd recommend the following :
  - An annotated diagram or flowchart for each of the three systems (Base Case, What if 1, What if 2) and their processes
  - A list of assumptions and simplifications that you will need to make
  - A list of the specific data you will require to build the model
- Keep things as simple as possible. You will have 20 minutes to complete this exercise.

# The model – Conceptual model

Let's see what you have come up with

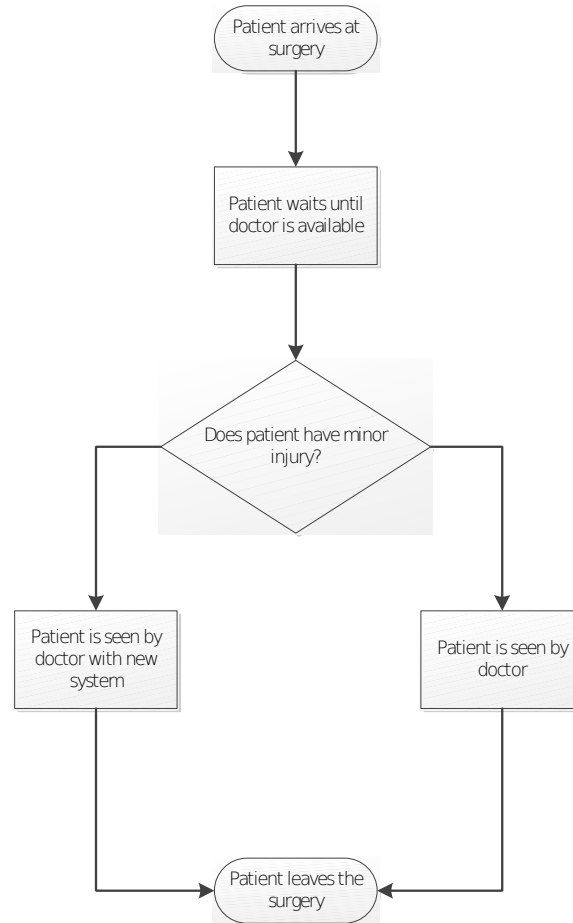
# The model – Conceptual model

- **Base case**



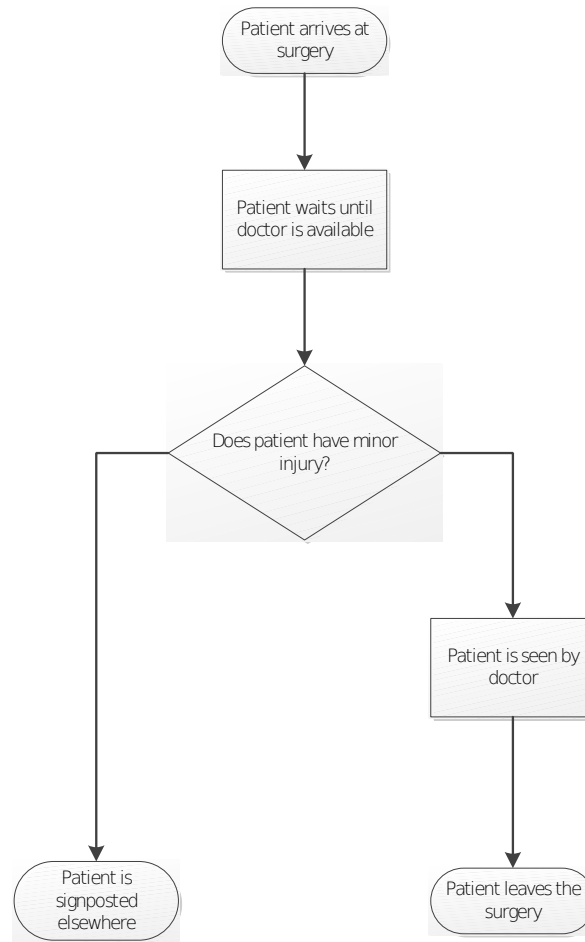
# The model – Conceptual model

- **Test of change 1**



# The model – Conceptual model

- **Test of change 2**



# The model – Conceptual model

## **Assumptions**

- There is no prioritisation of patients
- The day we are modelling is representative
- Travel times are trivial
- The surgery starts on time

## **Simplifications**

- The doctor doesn't get any breaks
- There is no check in time or queuing for check in
- The doctor continues working until the last patient has been seen
- The doors close at 5pm and patients aren't stopped from entering before

# The model – Conceptual model

**You have been supplied with the following data :**

## **Inter-Arrival Times**

- The times between patients arriving at the clinic follow a normal distribution. The mean is 5 minutes and the standard deviation is 1 minute.

## **Consultation Times with Doctor**

- The times patients spend with the doctor follow a normal distribution.
- For patients with minor injuries, the mean is 5 minutes and the standard deviation is 1 minute.
- For all other patients, the mean is 8 minutes and the standard deviation is 2 minutes.

