











# Training for Medical education via innovative eTechnology (MediTec)





# Training for Medical education via innovative eTechnology (MediTec)

ERASMUS+ PROGRAMME

Erasmus+ - Key Action 2

Capacity Building in the Field of Higher Education

Project No: 585980-EPP-1-2017-1-DE-CBHE-JP

# Training for Medical education via innovative eTechnology (MediTec)

**Workshop in Malta**

**26-28.06.2018**

**Venue: The ICT Lab (Level -1, Block B, Room 1).  
Faculty of Information and Communication  
Technology,  
ICT Building, University of Malta, Msida, Malta**



# Training for Medical education via innovative eTechnology (MediTec)

June 26th, 2019

8:30 – 9:00	Breakfast and Registration	
09:00 – 09:30	Welcome and Introduction	Dr Lalit Garg
9:30 – 10:30	Health Informatics	Dr Lalit Garg
10:30 – 11:00	Refreshment	
11:30 – 12:30	Health Systems Management, and Leadership	Prof. Sandra Buttigieg
12:30 – 13:30	Lunch	
13:30 – 14:30	Dental curriculum Review and Update	
14:30 – 15:00	Refreshment	
15:00 – 16:00	Dental curriculum Review and Update	

# Training for Medical education via innovative eTechnology (MediTec)

June 27th, 2019

8:30 – 9:00	Breakfast and Registration	
09:00 – 10:30	Advances in Medical Education	Prof. Isabel Stabile
10:30 – 11:00	Refreshment	
11:30 – 12:30	Advances in Medical Education	Prof. Isabel Stabile
12:30 – 13:30	Lunch	
13:30 – 14:30	Visit to the Mater Dei Hospital	
14:30 – 15:00	Refreshment	
15:00 – 16:00	Visit to the Mater Dei Hospital	



# Training for Medical education via innovative eTechnology (MediTec)

June 28th, 2019

8:30 – 9:00	Breakfast and Registration	
09:00 - 10:30	Management Committee Meeting	
10:30 – 11:00	Refreshment	
11:30 – 12:30	Management Committee Meeting	
12:30 – 13:30	Lunch	
13:30 – 14:30	Management Committee Meeting	
14:30 – 15:00	Refreshment	
15:00 – 16:00	Visit to the Simulation Lab	
June 29th, 2019		
09:00 – 12:30	Visit to Valletta and Mdina	

# Training for Medical education via innovative eTechnology (MediTec)

**PS: We will provide the refreshment and the Lunch (Halal or vegetarian).**

**If you have any questions, you can always call**

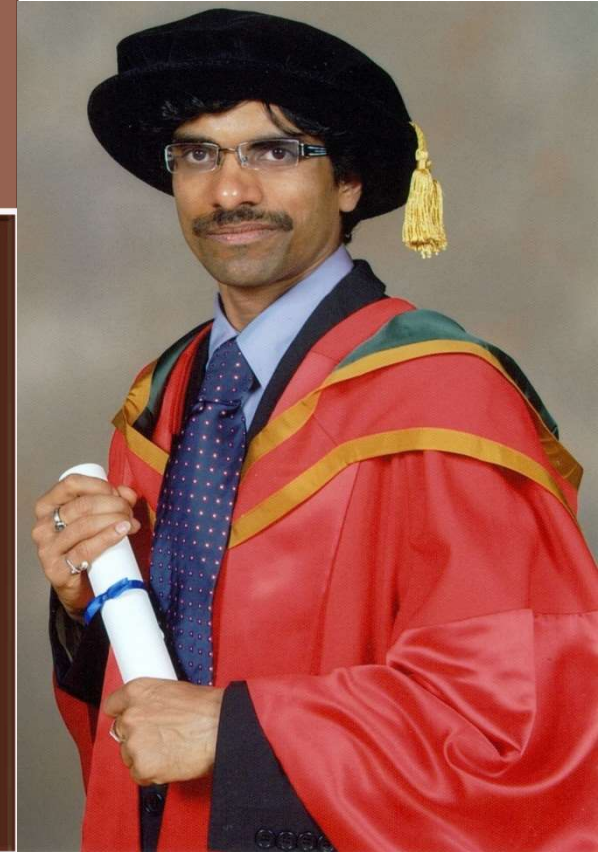
**Dr Lalit Garg at +356-79233327,**

**Mr Emeka Chukwu: +356-99330888**

**Mr Vijay Sony: +356-99705051.**

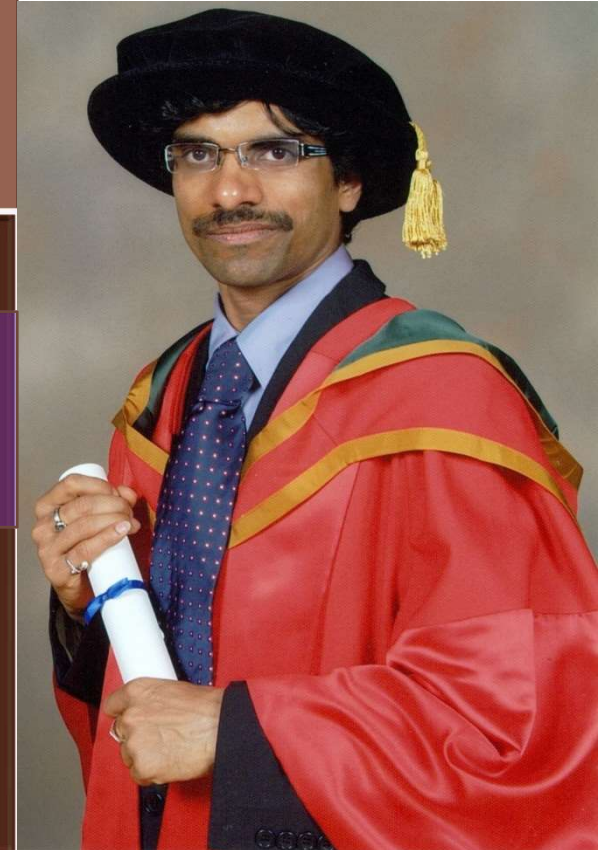


# PhD Computer Science University of Ulster, UK, 2010



PhD Computer Science  
University of Ulster, UK, 2010

Postgraduate: Information Technology,  
ABV-IIITM, Gwalior, India, 2001

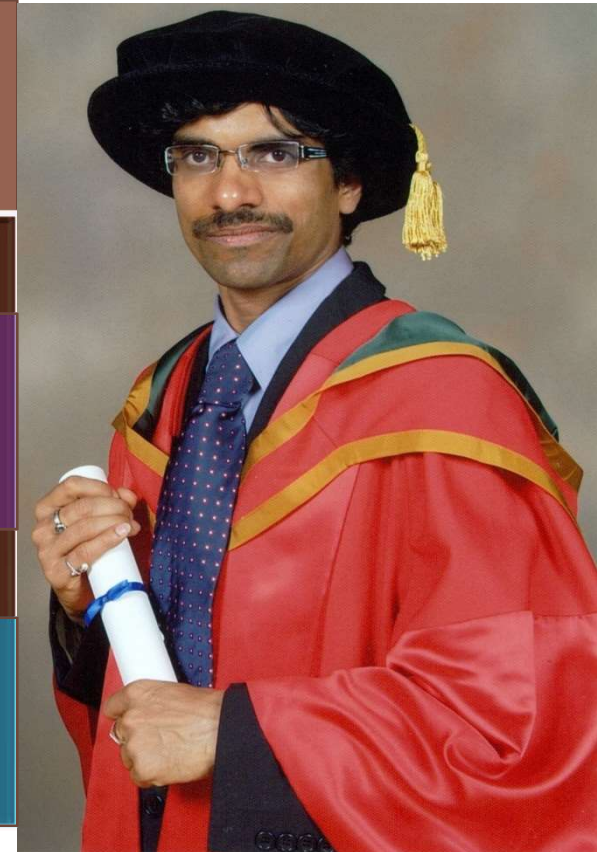




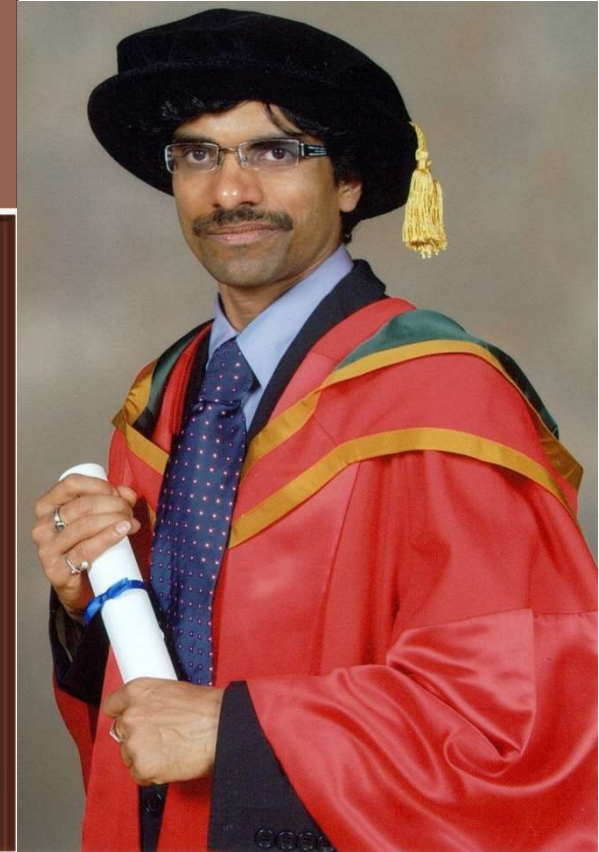
PhD Computer Science  
University of Ulster, UK, 2010

Postgraduate: Information Technology,  
ABV-IIITM, Gwalior, India, 2001

Bachelor: Electronics & Communication  
Engineering, Barkatullah University, India, 1999

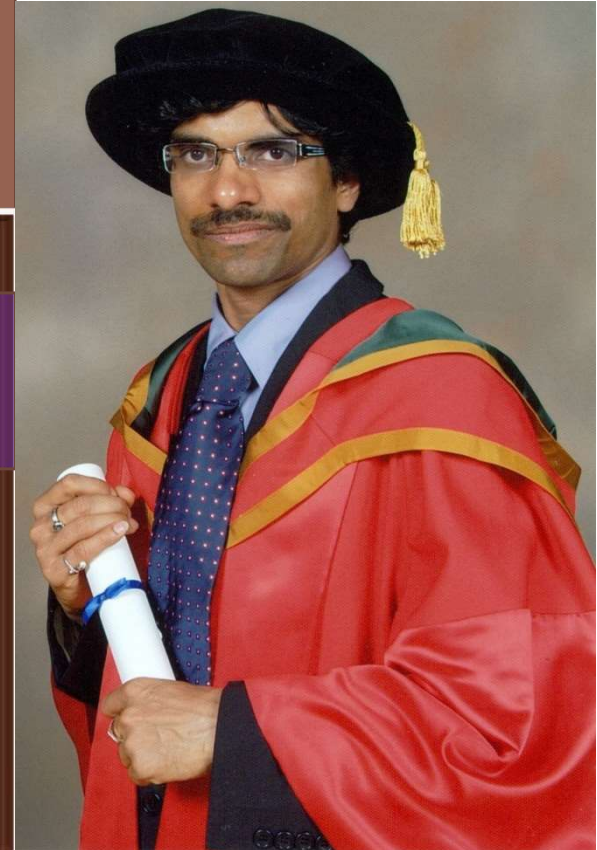


# Supervision experience



## Supervision experience

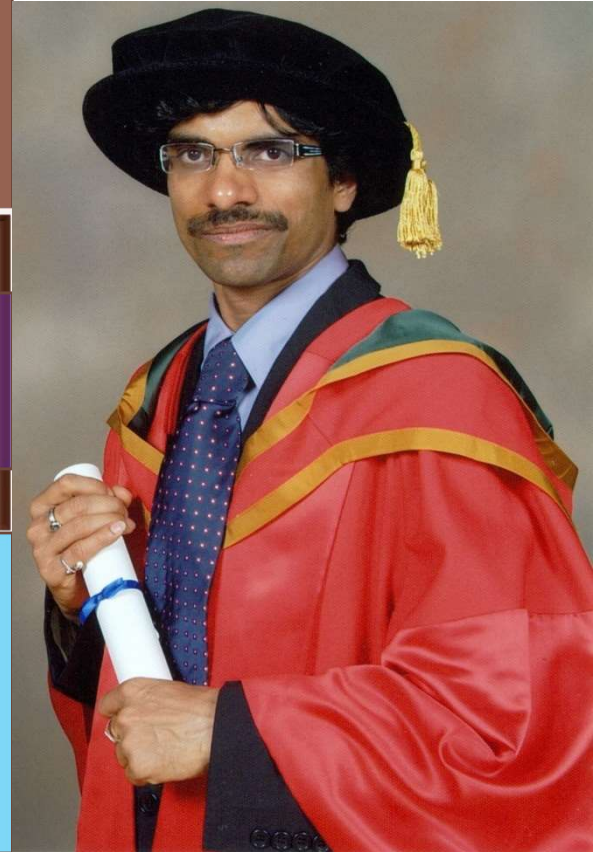
More than 200 successful masters dissertations (University of Liverpool, UK)



## Supervision experience

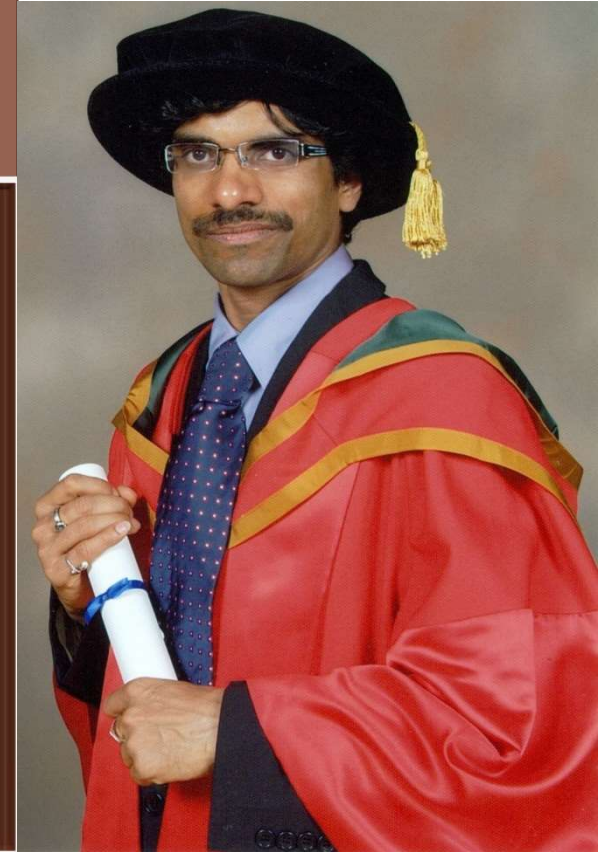
More than 200 successful masters dissertations (University of Liverpool, UK)

Many of these were sponsored by industrial organizations.





More than 18 years of teaching and  
research experience



More than 18 years of teaching and  
research experience

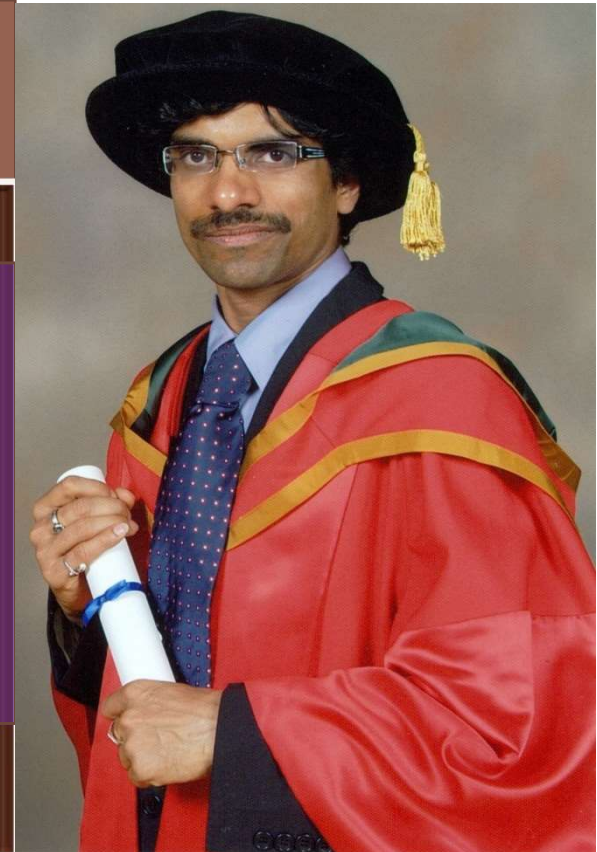
Thapar University, India

University of Liverpool, UK

University of Malta, Malta

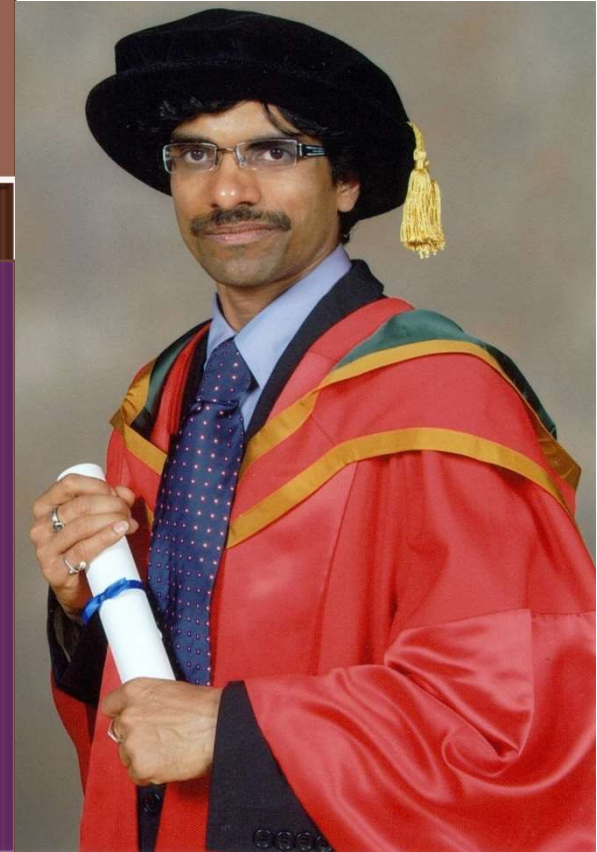
Nanyang Technological University, Singapore

University of Ulster, UK



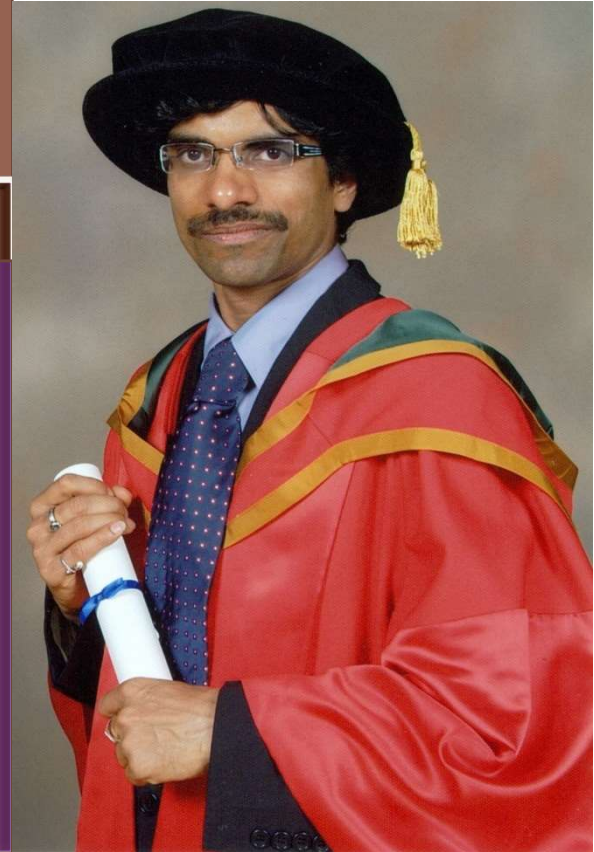
## Publication record

- 32 papers in refereed journals,
- 52 papers in refereed conferences,
- 10 refereed book chapters,
- 18 other (extended) abstracts,
- 1 contributed book and 1 edited book.



## Publication record

- Google Scholar: H-Index: 12 (total 557 citations and 77 indexed publications)
- Scopus: H-Index: 9 (total 297 citations and 50 indexed publications),
- The Web of knowledge: H-Index: 8 (total 187 citations and 38 indexed publications)





Visit me at  
<http://lalitgarg.info/>

Publication Record, Research  
Interests, Projects and .....



Visit me at  
<http://lalitgarg.info/>

Research Interests, Projects and  
.....

Other information



# Health data Analytics: Making Sense of Health Data to improve health services

**Lalit Garg,**

**Senior Lecturer, University of Malta, Malta**

**Honorary Lecturer, University of Liverpool, UK**

e-mail: [lalit.garg@um.edu.mt](mailto:lalit.garg@um.edu.mt)

web: <http://lalitgarg.info/>

Phone: +356-2340-2112



# Roadmap

- **Introduction**
  - **Complex Systems**
  - **Some interesting problems and observations**
- **Background**
  - **Phase type distribution**
  - **Phase type distribution survival trees**
- **Applications**
- **Publications**      16/10/2019



# Introduction

- Life expectancy has increased with improvement in health services and standard of living.

# Introduction

- Life expectancy has increased with improvement in health services and standard of living.
- Higher demand to the healthcare resources

# Introduction

- Life expectancy has increased with improvement in health services and standard of living.
- Higher demand to the healthcare resources
- Healthcare challenge is to continue providing the same quality of care

# Introduction

- Healthcare system facing major problems



# Introduction

- Healthcare system facing major problems
  - Lack of beds in hospitals

# Introduction

- Healthcare system facing major problems
  - Lack of beds in hospitals and
  - Lack of other hospital resources.

# Introduction

- To work with these problems the healthcare system needs :

# Introduction

- To work with these problems the healthcare system needs :
  - An efficient way to forecast the resources required

# Introduction

- To work with these problems the healthcare system needs :
  - An efficient way to forecast the resources required
  - To minimize the cost of care while maintaining the quality of care.



# Introduction

- When modelling the healthcare system it would help:

# Introduction

- When modelling the healthcare system it would help:
  - To better understand the process for the design of policies that can improve the quality of care

# Introduction

- When modelling the healthcare system it would help:
  - To better understand the process for the design of policies that can improve the quality of care
  - To ensure the optimal utilization of the available resources

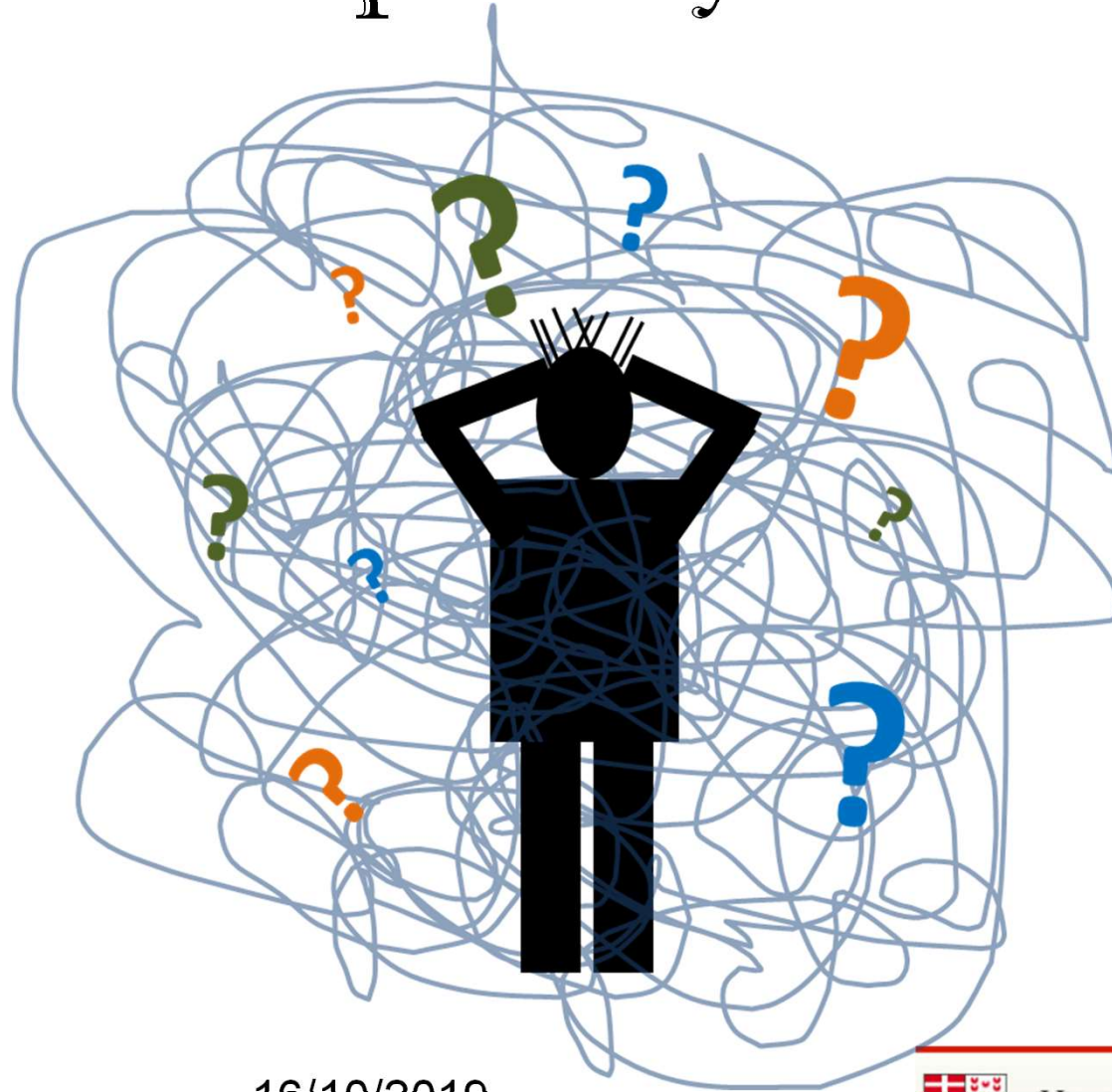
# Background



MediTec

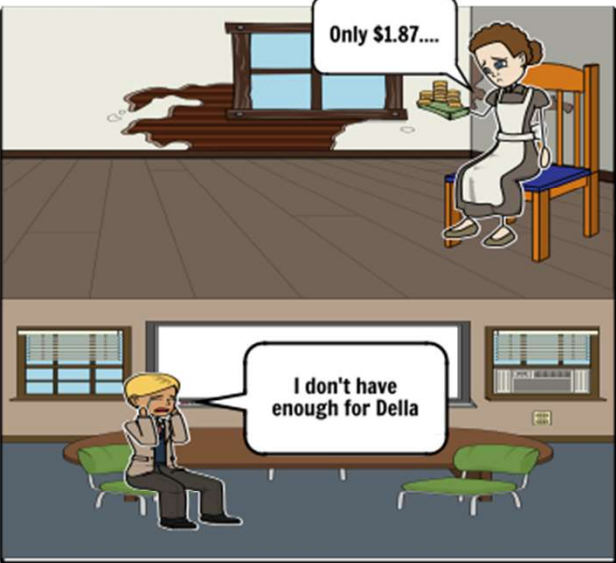
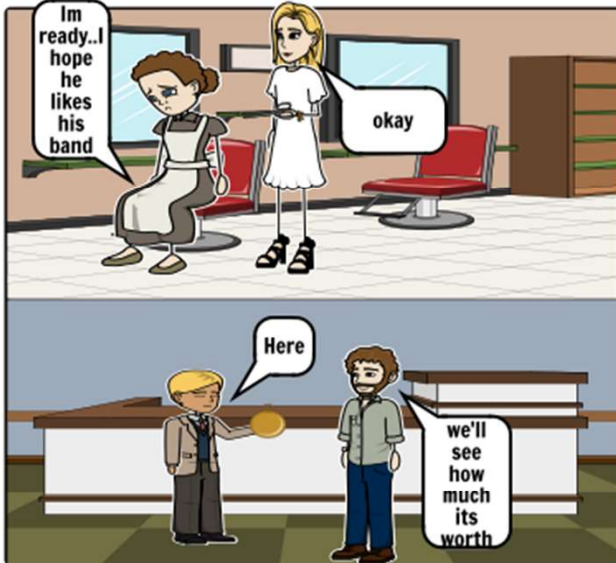
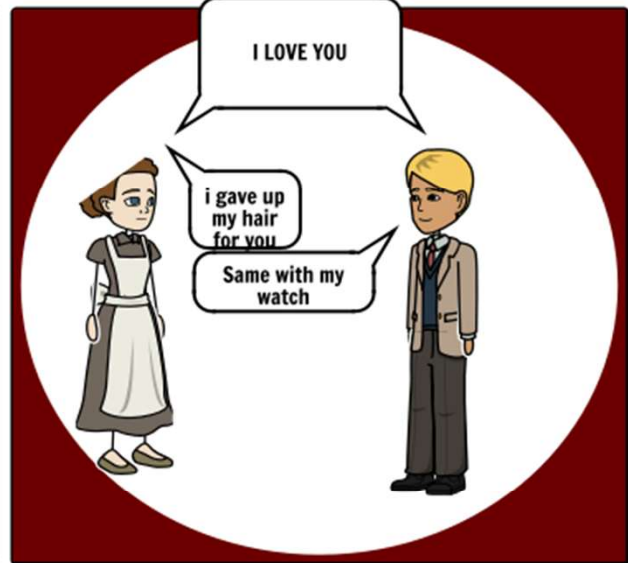
16/10/2019

# Complex Systems





# Our family system: One of the most complex systems

Rising Action	Climax	Falling Action
 <p>Only \$1.87....</p> <p>I don't have enough for Della</p>	 <p>I'm ready..I hope he likes his band</p> <p>okay</p> <p>Here</p> <p>we'll see how much its worth</p>	 <p>I LOVE YOU</p> <p>i gave up my hair for you</p> <p>Same with my watch</p>
<p>Della nor James had enough to exchange gifts. They were very poor. Both were very disappointed. Then, they had an idea.</p>	<p>Della gave up her most prized possession....her hair. James gave up his watch:also a prized possession. They did it for each other.</p>	<p>They bought gifts for each other with the possessions they gave up. Della received combs. James received a band for his watch. They really love each other.</p>

Create your own at Storyboard That

[https://www.storyboardthat.com/storyboards/baptist\\_snniper/the-gift-of-the-magi-story-elements](https://www.storyboardthat.com/storyboards/baptist_snniper/the-gift-of-the-magi-story-elements)

# Our family system: One of the most complex systems



# Our family system: One of the most complex systems



# Our family system: One of the most complex systems

Requires

1. Human Behavioural Modelling

# Our family system: One of the most complex systems

## Requires

1. Human Behavioural Modelling
2. Modelling the effect of others' Behaviour (using game theory),



# Our family system: One of the most complex systems

## Requires

1. Human Behavioural Modelling
2. Modelling the effect of others' Behaviour (using game theory),
3. Modelling of cultural, social, economical, financial and environmental effects (Big data analytics),

# Our family system: One of the most complex systems

## Requires

1. Human Behavioural Modelling
2. Modelling the effect of others' Behaviour (using game theory),
3. Modelling of cultural, social, economical, financial and environmental effects (Big data analytics),
4. Most difficult: modelling spontaneous (uncorrelated) changes in sentiments,

# Our family system: One of the most complex systems

## Requires

1. Human Behavioural Modelling
2. Modelling the effect of others' Behaviour (using game theory),
3. Modelling of cultural, social, economical, financial and environmental effects (Big data analytics),
4. Most difficult: modelling spontaneous (uncorrelated) changes in sentiments,
5. Reality vs perception.

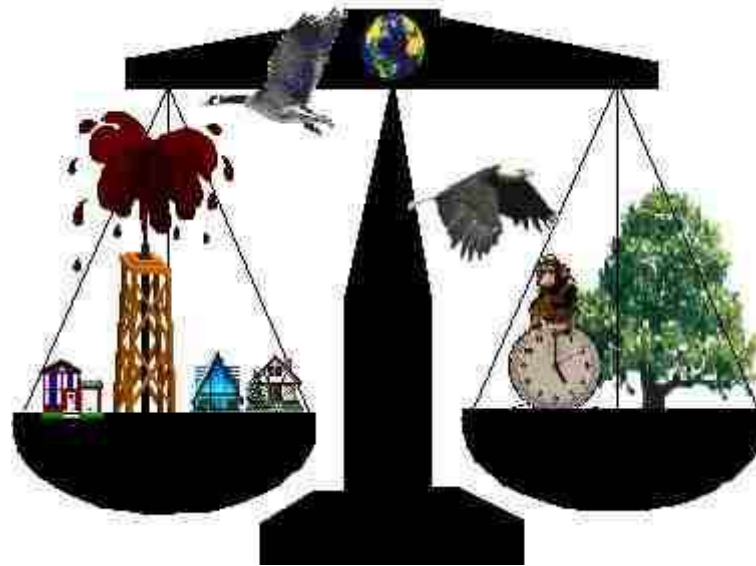
MediTec

# Reality vs Perception



MediTec

# Our Economy: A Complex System



MediTec



# Our Economy: A Complex System

With government intervention:

MediTec

# Our Economy: A Complex System

With government intervention:

More demand than supply = More subsidy to the  
buyer

# Our Economy: A Complex System

With government intervention:

More demand than supply = More buyer subsidy

More buyer subsidy = More profit

# Our Economy: A Complex System

- **All buyer subsidy will go to supplier/  
manufacturers**



# Our Economy: A Complex System

With government intervention:

More demand than supply = More buyer subsidy

More buyer subsidy = More profit

More profit = More attractive industry

# Our Economy: A Complex System

**All grants will ultimately go to the buyers**



MediTec



UNIVERSITY OF MALTA  
*L-Università ta' Malta*



# Our Economy: A Complex System

With government intervention:

More demand than supply = More buyer subsidy

More buyer subsidy = More profit

More profit = More attractive industry

= More players

# Our Economy: A Complex System

With government intervention:

More demand than supply = More buyer subsidy

More buyer subsidy = More profit

More profit = More attractive industry

= More players

= More supply than demand

# Our Economy: A Complex System

With government intervention:

More demand than supply = More buyer subsidy

More buyer subsidy = More profit

More profit = More attractive industry

= More players

= More supply than demand

= Less price = Less profit

# Our Economy: A Complex System

With government intervention:

More demand than supply = More buyer subsidy

More buyer subsidy = More profit

More profit = More attractive industry

= More players

= More supply than demand

= Less price = Less profit

= Some will leave the market with loss

# Our Economy: A Complex System

With government intervention:

More demand than supply = More buyer subsidy

More buyer subsidy = More profit

More profit = More attractive industry

= More players

= More supply than demand

= Less price = Less profit

= Some will leave the market with loss

= More demand than supply

# Our Economy: A Complex System

With government intervention:

More demand than supply = More grant to the  
supplier/manufacturer to ensure meeting the  
demand



# Our Economy: A Complex System

With government intervention:

More demand than supply = More supplier grant

More supplier subsidy = More profit

# Our Economy: A Complex System

With government intervention:

More demand than supply = More supplier grant

More supplier subsidy = More profit

More profit = More attractive industry

# Our Economy: A Complex System

With government intervention:

More demand than supply = More supplier grant

More supplier subsidy = More profit

More profit = More attractive industry  
= More players

# Our Economy: A Complex System

With government intervention:

More demand than supply = More supplier grant

More supplier subsidy = More profit

More profit = More attractive industry

= More players

= More supply than demand

# Our Economy: A Complex System

With government intervention:

More demand than supply = More supplier grant

More supplier subsidy = More profit

More profit = More attractive industry

= More players

= More supply than demand

= Less price = Less profit

# Our Economy: A Complex System

With government intervention:

More demand than supply = More supplier grant

More supplier subsidy = More profit

More profit = More attractive industry

= More players

= More supply than demand

= Less price = Less profit

= Some will leave the market with

losses

MediTec



# Our Economy: A Complex System

With government intervention:

More demand than supply = More supplier grant

More supplier subsidy = More profit

More profit = More attractive industry

= More players

= More supply than demand

= Less price = Less profit

= Some will leave the market with loss

= More demand than supply

# Our Economy: A Complex System

Without government intervention:

More demand than supply = More profit

# Our Economy: A Complex System

Without government intervention:

More demand than supply = More profit

More profit = More attractive industry

= More players

# Our Economy: A Complex System

Without government intervention:

More demand than supply = More profit

More profit = More attractive industry

= More players

= More supply than demand

# Our Economy: A Complex System

Without government intervention:

More demand than supply = More profit

More profit = More attractive industry

= More players

= More supply than demand

= Less price = Less profit

# Our Economy: A Complex System

Without government intervention:

More demand than supply = More profit

More profit = More attractive industry

= More players

= More supply than demand

= Less price = Less profit

= Some will leave the market with loss



# Our Economy: A Complex System

Without government intervention:

More demand than supply = More profit

More profit = More attractive industry

= More players

= More supply than demand

= Less price = Less profit

= Some will leave the market with loss

= More demand than supply

# Our Economy: A Complex System

For industries to solve the problem: One of the solution is **innovation**.

- ✓ Develop substitute products
- ✓ Make process more efficient
- ✓ Reduce production cost
- ✓ More efficient supply chain network
- ✓ Provide add-on services

# Our Economy: A Complex System

For other industries to solve the problem: One of the methods is **innovation**

And **duration of stay in the market**

- ✓ How long an industry would be attractive
- ✓ When to leave the market/industry
- ✓ How to increase this duration
- ✓ Alternate product development through **innovation**
- ✓ Plan to leave the market/industry

# Our Education System

According to MHRD:

- In 2015, there were more than 6000 engineering and technology institutes.
- Produced more than 2.9 million engineering graduates.
- Only 1.5 million got jobs in their engineering discipline.
- ?????

# Our Education System

- **The decision to pursue BE/BTech in their chosen discipline was taken 4 years back based on then current data.**



MediTec

# Food wastage

- **Should we reduce the food wastage or not?**



# Food wastage

- **Should we reduce the food wastage or not?**
- Assume there is 35% food wastage
- Means we are producing 135% food than required.
- Are food producers (farmers) getting too much profit?
- Are food product prices are inflated?



# Food wastage

- **Should we reduce the food wastage or not?**
- What if we reduce the food wastage by 50%?
- Then the demand will be 118% and supply will be 135%?
- What will be the food prices?
- What will happen with our farmers?