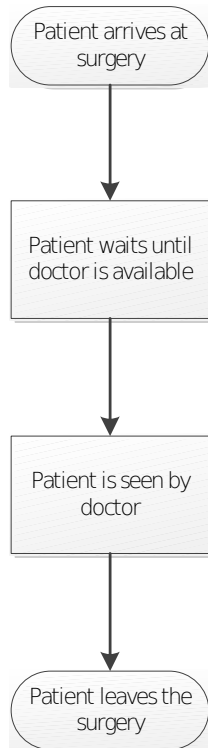
## **Presentations to the Surgery**

- 57% of patients attending present with minor injuries.



# The model – Building the model

We're now going to build the model together. Let's build the model for the base case first, and let's start by modelling only Minor Injuries patients.



**You have been supplied with the following data :**

## **Inter-Arrival Times**

The times between patients arriving at the clinic follow a normal distribution. The mean is 5 minutes and the standard deviation is 1 minute.

## **Consultation Times with Doctor**

The times patients spend with the doctor follow a normal distribution.

For patients with minor injuries, the mean is 5 minutes and the standard deviation is 1 minute.

For all other patients, the mean is 8 minutes and the standard deviation is 2 minutes.

## **Presentations to the Surgery**

57% of patients attending present with minor injuries.

# The model – The base case

We now have the answer to Question 1 :

**1. How long, on average are patients waiting to be seen now? What's the longest waiting time?**

Patient no.	Presentation Random Number	Minor Injury?	Inter-arrival time	Arrival time	Start consultation	Consultation time if Minor Injury	Consultation time if OTHER	End Consultation	Waiting time	Start time of surgery :	08:00		
1	0.235419782	Yes	03	08:03	08:03	04	05	08:08	00:00				
2	0.527878658	Yes	04	08:08	08:08	06	07	08:14	00:00				
3	0.070711386	Yes	04	08:12	08:14	05	09	08:19	00:02				
4	0.554216132	Yes	06	08:19	08:19	06	08	08:25	00:00				
5	0.149967956	Yes	06	08:25	08:25	04	08	08:30	00:00				
6	0.577135533	No	03	08:29	08:30	05	11	08:42	00:00				
7	0.093203528	Yes	04	08:34	08:42	03	05	08:45	00:07				
8	0.329538865	Yes	05	08:39	08:45	06	08	08:51	00:05				
9	0.234992523	Yes	05	08:45	08:51	04	08	08:56	00:06				
10	0.270607624	Yes	05	08:50	08:56	05	07	09:01	00:05				
11	0.622211371	No	03	08:54	09:01	04	10	09:12	00:07				
12	0.54893643	Yes	04	08:59	09:12	04	08	09:17	00:12				
13	0.253517258	Yes	04	09:03	09:17	02	05	09:20	00:13				
14	0.897885067	No	03	09:07	09:20	06	09	09:29	00:12				
15	0.553361614	Yes	05	09:12	09:29	06	12	09:35	00:16				
16	0.732871487	No	04	09:17	09:35	04	10	09:45	00:17				
17	0.521469771	Yes	05	09:22	09:45	05	08	09:51	00:23				
18	0.378734703	Yes	05	09:28	09:51	03	07	09:55	00:23				
19	0.024018067	Yes	03	09:31	09:55	05	09	10:01	00:24				
20	0.278389843	Yes	05	09:37	10:01	05	07	10:06	00:23				
21	0.92056032	No	06	09:43	10:06	05	07	10:13	00:22				
22	0.349986267	Yes	04	09:48	10:13	04	11	10:17	00:25				
23	0.778588214	No	04	09:52	10:17	06	07	10:25	00:25				
24	0.488021485	Yes	04	09:56	10:25	03	08	10:29	00:28				
25	0.436262093	Yes	05	10:02	10:29	05	06	10:35	00:26				
26	0.675801874	No	03	10:06	10:35	04	06	10:41	00:28				
27	0.262306589	Yes	03	10:09	10:41	05	08	10:46	00:31				
												Average Waiting Time :	00:48
												Longest Waiting Time :	01:55

# The model – Tests of change

**You now need to use your models to answer questions 2 and 3 :**

**2. What is the predicted impact on waiting times if the average consultation time for minor injuries was reduced by 20%?**

**3. What would be the impact on waiting times if a receptionist was hired to signpost patients arriving with minor injuries elsewhere, so they are not seen by the doctor?**