# Our healthcare system: A complex system



# Our healthcare system: A complex system



# Our healthcare system: A complex system

# Other complex systems: A complex system

## Private healthcare:

- Some patients want cheap healthcare

# Other complex systems: A complex system

## Private healthcare:

- Some patients want cheap healthcare
- Some patients want best (luxurious) healthcare

# Other complex systems: A complex system

## Private healthcare:

- Some patients want cheap healthcare
- Some patients want best (luxurious) healthcare
- Health providers want maximum profit



# Other complex systems: A complex system

## Private healthcare:

- Some patients want cheap healthcare
- Some patients want best (luxurious) healthcare
- Health providers want maximum profit
- maximum profit = maximum hospital visits

# Other complex systems: A complex system

## Private healthcare:

- Some patients want cheap healthcare
- Some patients want best (luxurious) healthcare
- Health providers want maximum profit
- maximum profit = maximum hospital visits
- = maximum readmissions
- + maximum hospital duration of stay



# Our healthcare system: A complex system

# Other complex systems: A complex system

## Public healthcare:

- Everyone gets the same healthcare

# Other complex systems: A complex system

## Public healthcare:

- Everyone gets the same healthcare
- Health providers want minimum cost

# Other complex systems: A complex system

## Public healthcare:

- Everyone gets the same healthcare
- Health providers want minimum cost
- Minimum cost
  - = Limited resources
  - = least duration in hospitals + minimum admissions

# Other complex systems: A complex system

## Public healthcare:

- Everyone gets the same healthcare
- Health providers want minimum cost
- Minimum cost
  - = Limited resources
  - = least duration in hospitals + waiting list



# Other complex systems: A complex system

## Public healthcare:

- Everyone gets the same healthcare
- Health providers want minimum cost
- Minimum cost
  - = Limited resources
  - = more readmissions + waiting list

# Other complex systems: A complex system

## Public healthcare:

- Everyone gets the same healthcare
- Health providers want minimum cost
- Minimum cost
  - = Limited resources
  - = more readmissions + waiting list
  - = longer waiting list

# Other complex systems: A complex system



<https://fineartamerica.com/featured/hospital-waiting-room-mark-thomasscience-photo-library.html>

MediTec



UNIVERSITY OF MALTA  
*L-Università ta' Malta*

# Other complex systems: A complex system

## Public healthcare:

- Everyone gets the same healthcare
- Health providers want minimum cost
- Minimum cost
  - = Limited resources
  - = longer waiting list
  - = **Poor healthcare**

# Other complex systems: A complex system

## Public healthcare:

- Everyone gets the same healthcare
- Health providers want minimum cost
- Minimum cost
  - = Limited resources
  - = longer waiting list
  - = **Poor healthcare**
  - = **Public outcry**
  - = **Preference**

# Other complex systems: A complex system

## Public healthcare:

- ~~Everyone gets the same healthcare~~
- Health providers want minimum cost
- Minimum cost
  - = Limited resources
  - = longer waiting list
  - = **Poor healthcare**
  - = **Public outcry**
  - = **Preference**

# Other complex systems: A complex system

## Public healthcare:

- **Corruption**
- Health providers want minimum cost
- Minimum cost
  - = Limited resources
  - = longer waiting list
  - = **Poor healthcare**
  - = **Public outcry**
  - = **Preference**

# Other complex systems: A complex system

Public healthcare:

- More resources



# Other complex systems: A complex system

Public healthcare:

- More resources = More cost

# Other complex systems: A complex system

## Public healthcare:

- More resources = short waiting lists

# Other complex systems: A complex system

## Public healthcare:

- More resources = short waiting lists
- Short waiting list = longer hospital stay

# Other complex systems: A complex system

## Public healthcare:

- More resources = short waiting lists
- Short waiting list = longer hospital stay  
minimum readmissions

# Other complex systems: A complex system

## Public healthcare:

- More resources = short waiting lists
- Short waiting list = longer hospital stay  
minimum readmissions  
**more patients**

# Other complex systems: A complex system

## Public healthcare:

- Even more resources = no waiting lists
- Short waiting list = longer hospital stay  
minimum readmissions  
**more patients**  
**underutilization**

# Other complex systems: A complex system

## Public healthcare:

- Even more resources = no waiting lists
- Short waiting list = longer hospital stay  
minimum readmissions  
**more patients**  
**underutilization**  
**misuse**

# Other complex systems: A complex system

## Public healthcare:

- Even more resources = no waiting lists
- Short waiting list = longer hospital stay  
minimum readmissions  
**more patients**  
**underutilization**  
**misuse**  
**more cost**



# Other complex systems: A complex system

## Public healthcare:

- Even more resources = no waiting lists
- Short waiting list = longer hospital stay  
minimum readmissions  
**more patients**  
**underutilization**  
**misuse**  
**more cost**  
**Some waiting lists**

# Other complex systems: A complex system

Public healthcare:

- Optimum resources = optimum waiting time

# Other complex systems: A complex system

## Public healthcare:

- Optimum resources = optimum waiting time
- = Optimum hospital stay

# Other complex systems: A complex system

## Public healthcare:

- Optimum resources = optimum waiting time  
= Optimum hospital stay  
= minimum readmissions

# Other complex systems: A complex system

## Public healthcare:

- Optimum resources = optimum waiting time  
= Optimum hospital stay  
= minimum readmissions  
**= optimum patients' number**

# Other complex systems: A complex system

## Public healthcare:

- Optimum resources = optimum waiting time  
= Optimum hospital stay  
= minimum readmissions  
**= optimum patients' number**  
**optimum utilization**

# Other complex systems: A complex system

## Public healthcare:

- Optimum resources = optimum waiting time  
= Optimum hospital stay  
= minimum readmissions  
**= optimum patients' number**  
**optimum utilization**  
**minimum misuse**

# Other complex systems: A complex system

## Public healthcare:

- Optimum resources = optimum waiting time  
= Optimum hospital stay  
= minimum readmissions  
**= optimum patients' number**  
**optimum utilization**  
**minimum misuse**  
**optimum cost**



# Other complex systems: A complex system

## Public healthcare:

- Optimum resources = optimum waiting time  
= Optimum hospital stay  
= minimum readmissions  
**= optimum patients' number**  
**optimum utilization**  
**minimum misuse**  
**optimum cost**  
**Some waiting lists**

# Other complex systems: A complex system

## Public healthcare:

- Optimum resources = Proper planning
  - = Continuously adding resources (if population is increasing/changing)
  - = **Resource requirement forecasting**

# Other complex systems: A complex system

## Public healthcare:

- Optimum resources = Proper planning
  - = Continuously adding resources (if population is increasing/changing)
  - = **Resource requirement forecasting**
  - = **Admission rate estimation**
  - = **Length of stay estimation**



# Coxian phase type distributions



# Coxian phase type distributions

Among popular choices to fit spell length of  
stay data.



# Coxian phase type distributions

Among popular choices to fit spell length of stay data.

Provide a simple interpretation of fit for the length of stay data.



# Coxian phase type distributions

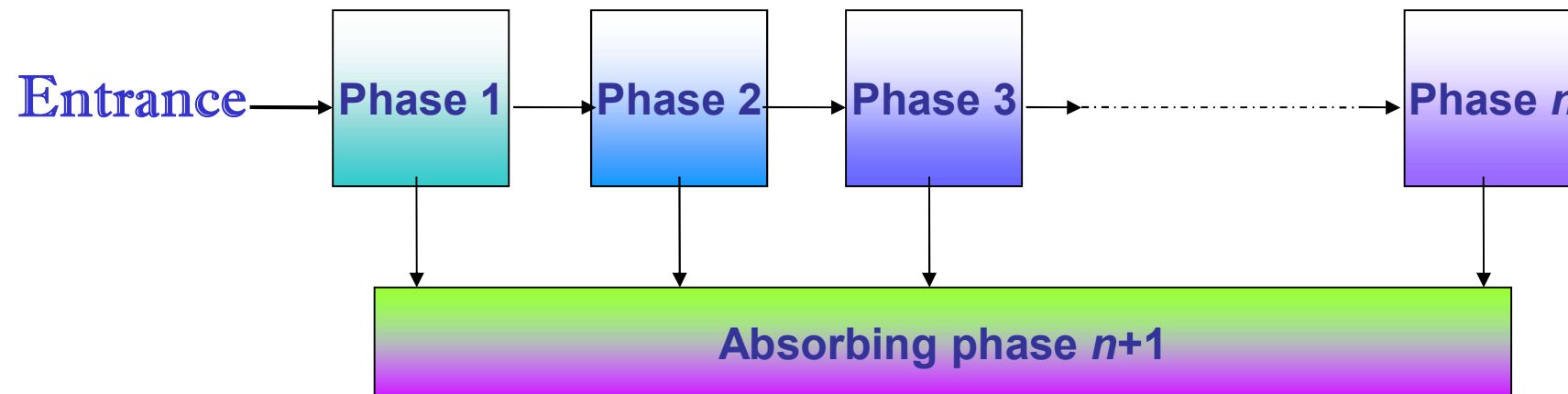
Among popular choices to fit spell length of stay data.

Provide a simple interpretation of fit for the length of stay data.

Parameter estimation is easier than other phase type distributions.



# A Markov chain





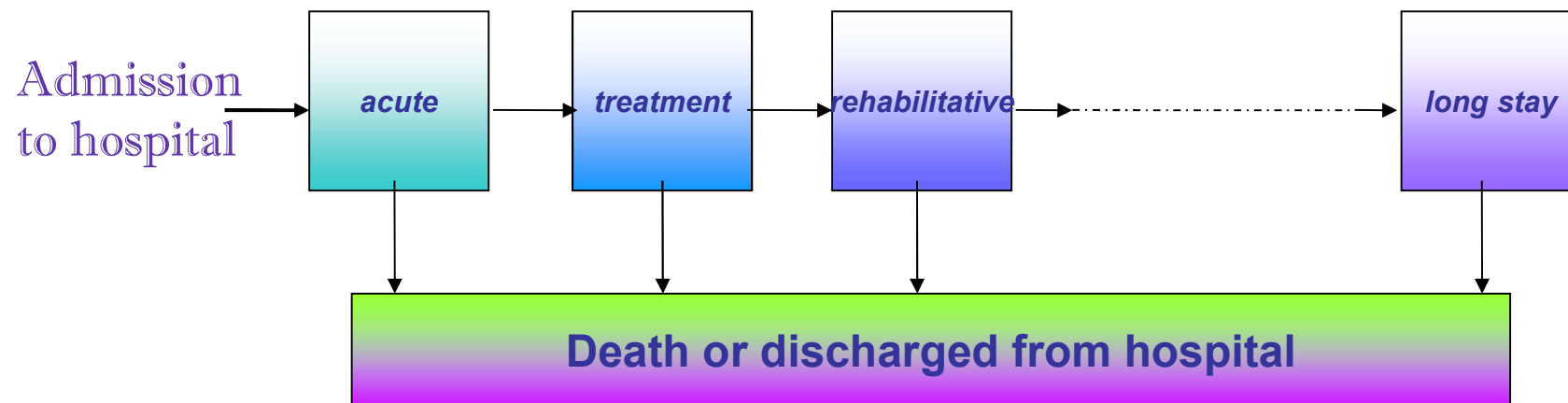
# A Markov chain





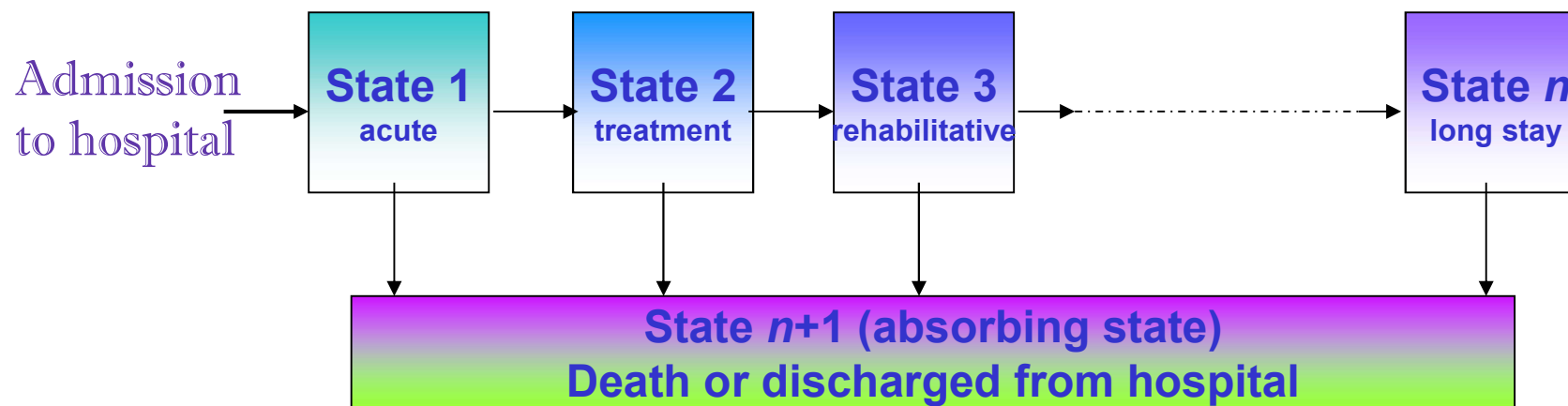
# Hospital care system as a Markov chain

Patient flow in the stroke care system can be modelled as an  $n$  state Markov process with Coxian phase type distributions



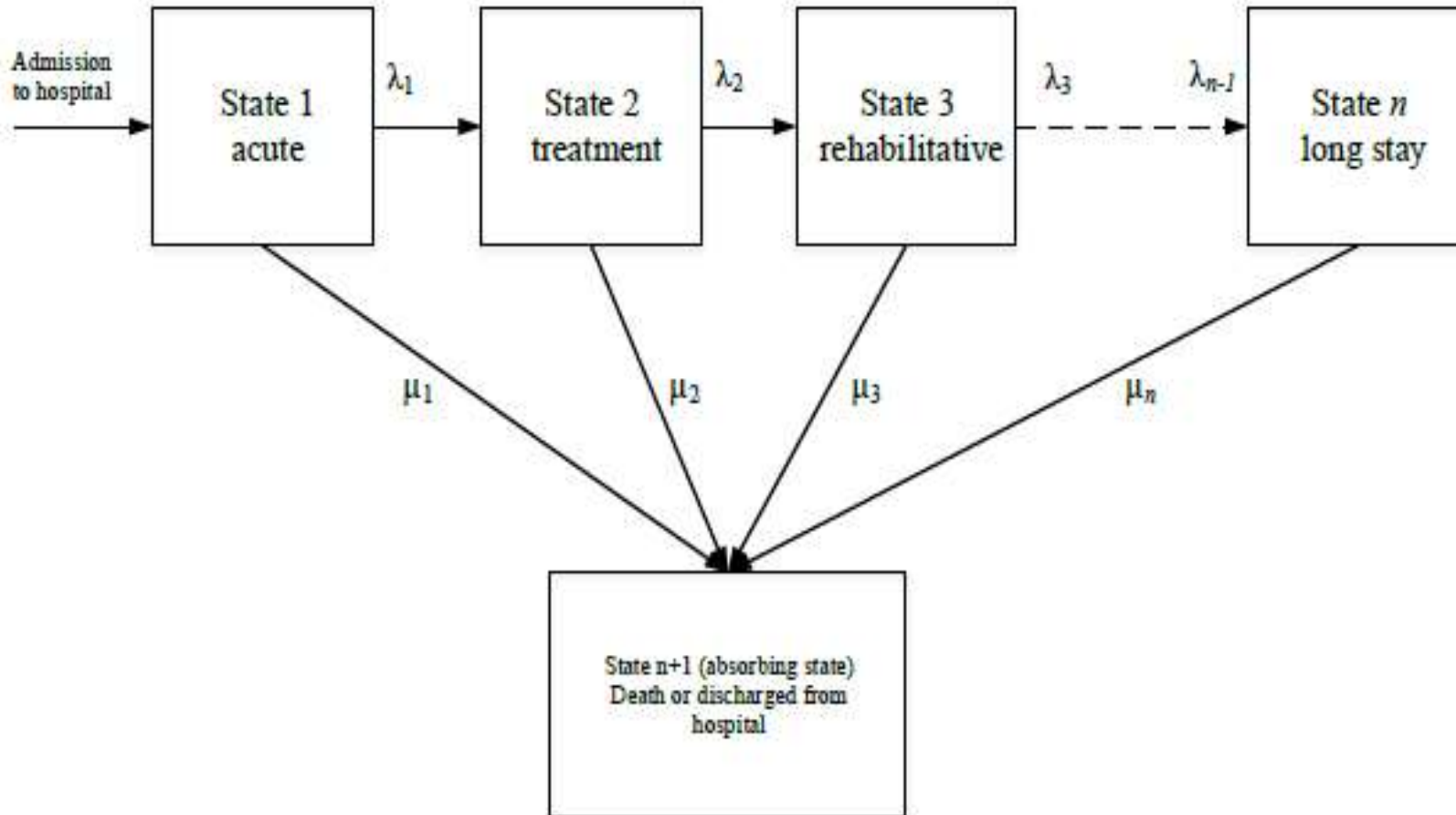


# Coxian phase type distributions





# Coxian phase type distributions





# Coxian phase type distributions

A process can start only in the first state (state 1).

Sequential transition rate is  $\lambda_k$ .

Also transition rate from any state  $k$  to the absorbing state  $n+1$  is  $\mu_k$ .

# Coxian phase type distributions

The PDF for the duration before absorption:

$$f(t) = \mathbf{p} \exp(\mathbf{Q}t) \mathbf{q}$$

where the initial state probability distribution

$$\mathbf{p} = (1 \ 0 \ 0 \ \dots \ 0 \ 0)$$

absorption probabilities

$$\mathbf{q} = (\mu_1 \ \mu_2 \ \dots \ \mu_{n-2} \ \mu_n)^T .$$

# Coxian phase type distributions

And the transition matrix

$$\mathbf{Q} = \begin{pmatrix} -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & -(\lambda_2 + \mu_2) & \lambda_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & -(\lambda_{n-1} + \mu_{n-1}) & \lambda_{n-1} \\ 0 & 0 & 0 & \dots & 0 & -\mu_n \end{pmatrix}$$

# Coxian phase type distributions

The likelihood function:

$$l = \prod_{i=1}^N (\mathbf{p} \exp\{\mathbf{Q}t_i\} \mathbf{q})$$

where  $N$  is the total number of patients in the care system.



# Coxian phase type distributions

The loglikelihood function

$$L = \sum_{i=1}^N \left( \log \left( \mathbf{p} \exp \{ \mathbf{Q} t_i \} \mathbf{q} \right) \right) .$$

Or

$$L = \sum_{i=1}^N f(t_i)$$

where  $f(t_i) = \log \left( \mathbf{p} \exp \{ \mathbf{Q} t_i \} \mathbf{q} \right)$

# Weighted-Average

## Information Criterion

WIC (Weighted-Average Information Criterion) is a weighted average of the Bayesian information criterion and the Akaike information criterion with a small sample size correction.

# Weighted-Average Information Criterion

WIC (Weighted-Average Information Criterion) is a weighted average of the Bayesian information criterion and the Akaike information criterion with a small sample size correction.

The splitting criteria based on the WIC combines the strengths of both the AIC and the BIC it works well with small and large sample sizes and in the situation where sample size is not known.

# Weighted-Average Information Criterion

The performance of WIC was compared with several other popular criteria in the study and the results showed that WIC is very reliable.

$$WIC = -2L + d + \frac{d(((\log(N) - 1) \log(N))(N - (d - 1))^2 + 2N(N + (d + 1)))}{(2N + (\log(N)(N - (d + 1))))(N - (d + 1))} .$$

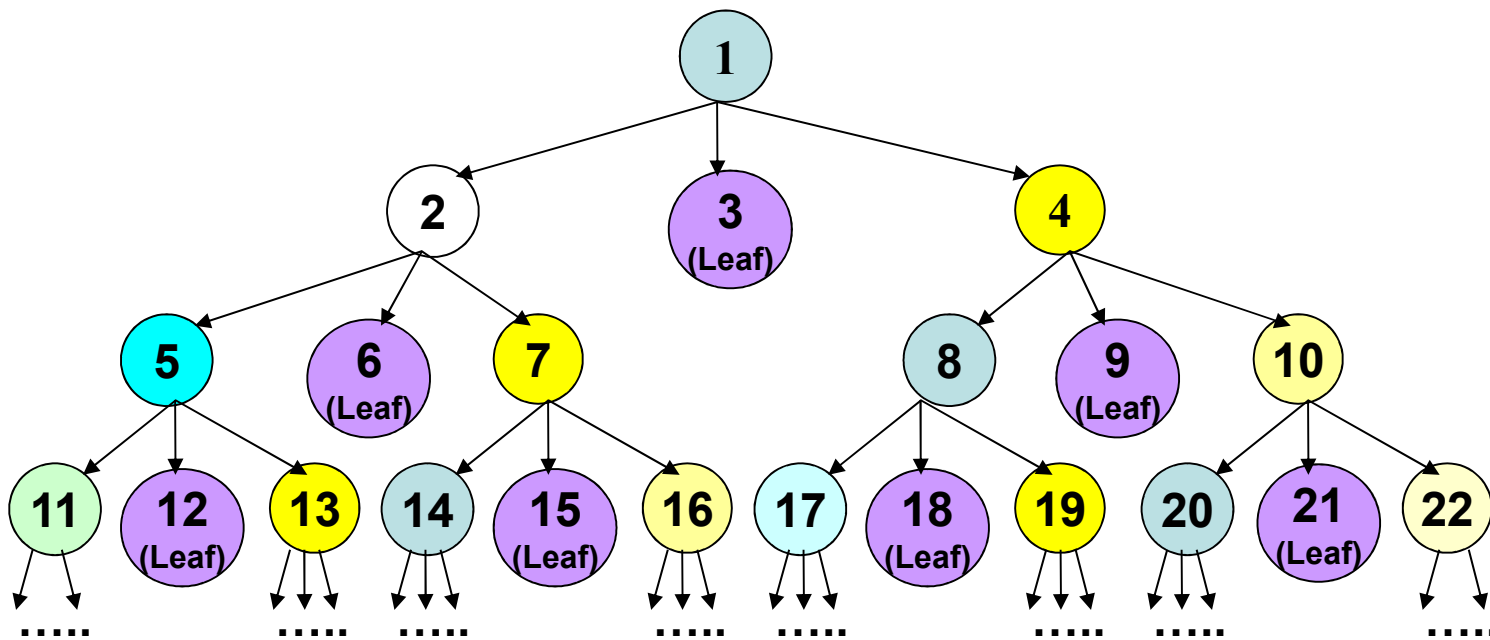


# Survival tree





# Survival tree



# Survival trees

- Decision trees in survival analysis

# Survival trees

- Decision trees in survival analysis
- A type of classification and regression trees



# Survival trees

- Decision trees in survival analysis
- A type of classification and regression trees
- Constructed by recursively partitioning the given dataset in to subsets based on some splitting and selection criteria.

# Phase type survival tree



# Phase type survival trees

- A powerful non-parametric method of clustering survival data for prognostication

# Phase type survival trees

- A powerful non-parametric method of clustering survival data for prognostication
- To determine importance and effect of various covariates (such as patient's characteristics)

# Phase type survival trees

- A powerful non-parametric method of clustering survival data for prognostication
  - To determine importance and effect of various covariates (such as patient's characteristics)
  - Their interrelation on patient's survival, treatment outcome, disease risk, disease progress or hospital length of stay

# Phase type survival tree

- Each node of *the survival tree* is separately modeled by *phase type distributions*

# Phase type survival tree

- Each node of *the survival tree* is separately modeled by *phase type distributions*
- It combines the merits of both phase type distributions and survival trees.

# Phase type survival tree

- Each node of *the survival tree* is separately modeled by *phase type distributions*
- It combines the merits of both phase type distributions and survival trees.
- Reduces the dimensionality of data and explains the variations in the data.





# Tree construction



Two steps

**Growing:** splitting a node into child nodes



# Tree construction



Two steps

**Growing:** splitting a node into child nodes

**Selection:** determining if a node is terminal node. If it is not then selecting the best possible partition by exploring all possible splits.



# Tree growing

**Growing:** by recursively partitioning into sub groups by the covariates based on some splitting criteria.

At each node apply one covariate at a time and repeat this with other covariates.



# Tree growing

**Splitting criteria:** maximizing either within node homogeneity or between node separation.

We used splitting criteria to maximize within node homogeneity based on improvement of WIC functions

# Tree growing

A covariate  $a$  can have any of the  $l$  values such that

$$N = N_{a1} + N_{a2} + \dots + N_{al} = \sum_{i=1}^l N_{ai} .$$

The loglikelihood of node  $a$  is

$$L = \sum_{j=1}^l \sum_{i=1}^{N_{aj}} f(t_{iaj}) = \sum_{i=1}^{N_{a1}} f(t_{ia1}) + \sum_{i=1}^{N_{a2}} f(t_{ia2}) + \dots + \sum_{i=1}^{N_{al}} f(t_{ial})$$

Or

$$L = L_{a1} + L_{a2} + \dots + L_{al} = \sum_{i=1}^l L_{ai} .$$

# Tree growing

Similarly, WIC of node  $a$  is

$$WIC = WIC_{a_1} + WIC_{a_2} + \dots + WIC_{a_l} = \sum_{i=1}^l WIC_{a_i} .$$



# Node selection



For each possible split of a node, record the total WIC after the split.

The split which maximizes the total WIC of sub-groups is determined as follows:

$$WIC_{\max} = \max(WIC_a, WIC_b, \dots, WIC_l)$$

# Node selection

If  $WIC_{\max}$  is greater than WIC of the node before the split, select the split with WIC equal to  $WIC_{\max}$  else record the node as a terminal node.



# Node selection

If  $WIC_{\max}$  is greater than WIC of the node before the split, select the split with WIC equal to  $WIC_{\max}$  else record the node as a terminal node.

**Terminal node:** A terminal node is the node at which within node homogeneity cannot significantly be improved by any possible split.

# Dataset

To evaluate the model we used the discharge dataset from the Emergency department at the Mater Dei Hospital Malta of all patients discharged in 2011-2014.

# Dataset

We used covariates that represent the patient characteristics:

Age

Gender

District

Source of Admissions

# Dataset

For the length of stay :

The continuous covariate was the patient's age

Three categorical covariates Gender, District and Source of Admission.

# Dataset

Categorical covariate data was divide in three groups.

The cut points of the age are:

1 to 40,

41 to 70 and

71 and over.

Patients with 0 age at admission were omitted from the data.

# Dataset

The gender covariate has two different values that are Female and Male.

# Dataset

The gender covariate has two different values that are Female and Male.

The district covariate has six different values that are the geographical districts of Malta.

# Dataset

The gender covariate has two different values that are Female and Male.

The district covariate has six different values that are the geographical districts of Malta.

Source of admission is from where the patient was admitted and has five different covariates.



# Dataset

The gender covariate has two different values that are Female and Male.

The district covariate has six different values that are the geographical districts of Malta.

Source of admission is from where the patient was admitted and has five different covariates.

Each cluster was given a group number for running the Coxian Phase fittings.

# Dataset

For the admissions:

The categorical covariate was the district of the patient and

# Dataset

For the admissions:

The categorical covariate was the district of the patient and

The categorical covariates are the age and the gender.

# Dataset

For the admissions:

The categorical covariate was the district of the patient and

The categorical covariates are the age and the gender.

Each value in the covariate is given a group number to run the Coxian phase fittings for each group.

# LOS-Phase type Survival tree

Node	Covariate	Covariate Value	Total Number of Patients	WIC	Mean LOS	Number of phases	Total WIC	Gain in WIC	
Level 1									
1 Root Node	All	Root Node	64439	351604.66	6.8411	6	351604.66	-	
	Age	<b>1 to 40</b>	<b>20631</b>	<b>87222.35</b>	<b>4.1304</b>	<b>6</b>	<b>341295.6</b>	<b>10309.1</b>	
		<b>41 to 70</b>	<b>22600</b>	<b>122877.8</b>	<b>6.7443</b>	<b>5</b>			
		<b>71 +</b>	<b>21208</b>	<b>131195.4</b>	<b>9.5813</b>	<b>5</b>			
	District	South		22237	121077.72	6.756	5	351775.15	-170.49
		Central		19480	107177.13	6.9864	4		
		West		8423	46460.1	7.0515	5		
		North		13542	72716.7	6.6032	4		
		Gozo		539	3227.25	8.3358	5		
		Unknown		218	1116.25	5.5	4		
	Source	Elderly Home		1925	11775.05	9.4732	6	351078.46	526.2
		Home		61356	332501.72	6.7339	6		
		Labour Ward		2	32.84	4.5	6		
		Other (Gov Hospital, Private, Mental and Abroad)		1060	6297.08	8.4632	6		
Police Custody			96	471.77	4.7604	2			
Gender	Female		32886	179393.48	6.8672	6	351637.51	-32.85	
	Male		31553	172244.02	6.814	5			

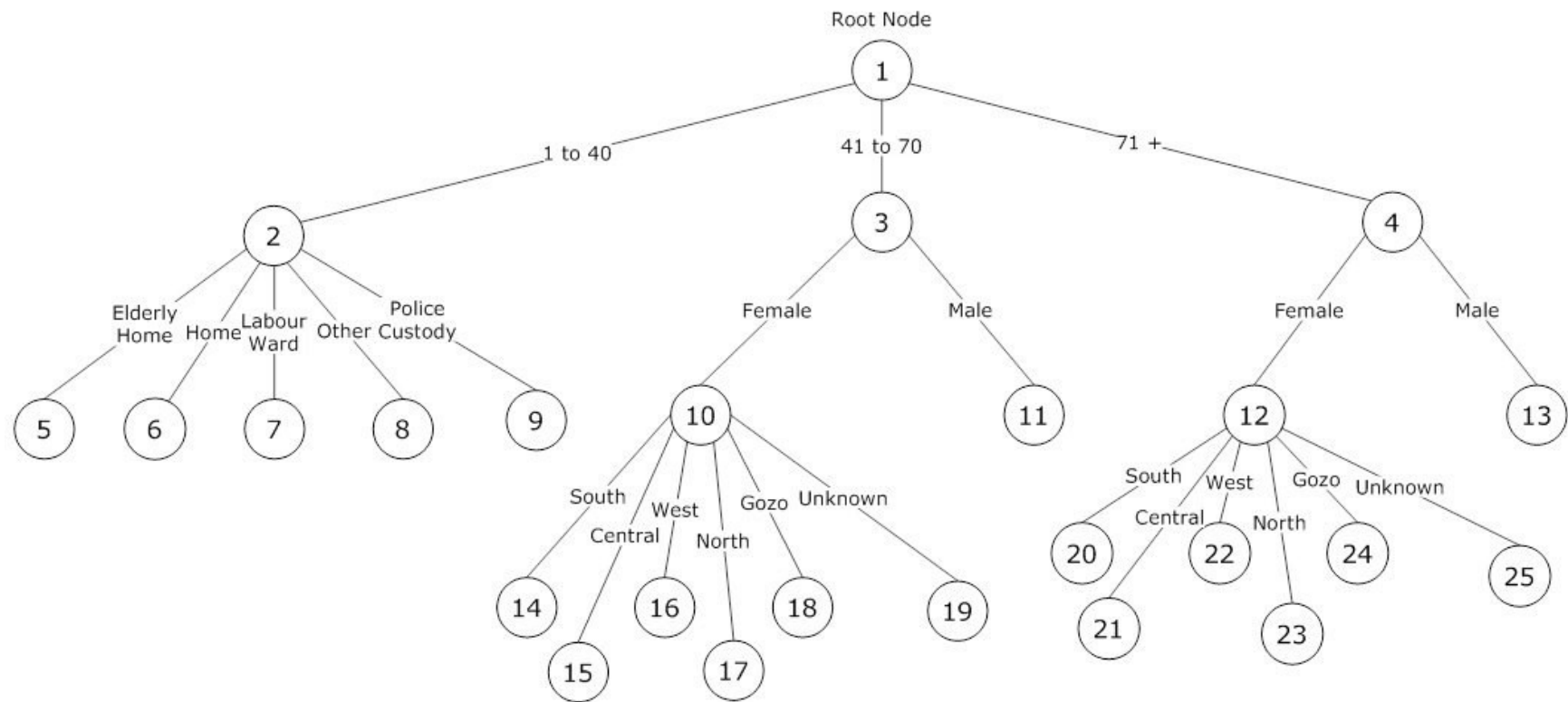
# LOS-Phase type Survival tree

Node	Covariate	Covariate Value	Total Number of Patients	WIC	Mean LOS	Number of phases	Total WIC	Gain in WIC
Level 3								
10 (Age 41 to 70, Female)	All	Age 41 to 70, Female	9088	49410.24	6.817	4	49410.24	-
	District	<b>41 to 70, South, F</b>	<b>3164</b>	<b>17051</b>	<b>6.8587</b>	<b>6</b>	49148.34	261.9
		<b>41 to 70, Central, F</b>	<b>2782</b>	<b>15094.21</b>	<b>6.8724</b>	<b>5</b>		
		<b>41 to 70, West, F</b>	<b>1123</b>	<b>6118.53</b>	<b>6.9154</b>	<b>5</b>		
		<b>41 to 70, North, F</b>	<b>1933</b>	<b>10357.31</b>	<b>6.5525</b>	<b>3</b>		
		<b>41 to 70, Gozo, F</b>	<b>55</b>	<b>366.03</b>	<b>9.9454</b>	<b>1</b>		
		<b>41 to 70, Unknown, F</b>	<b>31</b>	<b>161.25</b>	<b>4.9678</b>	<b>3</b>		
	Source of Admission	41 to 70, Elderly Home, F	81	561.03	12.4445	3	49396.46	13.78
		41 to 70, Home, F	8835	47791.27	6.7268	4		
		41 to 70, Labour Ward, F	1	3.89	7	1		
		41 to 70, Other (Gov Hospital, Private, Mental and Abroad), F	170	1038.88	8.8529	4		
		41 to 70, Police Custody, F	1	1.39	2	1		

# LOS-Phase type Survival tree

Node	Covariate	Covariate Value	Total Number of Patients	WIC	Mean LOS	Number of phases	Total WIC	Gain in WIC
Level 3								
12 (Age 71 +, Female)	All	Age 71 +, Female	11578	72543.24	9.9719	5	72543.24	-
	District	<b>71 +, South, F</b>	<b>3663</b>	<b>22859.81</b>	<b>9.8444</b>	<b>6</b>	<b>72219.66</b>	<b>323.58</b>
		<b>71 +, Central, F</b>	<b>3880</b>	<b>24104.55</b>	<b>9.8023</b>	<b>6</b>		
		<b>71 +, West, F</b>	<b>1736</b>	<b>11040.79</b>	<b>10.4919</b>	<b>4</b>		
		<b>71 +, North, F</b>	<b>2242</b>	<b>13837.2</b>	<b>10.0589</b>	<b>6</b>		
		<b>71 +, Gozo, F</b>	<b>40</b>	<b>287.23</b>	<b>12.825</b>	<b>1</b>		
		<b>71 +, Unknown, F</b>	<b>17</b>	<b>90.08</b>	<b>4.8235</b>	<b>1</b>		
	Source of Admission	71 +, Elderly Home, F	1257	7655.41	9.4121	4	72532.89	10.35
		71 +, Home, F	10093	63415.04	10.0396	6		
		71 +, Other (Gov Hospital, Private, Mental and Abroad), F	228	1462.44	10.057	4		

# LOS-Phase type Survival tree





# Admissions Phase-Type Survival Tree Construction

Node	Covariate	Covariate Value	Total Admissions	WIC	Mean	Number of Phases	Average WIC	Total WIC	Gain in WIC
Level 1									
1 (Root Node)	All	Root Node	32277	3171.43	89.43	22	3171.43	3171.43	-
	Age	1 to 40	10386	2561.57	29.45	10	853.86	2576.47	594.96
		41 to 70	11244	2590.39	31.81	10	863.46		
		71 +	10647	2577.45	30.17	10	859.15		
	Gender	Female	16510	2793.52	44.2	10	1396.76	2811.39	360.04
		Male	15767	2829.26	46.23	10	1414.63		
	District	South	11211	2581.18	31.72	10	430.2	1756.39	1415.04
		Central	9690	2491.79	27.55	10	415.3		
		West	4270	2051.09	12.7	10	341.85		
		North	6774	2289.19	19.56	10	381.53		
		Gozo	289	895.58	1.79	6	149.26		
		Unknown	43	229.51	1.12	10	38.25		

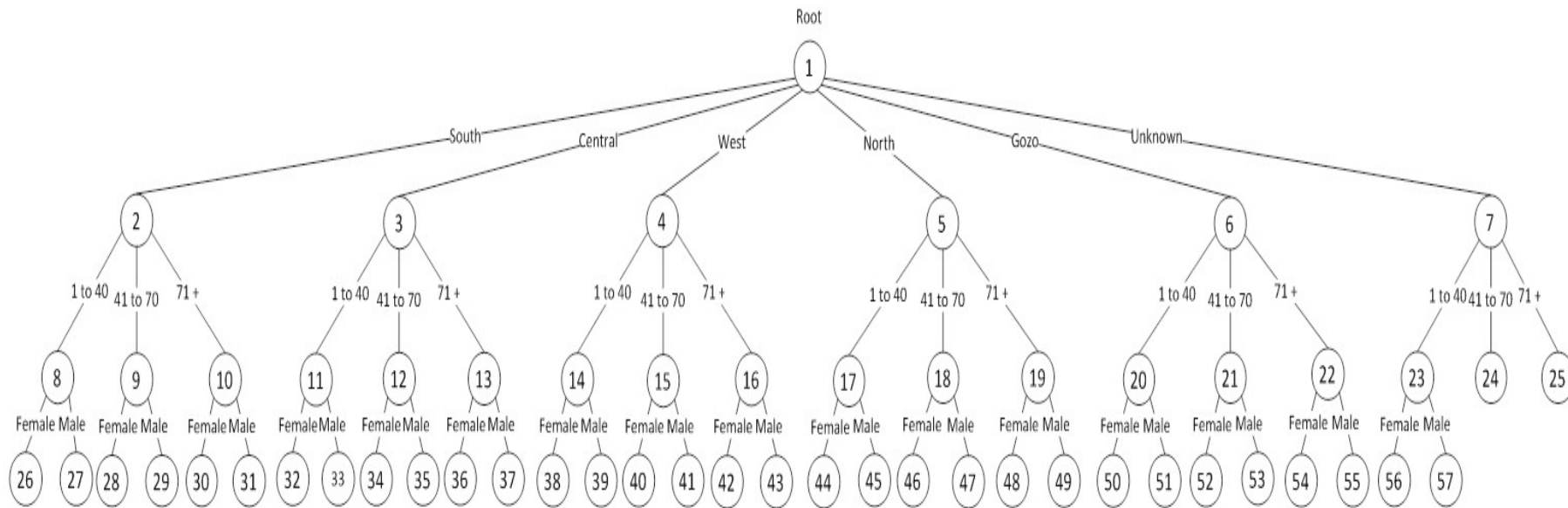
# Admissions Phase-Type Survival Tree Construction

Node	Covariate	Covariate Value	Total Admissions	WIC	Mean	Number of Phases	Average WIC	Total WIC	Gain in WIC
Level 3									
8 (South, 1 to 40)	Gender	Female	2263	1817.71	7.2	5	50.49	94.97	17.71
		Male	1518	1601.38	5.16	5	44.48		
9 (South, 41 to 70)	Gender	Female	1602	1617.75	5.39	5	44.94	94.31	18.11
		Male	2413	1777.52	7.61	7	49.38		
10 (South, 71 +)	Gender	Female	1804	1680.7	5.94	5	46.69	91.78	17.34
		Male	1611	1623.45	5.41	5	45.1		
11 (Central, 1 to 40)	Gender	Female	1761	1719.87	5.82	5	47.77	89.34	16.18
		Male	1191	1496.32	4.26	5	41.56		
12 (Central, 41 to 70)	Gender	Female	1325	1565.73	4.63	5	43.49	91.16	17.21
		Male	1942	1716.2	6.32	6	47.67		
13 (Central, 71 +)	Gender	Female	1934	1725.28	6.3	5	47.92	92.36	18.68
		Male	1537	1599.83	5.21	5	44.44		
14 (West, 1 to 40)	Gender	Female	820	1357.36	3.25	4	37.7	69.44	18.49
		Male	506	1142.3	2.39	4	31.73		
15 (West, 41 to 70)	Gender	Female	565	1200.36	2.55	4	33.34	70.41	16.41
		Male	840	1334.26	3.3	4	37.06		

# Admissions Phase-Type Survival Tree Construction

Node	Covariate	Covariate Value	Total Admissions	WIC	Mean	Number of Phases	Average WIC	Total WIC	Gain in WIC
Level 3									
16 (West, 71 +)	Gender	Female	908	1387.43	3.49	4	38.54	71.95	18.17
		Male	631	1202.62	2.73	4	33.41		
17 (North, 1 to 40)	Gender	Female	1304	1563.15	4.57	4	43.42	81.14	15.86
		Male	882	1357.83	3.42	4	37.72		
18 (North, 41 to 70)	Gender	Female	959	1411.44	3.63	4	39.21	84.06	17.54
		Male	1469	1614.66	5.02	5	44.85		
19 (North, 71 +)	Gender	Female	1125	1488.1	4.08	4	41.34	81.41	17.05
		Male	1035	1442.69	3.84	4	40.07		
20 (Gozo, 1 to 40)	Gender	Female	64	323.82	1.18	10	8.99	16.17	12.16
		Male	50	258.44	1.14	10	7.18		
21 (Gozo, 41 to 70)	Gender	Female	64	323.82	1.18	10	8.99	20.15	9.06
		Male	82	401.76	1.23	10	11.16		
22 (Gozo, 71 +)	Gender	Female	24	100.2	1.07	10	2.78	7.27	9.26
		Male	35	161.34	1.1	10	4.48		
23 (Unknown, 1 to 40)	Gender	Female	13	22.86	1.04	10	0.64	1.89	5.29
		Male	14	45.21	1.04	10	1.26		

# Admissions Phase-Type Survival Tree Construction



# Phase-Type Survival Tree Construction

- The Length of Stay phase-type survival tree has 19 leaf nodes and has a total Gain in WIC of 12619.16.

# Phase-Type Survival Tree Construction

- The Length of Stay phase-type survival tree has 19 leaf nodes and has a total Gain in WIC of 12619.16.
- The Admissions phase-type survival tree has 34 leaf nodes and a total Gain in WIC of 2111.41.

# Prognostication

- Both phase-type survival trees are showing
  - Analysis of the determined patient groups from our dataset.

# Prognostication

- Predictions can be made from the data used to construct the Phase-type survival tree
  - For the number of admissions by the patient grouping and



# Prognostication

- Predictions can be made from the data used to construct the Phase-type survival tree
  - For the number of admissions by the patient grouping and
  - We can predict the LOS of a patient by his/her characteristics.