**Use your conceptual models to guide how you will do this. Take a copy of your base case model.**

# The model – Test of change 1

Patient no.	Presentation Random Number	Minor Injury?	Inter-arrival time	Arrival time	Start consultation	Consultation time if Minor Injury	Consultation time if OTHER	End Consultation	Waiting time		Start time of surgery :	08:00		
1	0.235419782	Yes	03	08:03	08:03	03	05	08:07	00:00					
2	0.527878658	Yes	04	08:08	08:08	05	07	08:13	00:00				Average Waiting Time :	00:17
3	0.070711386	Yes	04	08:12	08:13	04	09	08:17	00:01				Longest Waiting Time :	00:52
4	0.554216132	Yes	06	08:19	08:19	05	08	08:24	00:00					
5	0.149967956	Yes	06	08:25	08:25	03	08	08:29	00:00					
6	0.577135533	No	03	08:29	08:29	04	11	08:41	00:00					
7	0.093203528	Yes	04	08:34	08:41	02	05	08:43	00:06					
8	0.329538865	Yes	05	08:39	08:43	05	08	08:48	00:03					
9	0.234992523	Yes	05	08:45	08:48	03	08	08:52	00:03					
10	0.270607624	Yes	05	08:50	08:52	04	07	08:56	00:01					
11	0.622211371	No	03	08:54	08:56	03	10	09:07	00:02					
12	0.54893643	Yes	04	08:59	09:07	04	08	09:11	00:08					
13	0.253517258	Yes	04	09:03	09:11	02	05	09:13	00:08					
14	0.897885067	No	03	09:07	09:13	05	09	09:22	00:06					
15	0.553361614	Yes	05	09:12	09:22	05	12	09:27	00:09					
16	0.732871487	No	04	09:17	09:27	03	10	09:38	00:10					
17	0.521469771	Yes	05	09:22	09:38	04	08	09:42	00:15					
18	0.378734703	Yes	05	09:28	09:42	02	07	09:45	00:14					
19	0.024018067	Yes	03	09:31	09:45	04	09	09:50	00:14					
20	0.278389843	Yes	05	09:37	09:50	04	07	09:54	00:13					
21	0.92056032	No	06	09:43	09:54	04	07	10:02	00:11					
22	0.349986267	Yes	04	09:48	10:02	03	11	10:05	00:13					
23	0.778588214	No	04	09:52	10:05	05	07	10:13	00:12					
24	0.488021485	Yes	04	09:56	10:13	02	08	10:15	00:16					
25	0.436262093	Yes	05	10:02	10:15	05	06	10:20	00:13					
26	0.675801874	No	03	10:06	10:20	03	06	10:26	00:14					
27	0.263206588	Yes	03	10:09	10:26	04	08	10:31	00:17					

# The model – Test of change 2

Patient no.	Presentation Random Number	Minor Injury?	Inter-arrival time	Arrival time	Start consultation	Consultation time if Minor Injury	Consultation time if OTHER	End Consultation	Waiting time		Start time of surgery :	08:00		
1	0.235419782	Yes	03	08:03	08:03	00	05	08:03	00:00					
2	0.527878658	Yes	04	08:08	08:08	00	07	08:08	00:00				Average Waiting Time :	00:02
3	0.070711386	Yes	04	08:12	08:12	00	09	08:12	00:00				Longest Waiting Time :	00:16
4	0.554216132	Yes	06	08:19	08:19	00	08	08:19	00:00					
5	0.149967956	Yes	06	08:25	08:25	00	08	08:25	00:00					
6	0.577135533	No	03	08:29	08:29	00	11	08:41	00:00					
7	0.093203528	Yes	04	08:34	08:41	00	05	08:41	00:06					
8	0.329538865	Yes	05	08:39	08:41	00	08	08:41	00:01					
9	0.234992523	Yes	05	08:45	08:45	00	08	08:45	00:00					
10	0.270607624	Yes	05	08:50	08:50	00	07	08:50	00:00					
11	0.622211371	No	03	08:54	08:54	00	10	09:05	00:00					
12	0.54893643	Yes	04	08:59	09:05	00	08	09:05	00:05					
13	0.253517258	Yes	04	09:03	09:05	00	05	09:05	00:01					
14	0.897885067	No	03	09:07	09:07	00	09	09:16	00:00					
15	0.553361614	Yes	05	09:12	09:16	00	12	09:16	00:03					
16	0.732871487	No	04	09:17	09:17	00	10	09:27	00:00					
17	0.521469771	Yes	05	09:22	09:27	00	08	09:27	00:05					
18	0.378734703	Yes	05	09:28	09:28	00	07	09:28	00:00					
19	0.024018067	Yes	03	09:31	09:31	00	09	09:31	00:00					
20	0.278389843	Yes	05	09:37	09:37	00	07	09:37	00:00					
21	0.92056032	No	06	09:43	09:43	00	07	09:51	00:00					
22	0.349986267	Yes	04	09:48	09:51	00	11	09:51	00:02					
23	0.778588214	No	04	09:52	09:52	00	07	10:00	00:00					
24	0.488021485	Yes	04	09:56	10:00	00	08	10:00	00:03					
25	0.436262093	Yes	05	10:02	10:02	00	06	10:02	00:00					