# LOS-Prediction

Gender	Age	District	Source	Admission Date	Discharge Date	Actual LOS	Predicted LOS
M	1	South	Home	15/12/2012	19/12/2012	5	4.122102
M	67	Central	Home	21/12/2012	31/12/2012	11	6.744455
F	86	South	Home	18/12/2012	24/12/2012	7	9.960199
F	24	West	Home	22/12/2012	24/12/2012	3	4.122102
M	64	South	Home	15/12/2012	18/12/2012	4	6.744455
M	77	West	Elderly Home	26/12/2012	31/12/2012	6	9.189538
M	16	North	Home	20/12/2012	20/12/2012	1	4.122102
F	94	South	Home	18/12/2012	20/12/2012	3	9.960199
M	57	Central	Home	15/12/2012	19/12/2012	5	6.744455
F	49	Central	Home	20/12/2012	21/12/2012	2	6.916771

# Admission Predictions

Admissions Date	Group	Actual Admissions	Predicted Admissions
31/12/2011	41 to 70 Unknown	0	0.04
28/12/2011	1 to 40, South, Male	3	4.15
24/12/2011	1 to 40, Central Males	2	3.28
28/12/2011	1 to 40, West, Males	1	1.39
26/12/2011	1 to 40, North, Males	1	2.45
27/12/2011	1 to 40, Gozo, Males	0	0.14
27/12/2011	1 to 40, Unknown, Males	0	0.04
19/12/2011	1 to 40, South, Females	7	6.30
29/12/2011	1 to 40, Central, Females	3	4.87
30/12/2011	1 to 40, West, Females	2	2.28
28/12/2011	1 to 40, North, Females	5	3.59
24/12/2011	1 to 40, Gozo, Females	0	0.18
24/12/2011	1 to 40, Unknown, Females	0	0.03
28/12/2011	41 to 70, South, Males	12	6.54
19/12/2011	41 to 70, Central, Males	7	5.30
26/12/2011	41 to 70, West, Males	5	2.27
15/12/2011	41 to 70, North, Males	6	2.61

# Admission Predictions

Admissions Date	Group	Actual Admissions	Predicted Admissions
29/12/2011	41 to 70, Gozo, Males	0	0.22
23/12/2011	41 to 70, South, Females	7	4.39
29/12/2011	41 to 70, Central, Females	3	3.63
20/12/2011	41 to 70, West, Females	0	1.59
25/12/2011	41 to 70, North, Females	4	4.02
28/12/2011	41 to 70, Gozo, Females	0	0.18
24/12/2011	71 +, South, Males	8	4.41
30/12/2011	71 +, Central, Males	4	4.16
31/12/2011	71 +, West, Males	1	1.71
17/12/2011	71 +, North, Males	3	2.80
26/12/2011	71 +, Gozo, Males	1	0.10
17/12/2011	71 +, South, Females	4	4.87
16/12/2011	71 +, Central, Females	6	5.15
30/12/2011	71 +, West, Females	3	2.47
16/12/2011	71 +, North, Females	3	3.06
31/12/2011	71 +, Gozo, Females	0	0.07

# Construction of Phase-Type Survival Tree showing Effect of Weather on LOS

Level	Covariate	Covariate Group	No. of Patients	Mean LOS	WIC	Total WIC	WIC Gain
1 (Root)	<i>All</i>	<i>Root</i>	<i>66166</i>	<i>6.88</i>	<i>361646.80</i>	<i>361646.80</i>	
	MinTemp	0°C-10°C (1)	16465	7.19	91916.01	361631.50	15.30
		11°C-20°C (2)	33516	6.76	181607.62		
		21°C-30°C (3)	16185	6.83	88107.87		
	MaxTemp	0°C-10°C (1)	303	8.13	1786.56	349779.14	11867.67
		11°C-20°C (2)	28333	6.95	143924.01		
		21°C-30°C (3)	25205	6.83	137012.30		
		31+°C (4)	12325	6.82	67056.27		
	AvgTemp	0°C-10°C (1)	4834	7.23	26828.01	361381.17	265.63
		11°C- 20°C (2)	34493	6.87	188586.75		
		21°C-30°C (3)	26090	6.83	141956.96		
		31+°C (4)	749	6.88	4009.44		
	MaxVar	$x < -2^{\circ}\text{C}$ (1)	4032	7.02	22086.49	361419.43	227.37
		$-2^{\circ}\text{C} \leq x \leq -1^{\circ}\text{C}$ (2)	18199	6.78	99118.57		
		0°C (3)	19042	6.79	103741.30		
		$1^{\circ}\text{C} \leq x \leq 2^{\circ}\text{C}$ (4)	21365	7.02	117284.96		
$x > 2^{\circ}\text{C}$ (5)		3528	6.88	19188.12			

# Construction of Phase-Type Survival Tree showing Effect of Weather on LOS

Level	Covariate	Covariate Group	No. of Patients	Mean LOS	WIC	Total WIC	WIC Gain
2 (0°C-10°C Max)	All	0°C-10°C (1)	303	8.13	1786.56	1786.56	
	MinTemp	0°C-10°C (1)	303	8.13	1786.56	1786.56	0.00
		11°C-20 (2)	0	0.00	0.00		
		21°C-30°C (3)	0	0.00	0.00		
	AvgTemp	0°C-10°C (1)	303	8.13	1786.56	1786.56	0.00
		11°C-20°C (2)	0	0.00	0.00		
		21°C-30°C (3)	0	0.00	0.00		
		31+°C (4)	0	0.00	0.00		
	MaxVar	x < -2°C (1)	104	9.50	619.01	1809.80	-23.24
		-2°C ≤ x ≤ -1°C (2)	97	7.59	584.81		
		0°C (3)	102	7.25	605.99		
		1°C ≤ x ≤ 2°C (4)	0	0.00	0.00		
		x > 2°C (5)	0	0.00	0.00		



# Construction of Phase-Type Survival Tree showing Effect of Weather on LOS

Level	Covariate	Covariate Group	No. of Patients	Mean LOS	WIC	Total WIC	WIC Gain
2 (11°C-20°C Max)	All	11°C-20°C (2)	28333	6.83	143924.01	143924.01	
	MinTemp	0°C-10°C (1)	15983	7.19	88145.39	154784.63	-10860.62
		11°C-20°C (2)	12350	6.63	66639.23		
		21°C-30°C (3)	0	0.00	0.00		
	AvgTemp	0°C-10°C (1)	4531	7.17	25082.63	155610.36	-11686.34
		11°C-20°C (2)	23802	6.90	130527.73		
		21°C-30°C (3)	0	0.00	0.00		
		31+°C (4)	0	0.00	0.00		
	MaxVar	x < -2°C (1)	1818	6.98	10045.36	154715.89	-10791.88
		-2°C ≤ x ≤ -1°C (2)	8495	6.78	45964.04		
0°C (3)		8287	6.72	44646.01			
1°C ≤ x ≤ 2°C (4)		8551	7.23	47346.72			
x > 2°C (5)		1182	7.59	6713.75			

# Construction of Phase-Type Survival Tree showing Effect of Weather on LOS

Level	Covariate	Covariate Group	No. of Patients	Mean LOS	WIC	Total WIC	WIC Gain
2(21°C-30°C Max)	<i>All</i>	21°C-30°C (3)	25205	6.83	137012.30	137012.30	
	MinTemp	0°C-10°C (1)	179	6.13	967.79	136794.76	217.54
		11°C-20°C (2)	20347	6.83	110539.41		
		21°C-30°C (3)	4679	6.84	25287.56		
	AvgTemp	0°C-10°C (1)	0	0.00	0.00	136265.41	746.88
		11°C-20°C (2)	10691	6.81	57269.63		
		21°C-30°C (3)	14514	6.84	78995.78		
		31+°C (4)	0	0.00	0.00		
	MaxVar	x < -2°C (1)	1203	6.92	6576.60	136579.50	432.80
		-2°C ≤ x ≤ -1°C (2)	6861	6.77	36677.42		
		0°C (3)	7826	6.89	42694.92		
		1°C ≤ x ≤ 2°C (4)	8472	6.88	46191.76		
		x > 2°C (5)	843	6.04	4438.80		



# Construction of Phase-Type Survival Tree showing Effect of Weather on LOS

Level	Covariate	Covariate Group	No. of Patients	Mean LOS	WIC	Total WIC	WIC Gain
3(21°C-30°C Max, 11°C-20°C Avg)	All	11°C-20°C (2)	10691	6.81	57269.63	57269.63	
	MinTemp	0°C-10°C (1)	179	6.13	961.29	58083.83	-814.20
		11°C-20°C (2)	10512	6.82	57122.54		
		21°C-30°C (3)	0	0.00	0.00		
	MaxVar	x < -2°C (1)	397	5.66	2036.82	57493.12	-223.49
		-2°C ≤ x ≤ -1°C (2)	2405	6.95	13061.90		
0°C (3)		2666	6.88	14272.56			
1°C ≤ x ≤ 2°C (4)		4736	6.87	25550.60			
x > 2°C (5)		487	6.00	2571.25			
3(21°C-30°C Max, 21°C-30°C Avg)	All	21°C-30°C (3)	14514	6.84	78995.78	78995.78	
	MinTemp	0°C-10°C (1)	0	0.00	0.00	78354.56	641.22
		11°C-20°C (2)	9835	6.85	52787.28		
		21°C-30°C (3)	4679	6.84	25567.28		
	MaxVar	x < -2°C (1)	806	7.54	4555.23	78571.35	424.43
		-2°C ≤ x ≤ -1°C (2)	4456	6.68	23671.22		
0°C (3)		5160	6.90	28290.46			
1°C ≤ x ≤ 2°C (4)		3736	6.88	20163.21			
x > 2°C (5)		356	6.10	1891.23			

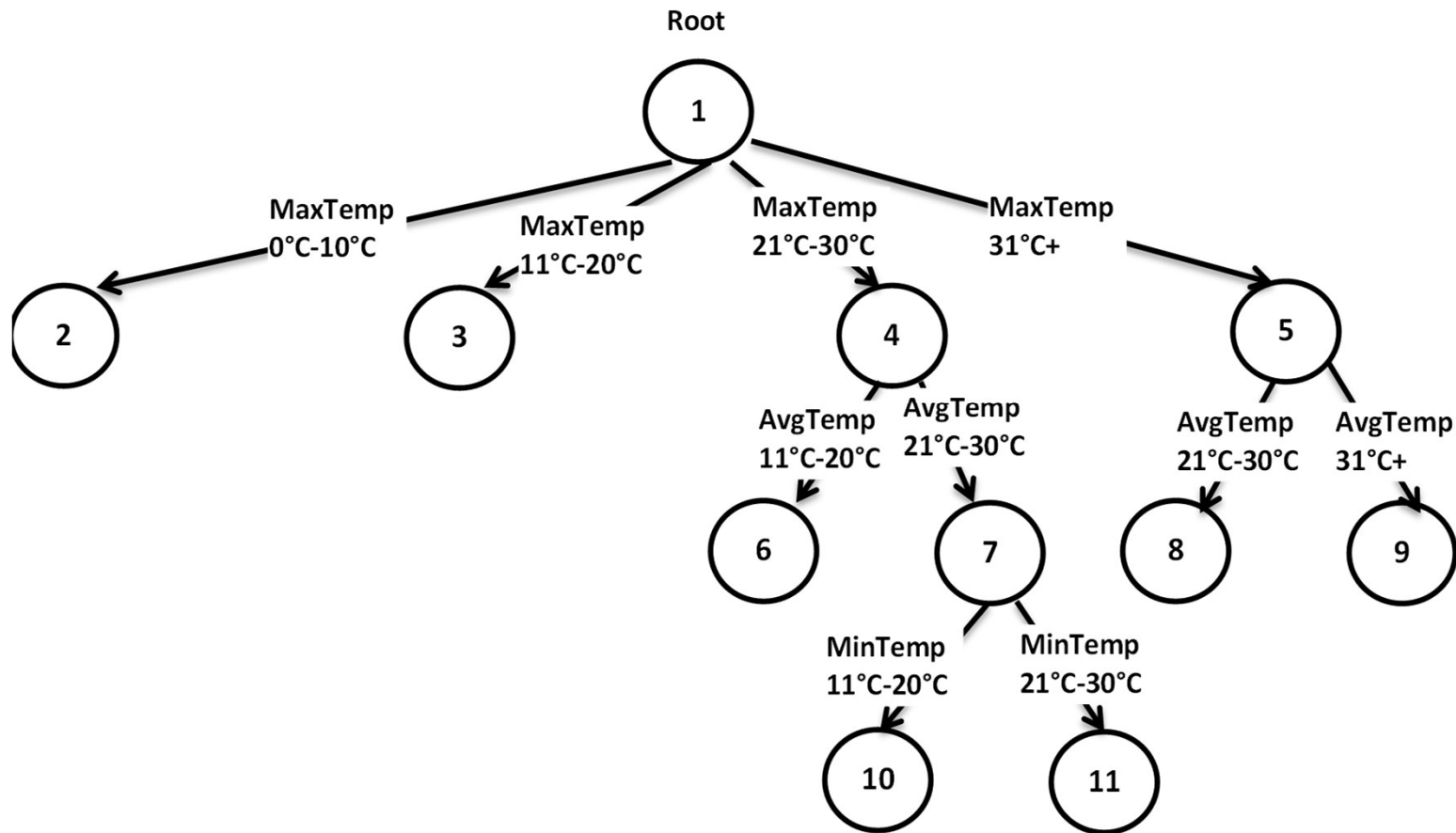
# Construction of Phase-Type Survival Tree showing Effect of Weather on LOS

Level	Covariate	Covariate Group	No. of Patients	Mean LOS	WIC	Total WIC	WIC Gain
2(31+°C Max)	<i>All</i>	<i>31+°C (4)</i>	<i>12325</i>	<i>6.82</i>	<i>67056.27</i>	<i>67056.27</i>	
	MinTemp	0°C-10°C (1)	0	0.00	0.00	67053.54	2.73
		11°C-20°C (2)	820	6.70	4466.70		
		21°C-30°C (3)	11505	6.83	62586.84		
	AvgTemp	0°C-10°C (1)	0	0.00	0.00	66238.27	818.00
		11°C-20°C (2)	0	0.00	0.00		
		21°C-30°C (3)	11576	6.82	62203.93		
		31+°C (4)	749	6.88	4034.35		
	MaxVar	x < -2°C (1)	907	6.96	4967.36	66443.03	613.24
		-2°C ≤ x ≤ -1°C (2)	2746	6.79	14698.72		
		0°C (3)	2827	6.70	15140.27		
		1°C ≤ x ≤ 2°C (4)	4342	6.90	23535.51		
		x > 2°C (5)	1503	6.80	8101.17		

# Construction of Phase-Type Survival Tree showing Effect of Weather on LOS

Level	Covariate	Covariate Group	No. of Patients	Mean LOS	WIC	Total WIC	WIC Gain	
3(31+°C Max, 21°C-30°C Avg)	<i>All</i>	<i>21°C-30°C (3)</i>	<i>11576</i>	<i>6.82</i>	<i>62203.93</i>	<i>62203.93</i>		
	MinTemp	0°C-10°C (1)	0	0.00	0.00	63023.76	-819.83	
		11°C-20°C (2)	820	6.70	4466.70			
		21°C-30°C (3)	10756	6.83	58557.06			
	MaxVar	x < -2°C (1)	820	7.04	4489.42	62465.66	-261.73	
		-2°C ≤ x ≤ -1°C (2)	2546	6.80	13659.83			
			0°C (3)	2827	6.70	15140.27		
			1°C ≤ x ≤ 2°C (4)	4243	6.90	23001.58		
			x > 2°C (5)	1140	6.68	6174.55		
	3(31+°C Max, 31+°C Avg)		<b>31+°C (4)</b>	<b>749</b>	<b>6.88</b>	<b>4034.35</b>	<b>4034.35</b>	
MinTemp		0°C-10°C (1)	0	0.00	0.00	4061.80	-27.45	
		11°C-20°C (2)	0	0.00	0.00			
		21°C-30°C (3)	749	6.88	4061.80			
MaxVar		x < -2°C (1)	87	6.20	485.00	4082.57	-48.23	
		-2°C ≤ x ≤ -1°C (2)	200	6.62	1074.44			
		0°C (3)	0	0.00	0.00			
	1°C ≤ x ≤ 2°C (4)	99	6.93	554.99				
		x > 2°C (5)	363	7.17	1968.14			

# Phase-Type Survival Tree showing Effect of Weather on LOS



# Phase-Type Survival Tree showing Effect of Weather on LOS

- Most significant prognostic factor affecting the patients' length of stay (LOS) is the maximum temperature.



# Phase-Type Survival Tree showing Effect of Weather on LOS

- Most significant prognostic factor affecting the patients' length of stay (LOS) is the maximum temperature.
- The average temperature affects the patients' length of stay only when the maximum temperature rises beyond 20°C.

# Phase-Type Survival Tree showing Effect of Weather on LOS

- The minimum temperature does not significantly affect the patients' length of stay.

# Phase-Type Survival Tree showing Effect of Weather on LOS

- The minimum temperature does not significantly affect the patients' length of stay.
- Also, the maximum variability in the average temperature between one day and the next does not affect patients' length of stay as patients usually stay inside.



# Phase-Type Survival Tree showing Effect of Weather on LOS

- These results might be different for different geographic regions due to different weather conditions and different genetic profile of inhabitants there.

# Phase-Type Survival Tree showing Effect of Weather on LOS

## Predictions and Accuracy Tests

Group	No. of Patients	Actual Mean LOS	Predicted Mean LOS	Forecast Error	Squared Error	Absolute Error	Percentage Error (%)
MaxTemp(0°C-10°C)	0	-	8.13	-	-	-	-
MaxTemp(11°C-20°C)	13406	7.19	6.83	-0.36	0.13	0.36	5.01
MaxTemp(21°C-30°C), AvgTemp(11°C-20°C)	6003	7.01	6.81	-0.20	0.04	0.20	2.85
MaxTemp(21°C-30°C), AvgTemp(21°C-30°C), MinTemp(11°C-20°C)	5850	6.78	6.85	0.07	0.00	0.07	1.03
MaxTemp(21°C-30°C), AvgTemp(21°C-30°C), MinTemp(21°C-30°C)	4520	6.47	6.84	0.37	0.14	0.37	5.72
MaxTemp(31+°C), AvgTemp(21°C-30°C)	0	-	6.82	-	-	-	-
MaxTemp(31+°C), AvgTemp(31+°C)	4471	6.72	6.88	0.16	0.03	0.16	2.38

# Construction of Phase-Type Survival Tree showing Effect of Weather on Admissions

Level	Covariate	Covariate Group	No. of Records	Mean Admissions	WIC	Average WIC	Total Average WIC	WIC Gain
I (Root)	ALL	Root	721	91.04	6522.86	6522.86	6522.86	
	Min	0°C-10°C (1)	174	94.63	1653.37	551.12	2249.21	4273.65
		11°C-20°C (2)	376	89.14	3421.33	1140.44		
		21°C-30°C (3)	181	89.41	1672.91	557.64		
	Max	0°C-10°C (1)	3	101.00	38.93	9.73	1690.81	4832.05
		11°C-20°C (2)	306	92.59	2848.41	712.10		
		21°C-30°C (3)	283	89.07	2580.33	645.08		
		31+°C (4)	139	88.67	1295.56	323.89		
	Avg	0°C-10°C (1)	49	98.65	495.38	123.84	1690.54	4832.32
		11°C- 20°C (2)	379	91.01	3490.43	872.61		
		21°C-30°C (3)	295	88.44	2685.31	671.33		
		31+°C (4)	8	93.62	91.04	22.76		
	MaxVar	$x < -2^{\circ}\text{C}$ (1)	45	89.60	449.32	89.86	1369.18	5153.68
		$-2^{\circ}\text{C} \leq x \leq -1^{\circ}\text{C}$ (2)	200	91.00	1867.44	373.49		
		$0^{\circ}\text{C}$ (3)	212	89.82	1956.63	391.33		
$1^{\circ}\text{C} \leq x \leq 2^{\circ}\text{C}$ (4)		236	90.53	2186.84	437.37			
$x > 2^{\circ}\text{C}$ (5)		38	92.87	385.68	77.14			

# Construction of Phase-Type Survival Tree showing Effect of Weather on Admissions

Level	Covariate	Covariate Group	No. of Records	Mean Admissions	WIC	Average WIC	Total Average WIC	WIC Gain
2(MaxVar, $x < -2^{\circ}\text{C}$ (1))	All	$x < -2^{\circ}\text{C}$ (1)	45	89.60	449.32	89.86	89.86	
	Min	0°C-10°C (1)	11	91.64	121.58	40.53	161.37	-71.51
		11°C-20°C (2)	19	87.37	200.39	66.80		
		21°C-30°C (3)	15	90.93	162.15	54.05		
	Max	0°C-10°C (1)	1	104.00	7.07	1.77	120.11	-30.25
		11°C-20°C (2)	20	90.90	211.72	52.93		
		21°C-30°C (3)	14	85.93	150.35	37.59		
		31+°C (4)		90.70	111.31	27.83		
	Avg	0°C-10°C (1)	6	91.83	68.79	17.20	119.49	-29.63
		11°C-20°C (2)	20	88.40	210.66	52.66		
		21°C-30°C (3)	18	90.33	191.81	47.95		
		31+°C (4)	1	87.00	6.71	1.68		



# Construction of Phase-Type Survival Tree showing Effect of Weather on Admissions

Level	Covariate	Covariate Group	No. of Records	Mean Admissions	WIC	Average WIC	Total Average WIC	WIC Gain
2(MaxVar, $-2^{\circ}\text{C} \leq x \leq -1^{\circ}\text{C}$ (2))	All	$-2^{\circ}\text{C} \leq x \leq -1^{\circ}\text{C}$ (2)	200	91.00	1867.44	373.49	373.49	
	Min	0°C-10°C (1)	44	96.32	454.40	151.47	650.97	-277.48
		11°C-20°C (2)	106	89.50	1003.42	334.47		
		21°C-30°C (3)	50	89.48	495.09	165.03		
	Max	0°C-10°C (1)	1	97.00	6.93	1.73	487.78	-114.29
		11°C-20°C (2)	92	92.34	896.50	224.12		
		21°C-30°C (3)	76	90.28	730.09	182.52		
		31+°C (4)	31	88.58	317.61	79.40		
	Avg	0°C-10°C (1)	1	99.63	175.32	43.83	484.97	-111.48
		11°C-20°C (2)	103	91.29	992.95	248.24		
		21°C-30°C (3)	79	88.63	761.01	190.25		
		31+°C (4)	2	100.00	10.59	2.65		

# Construction of Phase-Type Survival Tree showing Effect of Weather on Admissions

Level	Covariate	Covariate Group	No. of Records	Mean Admissions	WIC	Average WIC	Total Average WIC	WIC Gain
2(MaxVar, 0°C (3))	All	0°C (3)	212	89.82	1956.63	391.33	391.33	
	Min	0°C-10°C (1)	60	93.87	593.16	197.72	682.65	-291.33
		11°C-20°C (2)	109	88.20	1025.83	341.94		
		21°C-30°C (3)	43	88.28	428.98	142.99		
	Max	0°C-10°C (1)	1	102.00	7.17	1.79	510.66	-119.34
		11°C-20°C (2)	90	92.08	864.84	216.21		
		21°C-30°C (3)	89	87.93	844.45	211.11		
		31+°C (4)	32	88.34	326.19	81.55		
	Avg	0°C-10°C (1)	14	100.79	154.78	38.70	507.59	-116.27
		11°C-20°C (2)	108	89.30	1021.80	255.45		
		21°C-30°C (3)	90	88.74	853.79	213.45		
		31+°C (4)	0	0.00	0.00	0.00		

# Construction of Phase-Type Survival Tree showing Effect of Weather on Admissions

Level	Covariate	Covariate Group	No. of Records	Mean Admissions	WIC	Average WIC	Total Average WIC	WIC Gain
2(MaxVar, $1^{\circ}\text{C} \leq x \leq 2^{\circ}\text{C}$ (4))	All	$1^{\circ}\text{C} \leq x \leq 2^{\circ}\text{C}$ (4)	236	90.53	2186.84	437.37	437.37	
	Min	0°C-10°C (1)	50	93.86	507.91	169.30	761.50	-324.13
		11°C-20°C (2)	128	89.68	1208.86	402.95		
		21°C-30°C (3)	58	89.56	567.72	189.24		
	Max	0°C-10°C (1)	0	0.00	0.00	0.00	369.35	68.02
		11°C-20°C (2)	92	92.95	896.15	23.24		
		21°C-30°C (3)	95	89.18	899.29	224.82		
		31+°C (4)	49	88.61	485.15	121.29		
	Avg	0°C-10°C (1)	10	99.00	113.07	28.27	563.66	-126.29
		11°C-20°C (2)	134	91.77	1271.13	317.78		
		21°C-30°C(3)	91	87.68	863.45	215.86		
		31+°C (4)	1	99.00	6.97	1.74		

# Construction of Phase-Type Survival Tree showing Effect of Weather on Admissions

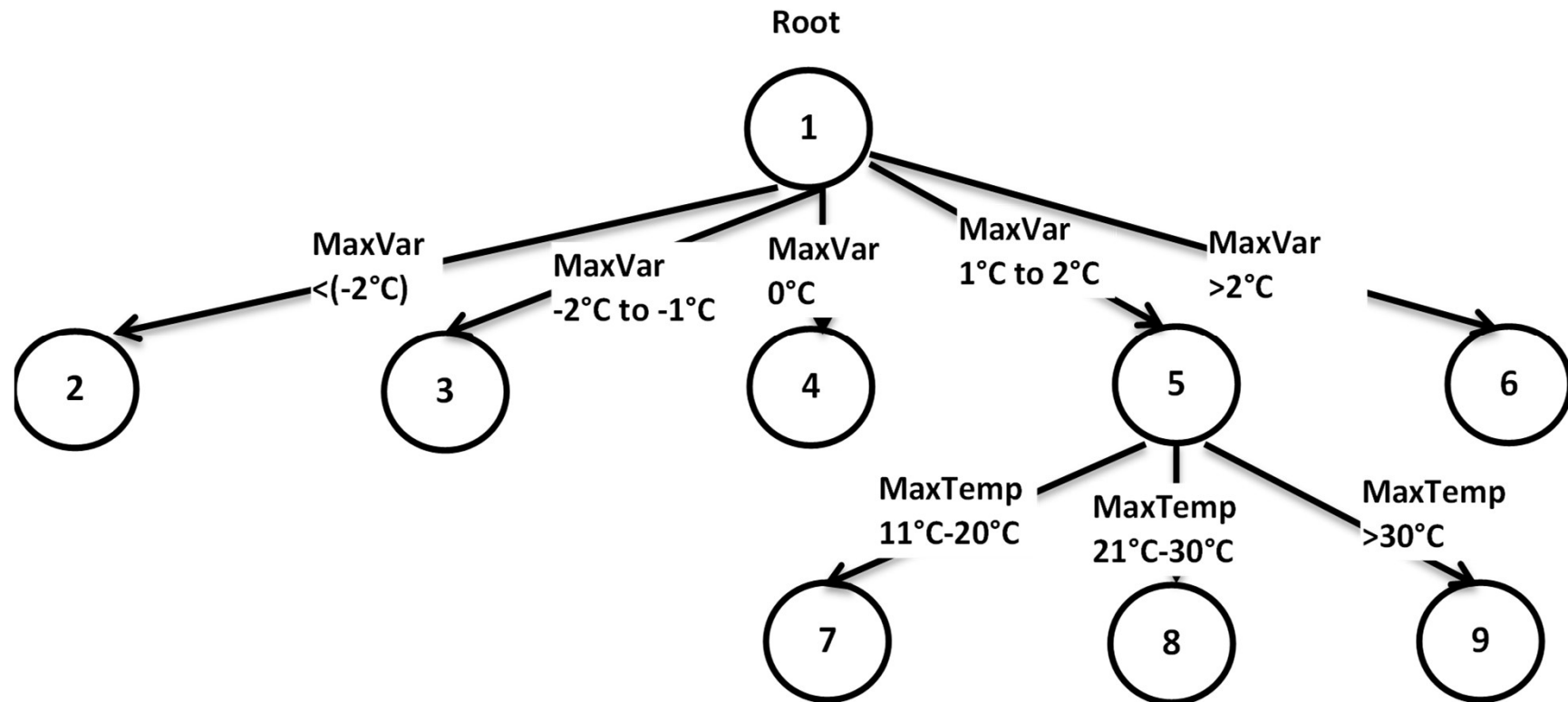
Level	Covariate	Covariate Group	No. of Records	Mean Admissions	WIC	Average WIC	Total Average WIC	WIC Gain
3(MaxVar (4), Max (2))	All	11- 20 (2)	92	92.95	896.15	23.24	23.24	
	Min	0°C-10°C (1)	49	93.78	498.37	124.59	233.29	-210.05
		11°C-20°C (2)	43	92.00	434.78	108.69		
		21°C-30°C (3)	0	0.00	0.00	0.00		
	Avg	0°C-10°C (1)	49	93.78	498.37	124.59	233.29	-210.05
		11°C-20°C (2)	43	92.00	434.78	108.69		
		21°C-30°C (3)	0	0.00	0.00	0.00		
31+°C (4)		0	0.00	0.00	0.00			
3(MaxVar (4), Max (3))	All	21-30 (3)	95	89.18	899.29	224.82	224.82	
	Min	0°C-10°C (1)	1		BAD WIC		BAD WIC	BAD WI
		11°C-20°C (2)	81	89.05	773.46	257.82		
		21°C-30°C (3)	13	89.31	141.15	47.05		
	Avg	0°C-10°C (1)	0	0.00	0.00	0.00	235.24	-10.41
		11°C-20°C (2)	52	91.08	512.81	128.20		
		21°C-30°C (3)	43	86.88	428.13	107.03		
31+°C (4)		0	0.00	0.00	0.00			



# Construction of Phase-Type Survival Tree showing Effect of Weather on Admissions

Level	Covariate	Covariate Group	No. of Records	Mean Admissions	WIC	Average WIC	Total Average WIC	WIC Gain
3(MaxVar (4), Max (4))	All	31+ (4)	49	88.61	485.15	121.29	121.29	
	Min	0°C-10°C (1)	0	0.00	0.00	0.00	164.93	-43.64
		11°C-20°C (2)	4	77.50	46.13	15.38		
		21°C-30°C (3)	45	89.60	448.66	149.55		
	Avg	0°C-10°C (1)	0	0.00	0.00	0.00	BAD WIC	BAD WI
		11°C-20°C (2)	0	0.00	0.00	0.00		
		21°C-30°C (3)	48	88.40	476.47	119.12		
31+°C (4)		1		BAD WIC				
2(MaxVar, x > 2°C (5))	All	$x > 2^\circ C$ (5)	38	92.87	385.68	77.14	77.14	
	Min	0°C-10°C (1)	9	99.33	103.12	34.37	138.99	-61.85
		11°C-20°C (2)	14	91.29	151.95	50.65		
		21°C-30°C (3)	15	90.47	161.90	53.97		
	Max	0°C-10°C (1)	0	0.00	0.00	0.00	104.15	-27.01
		11°C-20°C (2)	12	98.50	133.30	33.32		
		21°C-30°C (3)	9	93.78	102.12	25.53		
	Avg	31+°C (4)	17	88.41	181.18	45.29	105.29	-28.15
		0°C-10°C (1)	3	96.00	38.63	9.66		
		11°C-20°C (2)	14	98.64	154.06	38.52		
		21°C-30°C (3)	17	88.06	181.07	45.27		
		31+°C (4)	4	90.75	47.39	11.85		

# Phase-Type Survival Tree showing Effect of Weather on Admissions



# Phase-Type Survival Tree showing Effect of Weather on Admissions

- Most significant prognostic factor affecting the number of admissions is the maximum variability in the average temperature between one day and the next.

# Phase-Type Survival Tree showing Effect of Weather on Admissions

- Most significant prognostic factor affecting the number of admissions is the maximum variability in the average temperature between one day and the next.
- The maximum temperature affects the number of admissions only when the average temperature increases by 1°C-2°C than the previous day.

# Phase-Type Survival Tree showing Effect of Weather on Admissions

- The minimum temperature and average temperature do not affect number of admissions.

# Phase-Type Survival Tree showing Effect of Weather on Admissions

- The minimum temperature and average temperature do not affect number of admissions.
- These results might be different for different geographic regions due to different weather conditions and different genetic profile of inhabitants there.

# Phase-Type Survival Tree showing Effect of Weather on Admissions

## Predictions and Accuracy Tests

Group	No. of Records	Actual Mean Adm.	Predicted Mean Adm.	Forecast Error	Squared Error	Absolute Error	Percentage Error (%)
MaxVar( $x < -2^{\circ}\text{C}$ )	31	92.13	89.60	-2.53	6.40	2.53	2.75
MaxVar( $-2^{\circ}\text{C} \leq x \leq -1^{\circ}\text{C}$ )	99	92.34	91.00	-1.34	1.80	1.34	1.45
MaxVar( $x = 0^{\circ}\text{C}$ )	93	92.77	89.82	-2.95	8.70	2.95	3.18
MaxVar( $x > 2^{\circ}\text{C}$ )	19	97.63	92.87	-4.76	22.66	4.76	4.88
MaxVar( $1^{\circ}\text{C} \leq x \leq 2^{\circ}\text{C}$ ), MaxTemp ( $11^{\circ}\text{C} - 20^{\circ}\text{C}$ )	42	100.95	92.95	-8.00	64.00	8.00	7.92
MaxVar( $1^{\circ}\text{C} \leq x \leq 2^{\circ}\text{C}$ ), MaxTemp ( $21^{\circ}\text{C} - 30^{\circ}\text{C}$ )	54	91.63	89.18	-2.45	6.00	2.45	2.67
MaxVar( $1^{\circ}\text{C} \leq x \leq 2^{\circ}\text{C}$ ), MaxTemp ( $31+^{\circ}\text{C}$ )	27	95.48	88.61	-6.87	47.20	6.87	7.20



# Accuracy test for all predictions

		MSE	RMSE	MAD	BIAS
<i>LOS</i>	<i>Weather</i>	0.08	0.28	0.26	-0.09
	<i>Personal Characteristics</i>	1.15	1.07	0.74	-0.69
<i>Admissions</i>	<i>Weather</i>	16.17	4.02	3.37	-3.37
	<i>Personal Characteristics</i>	1.38	1.17	0.96	-0.82

MSE: Mean Square Error,  
RMSE: Root Mean Square Error,  
MAD: Mean Absolute Deviation  
BIAS: Bias



# Conclusions

- We can use phase-type survival tree analysis to
  - Effectively prognosticate survival data and

# Conclusions

- We can use phase-type survival tree analysis to
  - Effectively prognosticate survival data and
  - Cluster survival data into groups of patients following homogeneous patient pathways.

# Conclusions

- Our models can be used to forecast bed occupancy and the requirements.

# Conclusions

- Our models can be used to forecast bed occupancy and the requirements.
- The LOS can be predicted at admission by the use of this model.

# Conclusions

- Our models can be used to forecast bed occupancy and the requirements.
- The LOS can be predicted at admission by the use of this model.
- The number of admissions can be forecasted by the patients' characteristics.

# Conclusions

- These models can also be used to characterize the effect of weather on LOS and admissions.

# Conclusions

- These models can also be used to characterize the effect of weather on LOS and admissions.
- We can also use these models to predict effect of other factors affecting LOS and admissions.

# Conclusions

- These forecasts can help us better designing policies to ensure optimal utilization of scarce health resources.



# References

- Garg L, Dauwels J, Earnest A, Pang L (2013). Tensor based methods for handling missing data in quality-of-life questionnaires. *IEEE Journal of Biomedical and Health Informatics*. In press. doi: 10.1109/JBHI.2013.2288803. Impact Factor: 1.978, Journal H-Index(SJR:1996-2012): 55, SJR(2012):0.877.
- Garg L, McClean SI, Meenan BJ, Barton M, Fullerton K (2012). Intelligent patient management and resource planning for complex, heterogeneous, and stochastic healthcare systems. In press. *IEEE Transactions on Systems, Man, and Cybernetics--Part A: Systems and Humans* 42(6): 1332 – 1345. Impact Factor: 2.093, Journal H-Index(SJR:1996-2011): 54, SJR(2011):0.085, Scopus SNIP(2011):3.594, Scopus citations: 2, ISI Web of Knowledge Citations: 1, scHolar index: 2.
- Garg L, McClean SI, Meenan BJ, Barton M, Fullerton K (2011). An Extended Mixture Distribution Survival Tree for Patient Pathway Prognostication. *Communications in Statistics: Theory and Methodology*. 42(16):2912-2934. Impact Factor: 0.351, Journal H-Index(SJR:1996-2011): 24, SJR(2011):0.035, Scopus SNIP(2011):0.550.
- Barton M, McClean SI, Gillespie J, Garg L, Wilson D, Fullerton K (2012). Is it beneficial to increase the provision of thrombolysis?- A discrete-event simulation model, *QJM: An International Journal of Medicine*. 105(7), (art. no. hcs036):665-673. Impact Factor: 2.146, Journal H-Index(SJR:1993-2011): 71, SJR(2011):0.18, Scopus SNIP(2011):0.897, Scopus citations:3, ISI Web of Knowledge Citations: 1, scHolar index: 2.
- Garg L, McClean SI, Meenan BJ, Millard PH (2011). Phase-type survival trees and mixed distribution survival trees for clustering patients' hospital length of stay. *INFORMATICA*. 22(1): 57-72. Impact Factor: 1.786, Journal H-Index(SJR:1996-2010): 15, SJR(2011):0.048, Scopus SNIP(2011):1.814, Scopus citations: 8, ISI Web of Knowledge Citations: 7, scHolar index: 10.

# References

- McClean SI, Barton M, Garg L, Fullerton K (2011). A Modeling framework that combine Markov models and discrete event simulation for stroke patient care. *ACM Transactions on Modelling and Computer Simulation (TOMACS)*, 21(4). Impact Factor: 0.684, Journal H-Index(SJR:1991-2009): 26, SJR(2011): 0.043, Scopus SNIP(2011):2.053, ISI Web of Knowledge Citations: 2, Scopus citations: 4, scHolar citations: 6.
- Jahangirian M, Eldabi T, Garg L, Jun T, Naseer A, Patel B, Stergioulas L, Young TP (2011). A Rapid Review Method For Extremely Large Corpora of Literature: Applications to the domains of Modelling, Simulation, and Management. *International Journal of Information Management (IJIM)*, 31(3):234-243. Impact Factor: 1.554, Journal H-Index(SJR:1986-2010): 32, SJR(2011):0.043, Scopus SNIP(2011):2.517, Scopus citations: 3, ISI Web of Knowledge Citations: 2, scHolar citations: 5.
- Gillespie J, McClean SI, Scotney B, Garg L, Barton M, Fullerton K (2011). Costing Hospital Resources for Stroke Patients using Phase-type Models. *Health Care Management Science*. 14(3):279-291. Impact Factor: 0.394, SJI(2011): 0.063, Scopus SNIP(2011): 1.169, Journal H-Index(SJR:1998-2009): 23, Scopus citations: 1, scHolar citations: 3.
- Papadapoulou A, Tsaklidis G, McClean SI, Garg L (2011). On the moments and the distribution of the cost of a semi Markov model for healthcare systems. *Methodology and Computing in Applied Probability*. 14(3):717-737. Impact Factor: 0.774, Journal H-Index(SJR:2004-2010): 11, SJR(2011):0.040, Scopus SNIP(2011):0.718.
- Garg G, Prasad G, Garg L, Coyle D (2011). Gaussian Mixture Models for Brain Activation Detection from fMRI Data, *International Journal of Bioelectromagnetism*. 13(4):255-260, scHolar citations: 2.

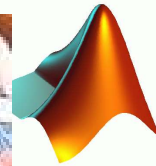
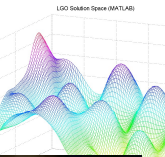
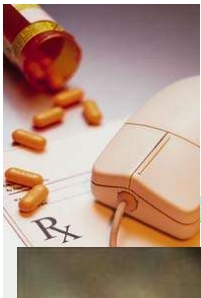
# References

- Garg L, McClean SI, Meenan BJ, Millard PH (2010). A non-homogeneous discrete time Markov model for admission scheduling and resource planning in a care system. *Health Care Management Science*. 13(2):155–169. Impact Factor: 0.394, Journal H-Index(SJR:1998-2009): 23, SJI(2011):0.063, Scopus SNIP(2011):1.169, Scopus citations: 8, ISI Web of Knowledge Citations: 5, scHolar citations: 18.
- Garg L., McClean SI, Barton M, Meenan BJ and Fullerton K (2010). The extended mixture distribution survival tree based analysis for clustering and patient pathway prognostication in a stroke care unit. *International Journal of Information Sciences and Application*. 2(4): 671-675.
- Garg L, McClean SI, Meenan BJ, Millard PH (2009). Non-homogeneous Markov Models for Sequential Pattern Mining of Healthcare Data. *IMA journal Management Mathematics*. 20(4): 327-344. Journal H-Index(SJR:2001-2009): 11, SJR(2011):0.033, Scopus SNIP(2011):0.671, Scopus citations: 1, ISI Web of Knowledge Citations: 2, scHolar citations: 4.
- Garg L, McClean SI, Barton M (2008). Is Management Science doing Enough to Improve Healthcare?. *International Journal of Human and Social Sciences* 3(2): 99-103. scHolar citations: 4.
- Garg L, McClean SI, Meenan BJ, Millard PH (2009). A phase type survival tree model for clustering patients according to their hospital length of stay. *Proceedings of the XIII International Conference on Applied Stochastic Models and Data Analysis (ASMDA 2009)*, pp. 497-502, Eds: Leonidas Sakalauskas, Christos Skiadas, Edmundas K. Zavadskas, ISBN: 978-9955-28-463-5, Publisher: Vilnius Gediminas Technical University Press. scHolar citations: 3.