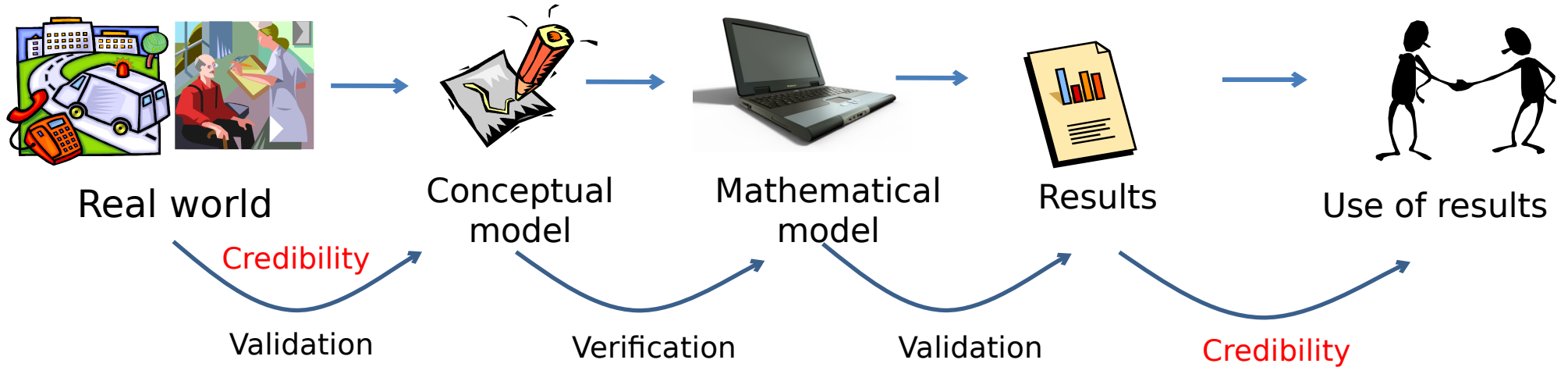
# The model – Completed

Congratulations!

You have built a model of a healthcare pathway  
and tested the impact of two changes

The story continues...

# Model validation and verification



- **Validation:** the process of determining whether a model is sufficiently accurate for the particular objectives of the study
- **Verification:** the process of determining if the conceptual model has been correctly translated into a mathematical or computer model (testing).
- **Credibility:** the model and its results are believed and trusted by users

# Conceptual model validation

*The CM contains details about the study objectives, how the real world works, assumptions, model simplifications, data sources and what-if scenarios.*

*E.g: the clinic model*

## **Model simplifications**

- No breaks for the doctor
- Same doctor each day working at same rate
- No check-in time with receptionist, no queue for receptionist
- The doctor continues working until the last patient has been seen
- The doors close at 5pm, patients aren't stopped from entering before

## **Assumptions**

- Clinic starts on time!
- Travel times are trivial
- No priority of patients



# Simple 'black box' validation

A simple approach to validating models is to compare the model's predicted outputs with data from the real system

*E.g. a comparison of modelled and historical average queuing times*

Issues?

- Ignores the internal workings of the model
- How do you validate a what-if scenario (no real world to compare against)?
- Often real world data are unavailable or inaccurate

Black box validation is a good start, but not enough! We need to:

- Validate smaller parts of the model
- Validate input data
- Work closely with subject matter experts
- Compare results to similar (sometimes simpler) models
- Assess uncertainty (to be discussed later)

# Confidence not validity

- It is not possible to prove that a model is valid
- But you can disprove that it is valid
- View validation as a process of building confidence that the model and its results are sufficiently accurate for the purpose it was built to address
- Confidence thresholds vary by user and modeller.
- This means that a valid model may not be viewed as credible by users
- This means that an invalid model may be viewed as credible by users!
- Do not fall into the trap of thinking that a valid model is the most 'realistic one'

# Final points

- A more complex model is not always a better model. In fact it rarely is!
- All models are wrong but some are useful
- An objectively built and well validated model can support decision making
- The real world is far more complex than a model so it should inform and not determine a decision
- Pathway models should be designed and built to answer a particular question
- Whole system models are rarely useful – better to focus on a known problem area
- Modelling is an iterative process – you will go back and forth between conceptual modelling, the formal model and output analysis
- A good model requires good data
- Enjoy building models, the process should be fun and interesting

# Well done!

## We have made it to the end of the session

Check out our websites for more information and advanced training materials

<http://clahrc-peninsula.nihr.ac.uk/>

<http://penchord.org>

<https://pythonhealthcare.org/>

More to come soon!!