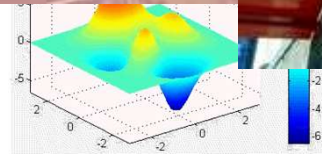
# References

- Barton M, McClean SI, Garg L, Fullerton K (2009). Modelling Stroke Patient Pathways using Survival Analysis and Simulation Modelling. Proceedings of the XIII International Conference on Applied Stochastic Models and Data Analysis (ASMDA 2009), pp. 370-373, Eds: Leonidas Sakalauskas, Christos Skiadas, Edmundas K. Zavadskas, ISBN: 978-9955-28-463-5, Publisher: Vilnius Gediminas Technical University Press. scHolar citations: 4.
- McClean SI, Garg L, Barton M, Fullerton K, Millard PH (2009). Using Markov Systems to plan Stroke Services. In Intelligent Patient Management (Eds. McClean, S.; Millard, P.; El-Darzi, E.; Nugent, C.D.), Book Series: Studies in Computational Intelligence, Vol. 189. pp. 241-256. Journal H-Index(SJR:1998-2010): 10, SJI(2011):0.029, Scopus SNIP(2011):0.310, Scopus citations: 2, ISI Web of Knowledge Citations: 4, scHolar citations: 3.
- McClean SI, Millard PH, Garg L (2008). Using Markov Models for Decision Support in Management of High Occupancy Hospital Care. In Intelligent Techniques and Tools for Novel System Architectures (P. Chountas, I. Petrounias, J. Kacprzyk Eds), Springer's book series Studies in Computational Intelligence (SCI). Vol. 109/2008. pp. 187–199. Journal H-Index(SJR:1998-2010): 10, SJI(2011):0.029, Scopus SNIP(2011):0.310.
- McClean SI, Garg L, Meenan BJ, Millard PH. (2007). Non-Homogeneous Markov Models for Performance Monitoring in Healthcare. (In C.H. Skiadas, Eds.) Recent Advances In Stochastic Modelling and Data Analysis. World Scientific Publishing, pp. 146-153. scHolar citations: 4.





PLEASE



MALTA

# Smart Sensor for EEG Acquisition and Epileptic Seizure Detection and prediction

With

Prof Justin Dauwels, Nanyang Technological University, Singapore

Prof Alok Mishra, Atilim University, Turkey

Prof K Ramesh, Karnataka State Women's University, Vijayapura, India

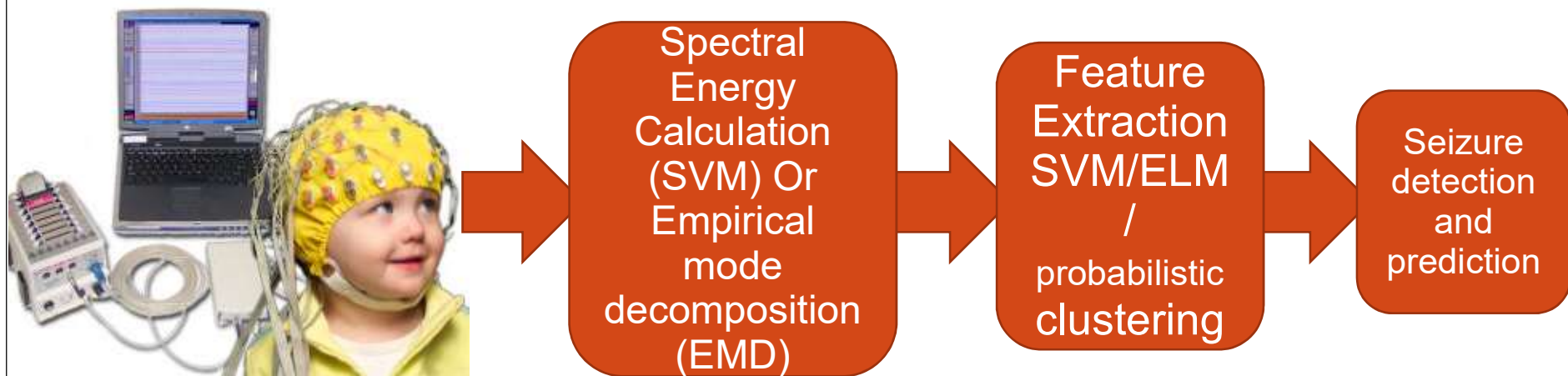
Dr Gaurav Garg, Ulster University, UK

Students:

Ms Sylvia Bugeja, Mr Eliazar Elisha Audu, Ms. Noela Galea, Ms. Yezi

Ali Kadhim, Mr Sean Bugeja, Mr. James Bonello

# Smart Sensor for EEG Acquisition and Epileptic Seizure Detection and prediction



# Smart Sensor for EEG Acquisition and Epileptic Seizure Detection

- **Collaborative partners:** Nanyang Technological University, Singapore and Massachusetts General Hospital, MIT, USA.
- **Approach:** Singular Vector Machine, Extreme learning machine, probabilistic clustering, Empirical mode decomposition.
- **Funding Body:** MNN-RIDT
- **Data:** Massachusetts General Hospital, MIT, USA.



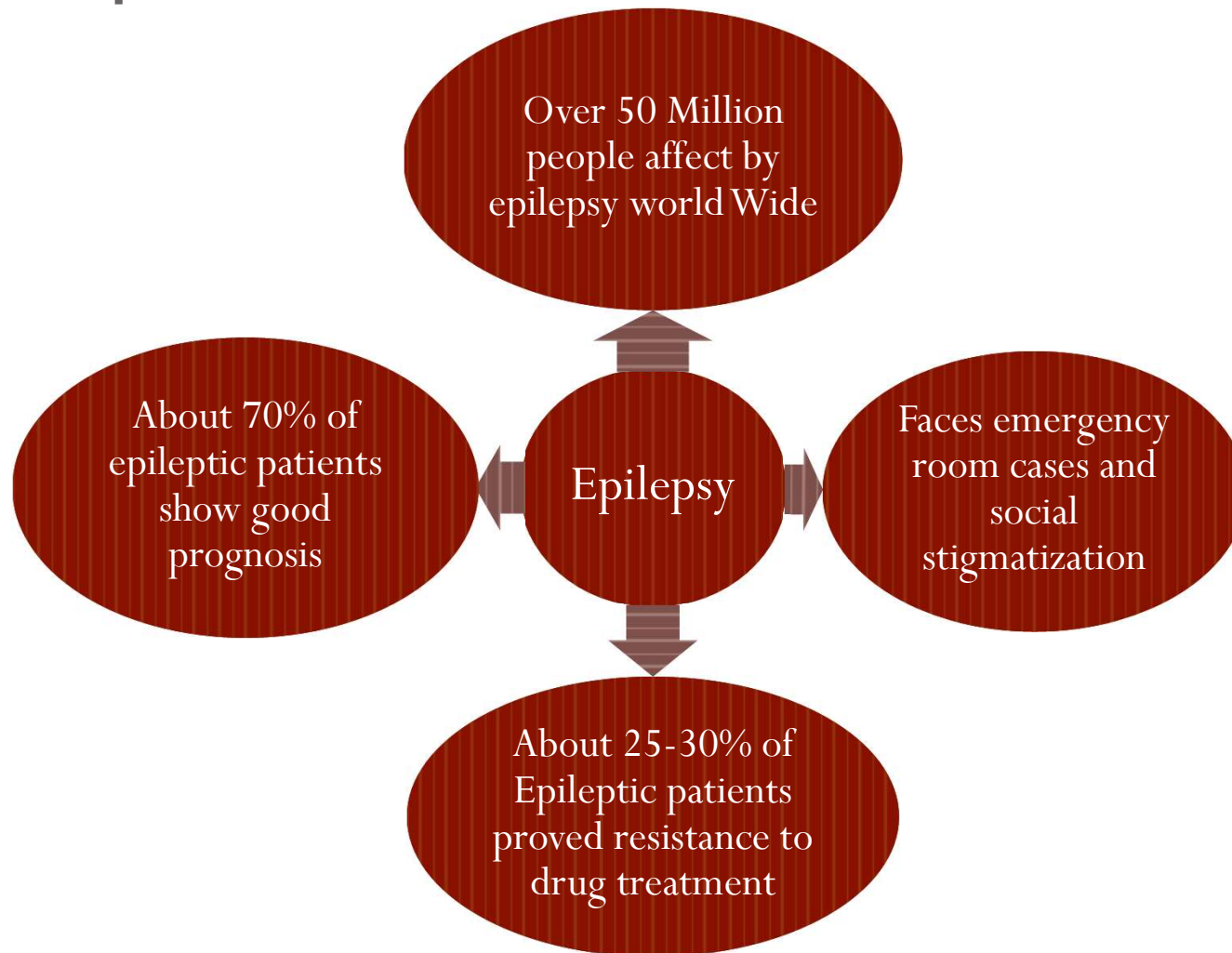
# Epilepsy

- Epilepsy is
- A medical condition
- Associated with recurrent seizures,
- Disrupt normal electrical function of the brain
- Due to excessive synchronization (hyper-synchronization) of cortical neural network.

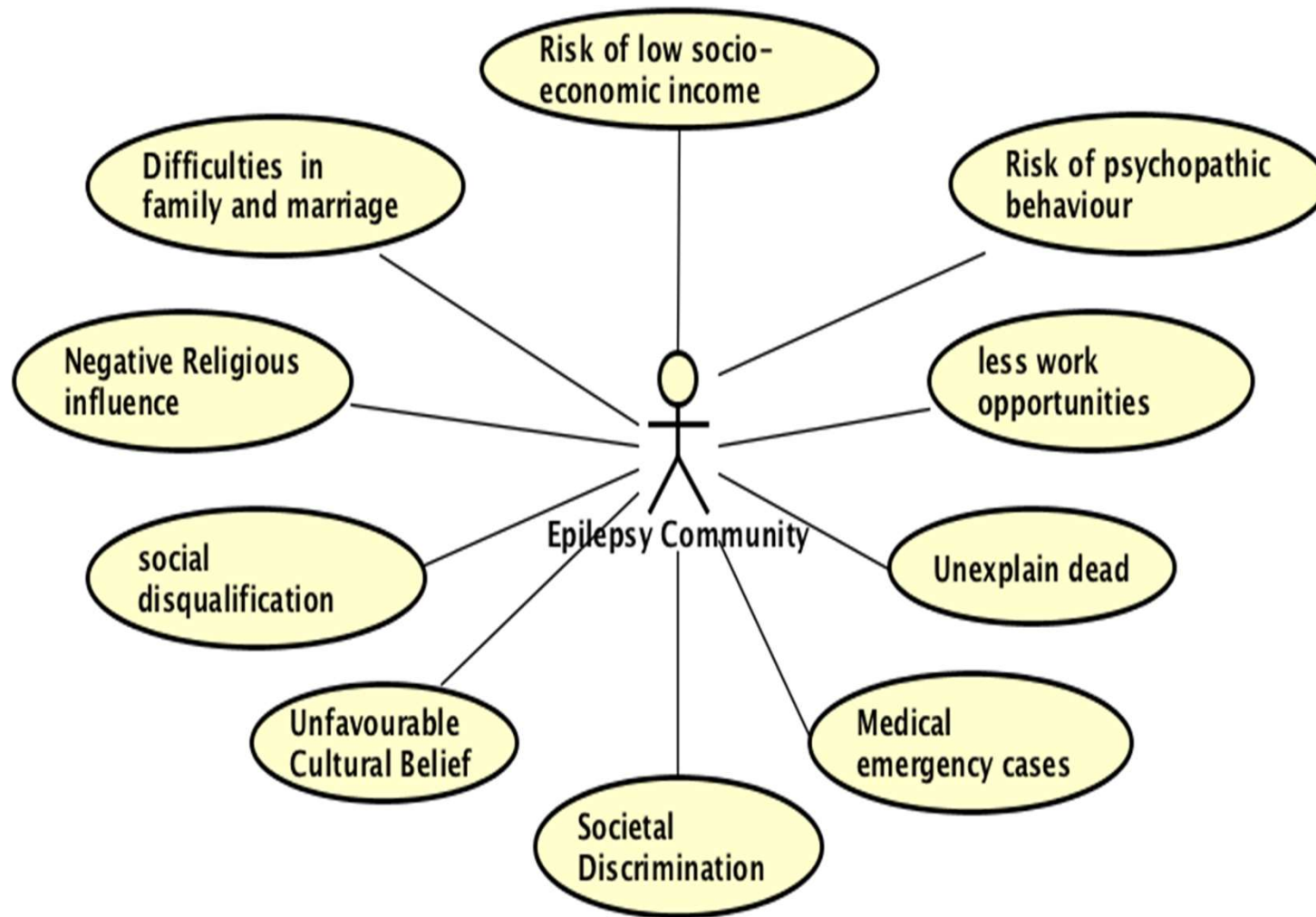
# Epilepsy

- Has profound effects on the state of consciousness, cognitive function and bodily motor control of the affected persons at the onset of seizures.

# Epileptic Seizure



# Epileptic Seizure



# Epileptic Seizure Detection and prediction

Clinical management of epilepsy:

- Through the application of signal processing and machine learning techniques.

# Epileptic Seizure Detection and prediction

Epileptic seizure detection:

- To develop systems which monitor patient EEG, learn to classify whether it is seizure or non-seizure EEG and act upon such a decision

# Epileptic Seizure Detection and prediction

Epileptic seizure prediction:

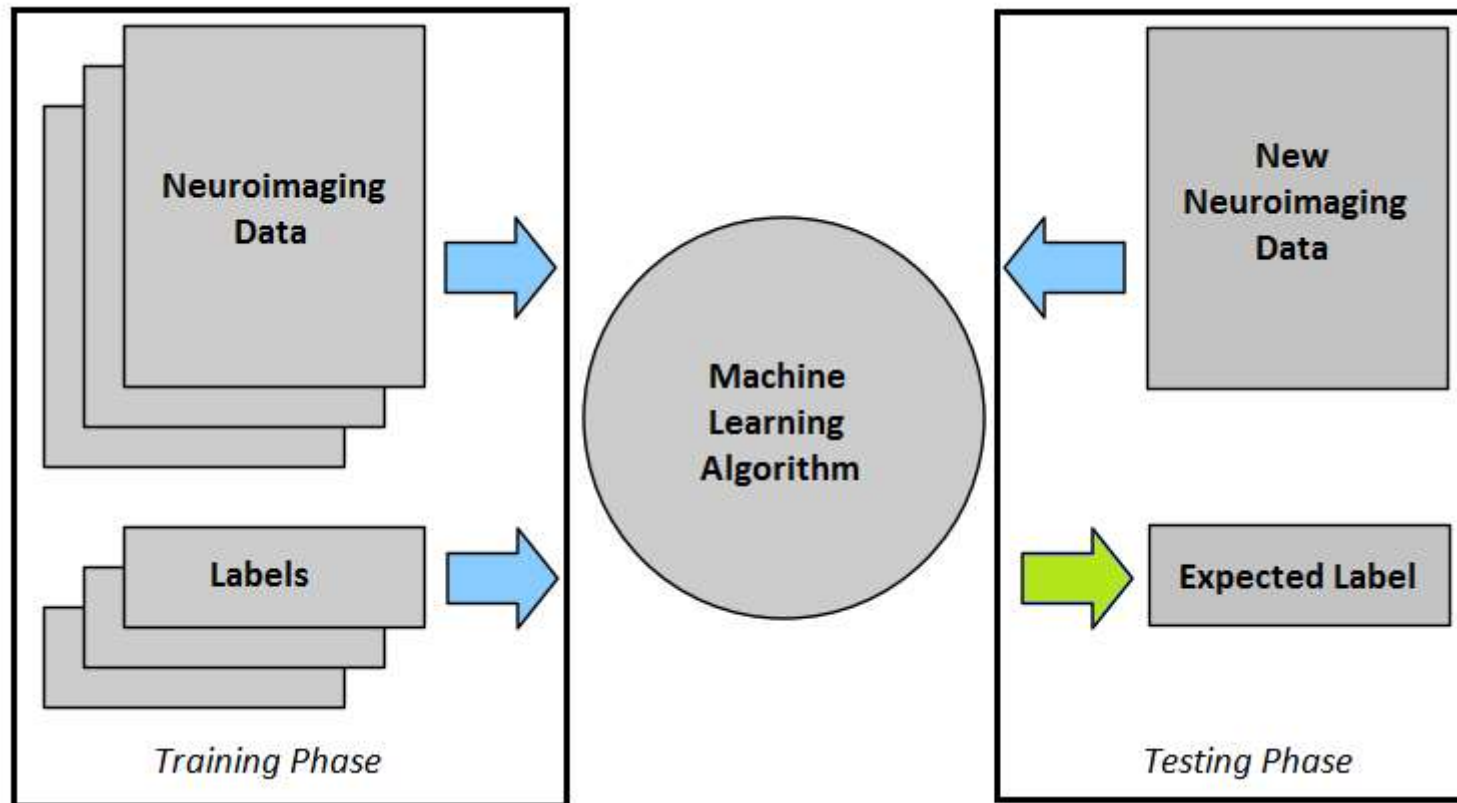
- To develop systems which monitor patient EEG, learn to predict whether the present signal is indicating the provability of occurrence of a seizure in a given time.

# Epileptic Seizure Prediction

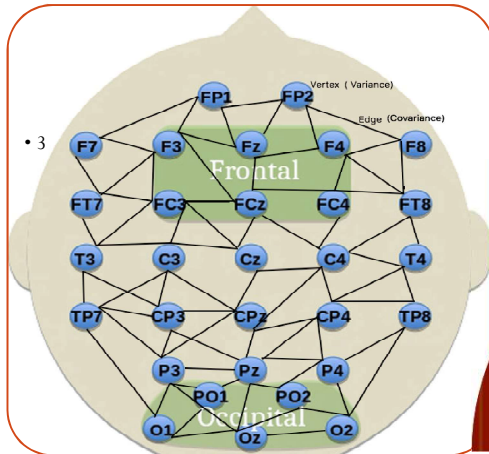
- Use signal processing and machine learning techniques
  - Extract features -> Create feature space -> Train -> Learn
- Patient-specific vs. Patient-non-specific systems
  - Patient-non-specific systems do not perform well across a large patient population, therefore not practical



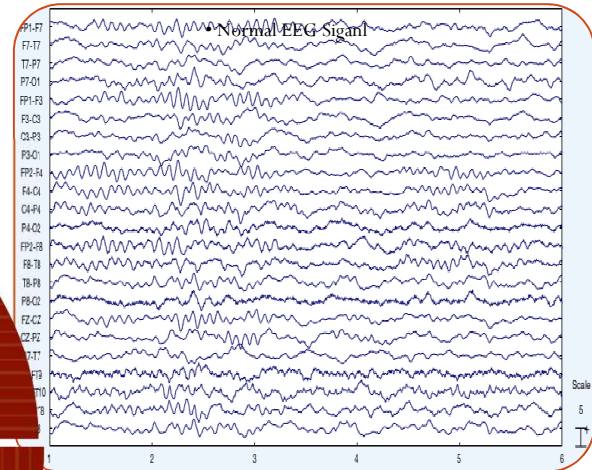
# Epileptic Seizure Prediction



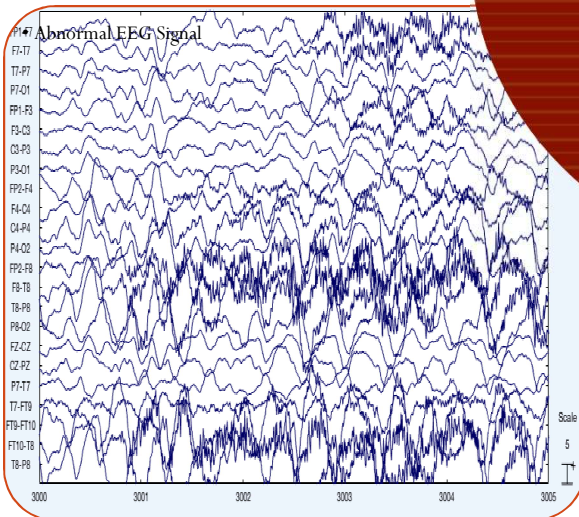
# Epileptic Seizure Prediction



**Graph Theoretic:**  
Each electrode represents vertex and the edge or link is the

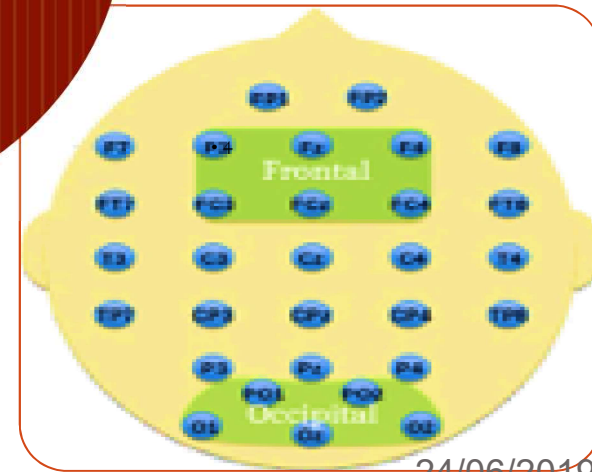


Normal EEG Signal



Abnormal EEG

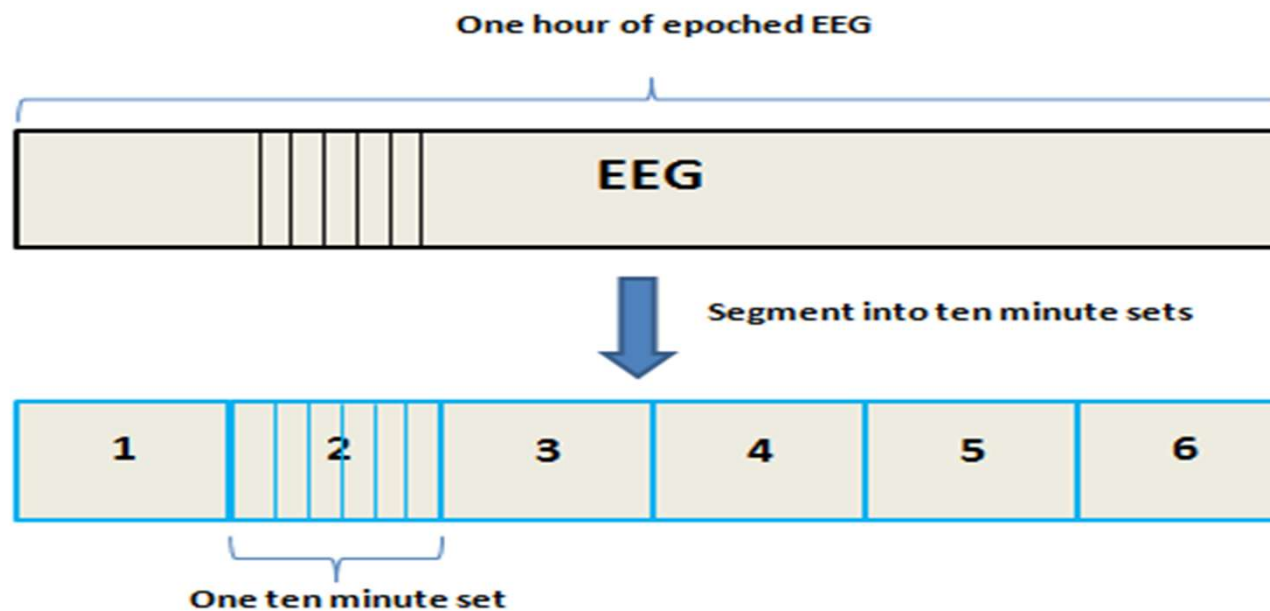
Scalp EEG



24/06/2019

# Extracting seizure and non-seizure sets

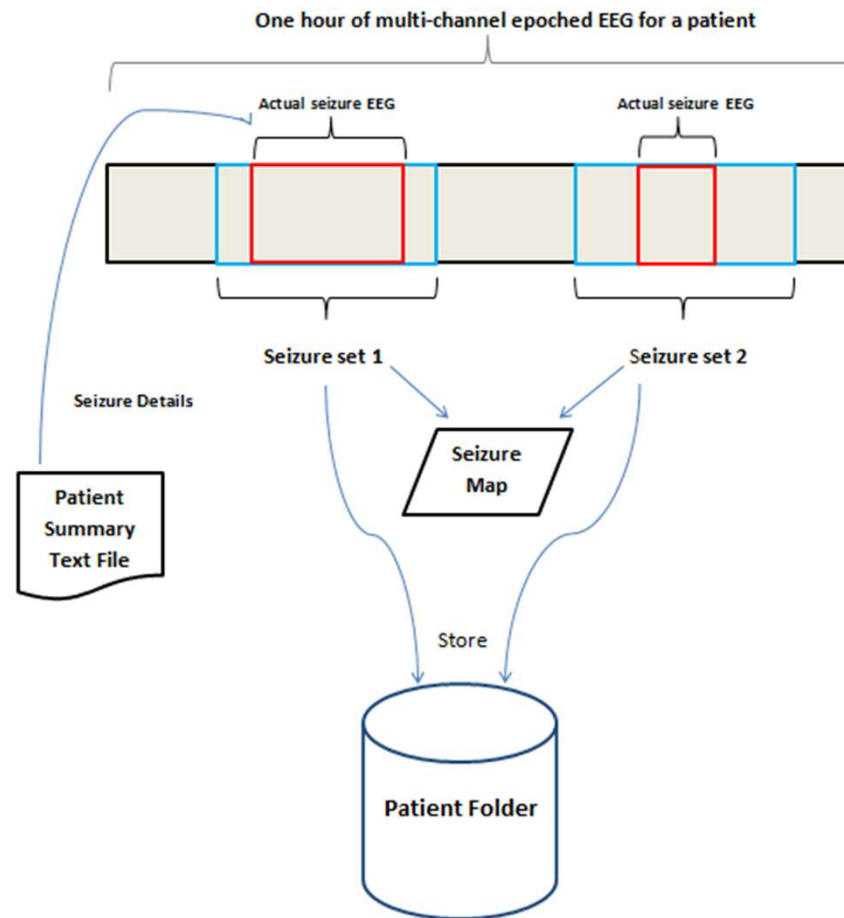
- Each hour is first segmented into 2-second epochs of EEG
- Extracting non-seizure sets is very simple
  1. Divide non-seizure hour into smaller sets of some decide equivalent length (eg. 10 minutes)
  2. Store selected sets in patient folder



# Extracting seizure and non-seizure sets

- Extracting seizure sets is less trivial
  - May have multiple seizures recorded in one hour of EEG
- Get seizure hour (For every seizure in the hour)
  1. Extract set of appropriate length (eg. 10 minutes) such that no other seizure EEG is contained within the set
  2. Store details of acquired seizure set in reference table
  3. Store seizure set in patient folder
- This method acquires smaller sets that still contain the necessary seizure EEG
- Patient folder will contain the necessary EEG which should constitute the feature space rather than the whole hours

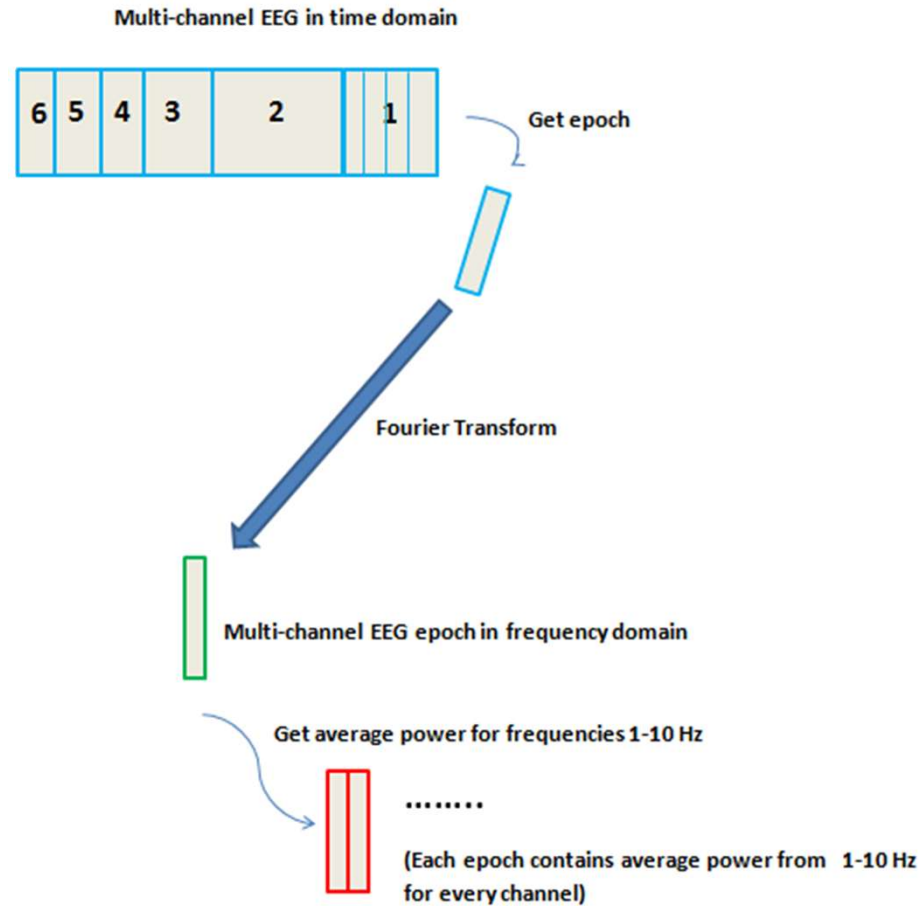
# Extracting seizure and non-seizure sets



# Extracting features

- The power is calculated from the frequency domain by squaring the amplitude
- For every EEG epoch
  - Convert from the time domain to the frequency domain
    - Uses Fourier Transform
  - Take the frequencies over the range  $\{1, 2, 3, \dots, 10\}$
  - Calculate the total power over these frequencies
  - Take the average
- Why the lower frequencies only?
  - Seizures generally act on the lower frequencies of EEG
- So the feature will be the average power over 1-10 Hz

# Extracting features



# Creating the feature space

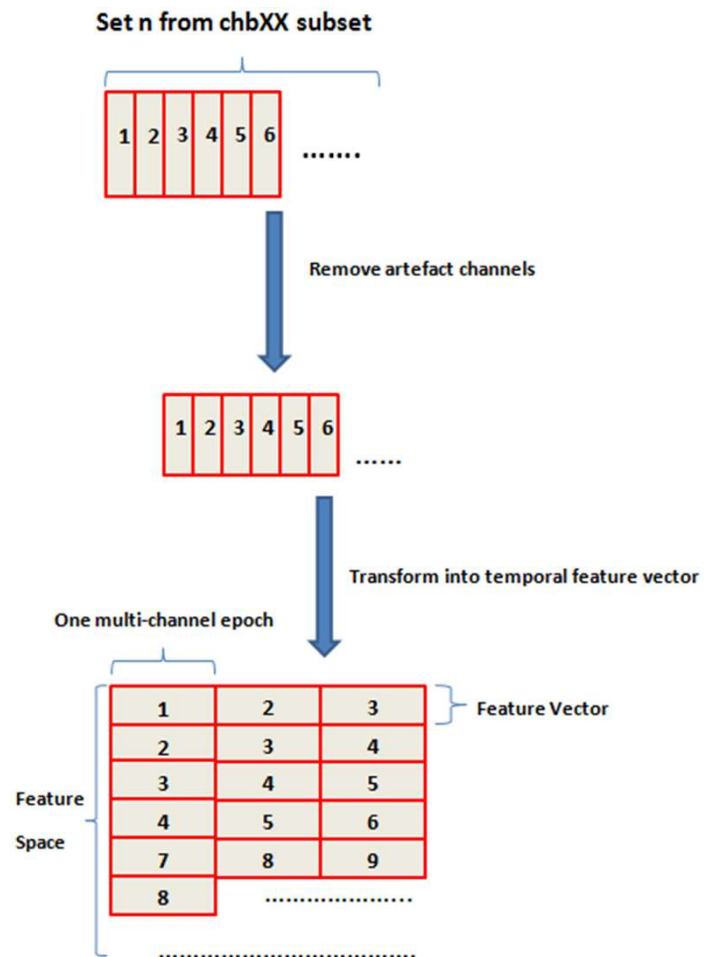
- Created by concatenating the patient folder sets
- Vary in size across different patients
  - Depending on number of channels and specified set length
- Represents temporal evolution of average power
  - 3-component feature vector containing 3 consecutive epochs
  - The “patient-specific” feature of the training set
- For every set in patient folder
  - Remove artefact channels
  - Transform into temporal representation
  - Add to feature space



# Creating the feature space

- Supervised learning approach
- Created by using seizure start and end details from the reference table
- Feature vectors that contain seizure epochs = 1
- Feature vectors that contain **no** seizure epochs = -1

# Creating the feature space



# Training

- For every patient we now have a feature space and its label vector
- SVM used for training on these structures
  - Provided by the supervisor
- Applies 3-fold cross validation

# Software and data used

- Method applied on CHB-MIT EEG datasets (as in Shoeb et. al.)
  - Pediatric EEG
- EEGLab is a toolbox plugin for MATLAB
- EEGLab used for visualisation, manipulation and processing of EEG data
- MATLAB used for building input structures and SVM training

# Results: Evaluation Criteria

- Sensitivity
  - percentage of seizure epochs correctly detected
- Latency
  - the delay between the actual start of the seizure (or seizure onset) and the time it took the classifier to react
- Number of False Positives (Selectivity)
  - number of non-seizure epochs falsely classified as seizure epochs
- **Results are prioritized by sensitivity followed by a tradeoff between latency and false positive number**

# Results

	FYP	Shoeb et. al.
Sensitivity	92.39%	96%
Latency	3.72 seconds	4.6 seconds
Selectivity	91.55%	
Total number of hours used	49.48 hours	916 hours

# Another method

- The proposed method creates a simple, yet very effective training set acquisition for epileptic seizure detection making the classifier's training phase faster.
- The proposed method was tested using CHB-MIT database, a dataset of 977 hours of EEG data containing 192 seizure instances from 22 pediatric patients collected at the Children's Hospital, Boston.

# Results

	10-Minute Subsets		20-Minute Subsets		30-Minutes Subset		Shoeb et al. <sup>[2]</sup> results
	SVM	ELM	SVM	ELM	SVM	ELM	
<b>Sensitivity(%)</b>	<b>95.33</b>	<b>99.48</b>	<b>95.42</b>	<b>99.48</b>	<b>97.98</b>	<b>98.99</b>	<b>96%</b>
<b>Specificity(%)</b>	<b>87.11</b>	<b>74.21</b>	<b>89.90</b>	<b>77.16</b>	<b>83.73</b>	<b>81.39</b>	-
<b>Latency(Seconds)</b>	<b>3.18</b>	<b>0.97</b>	<b>2.88</b>	<b>0.97</b>	<b>2.95</b>	<b>1.26</b>	<b>3</b>



# Epileptic Seizure localization

- If we can detect or predict seizure onsets by using the less number of channels (ideally only one)
- It will help us in making the seizure detection and prediction energy efficient.

# Smart Sensor for EEG Acquisition and Epileptic Seizure Detection

## More info:

- Agarwal A, Garg L, Audu EE, Pachori RB, and Dauwels J (2019) Early detection of epileptic seizures based on scalp EEG signals, In: R.S. Hegadi and K.C. Santosh (Eds.) Medical imaging: Use of AI, Image Recognition and Machine Learning Techniques, Elsevier.
- Bonello J, Garg L, Garg G, Audu EE (2018). Effective Data Acquisition for Machine Learning Algorithm in EEG Signal Processing. In Soft Computing: Theories and Applications (pp. 233-244). Springer, Singapore.

# Smart Sensor for EEG Acquisition and Epileptic Seizure Detection

## More info:

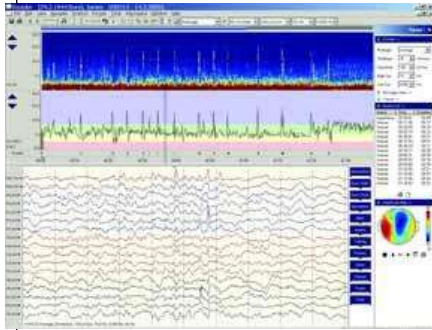
- Bugeja S, Garg L, Audu EE (2016) A novel method of EEG data acquisition, feature extraction and feature space creation for Early Detection of Epileptic Seizures, 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC-2016). Orlando, Florida USA, August 16-20, 2016.
- Bugeja S, Garg L (2015) Application of Machine Learning Techniques for the modelling of EEG data for diagnosis of Epileptic Seizures, The 3rd Workshop on Recognition and Action for Scene Understanding (REACTS 2015) Valletta, Malta, September 5 2015.

# Smart Sensor for EEG Acquisition and Epileptic Seizure Detection

## More info:

- 37. Agrawal A, Garg L, Dauwels J (2013) Application of empirical mode decomposition algorithm for epileptic seizure detection from scalp EEG, The 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'13) Osaka, Japan, 3-7 July 2013.
- Audu EE, Garg L, Falzon O, Giovanni GD (2017), Applications of machine learning in energy efficient, real-time, monitoring, prediction, detection and management of seizure: Localization of Abnormal (Seizure) EEG Source, Mediterranean Neuroscience Society – 6th Conference 2017, St Julian's Malta, June 12 – 15, 2017

# EEG and fMRI integration based models of brain disorders

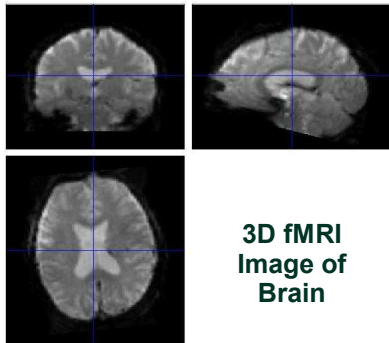


Electroencephalogram (EEG)

Pros: Fast Temporal Response

Cons: Poor Spatial Resolution

(CPP and 2-D)

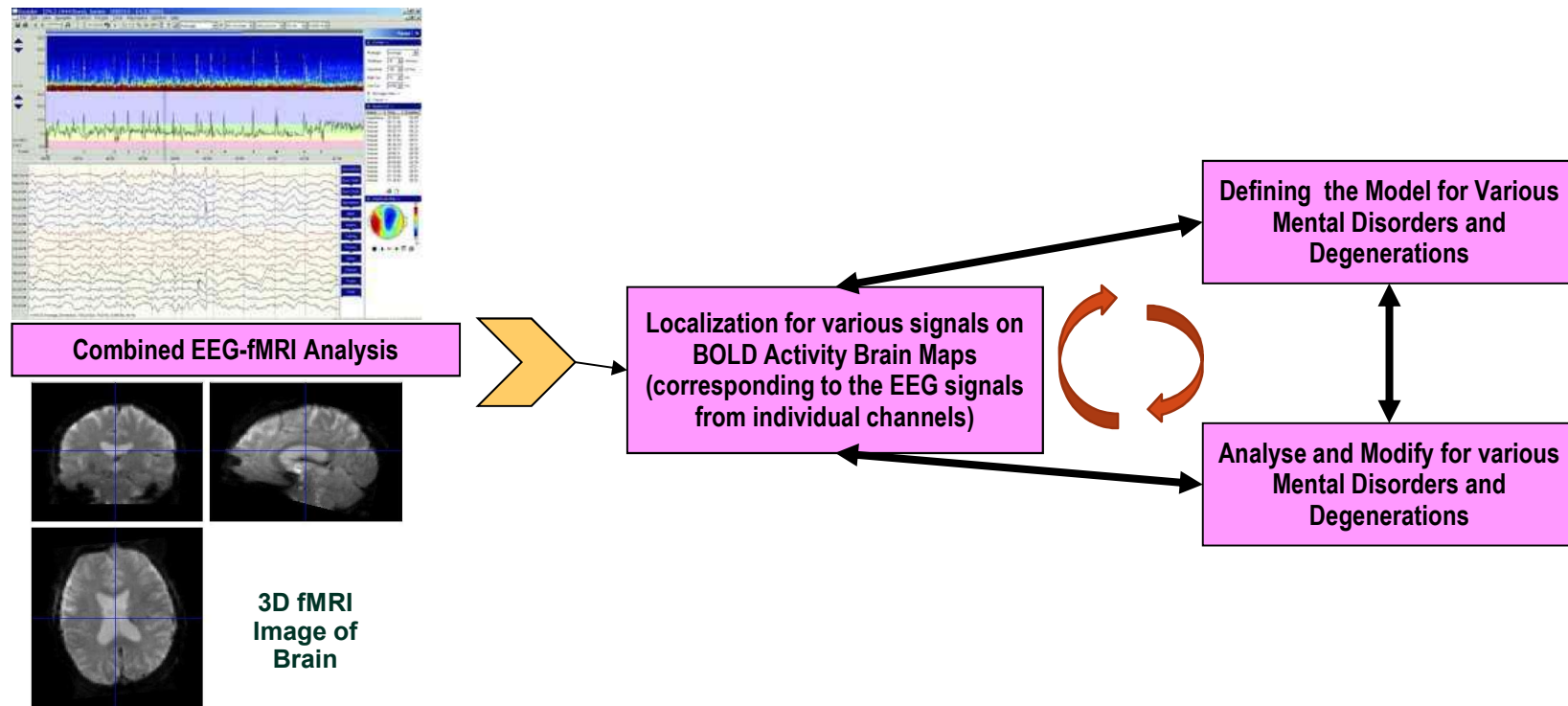


functional Magnetic Resonance Imaging (fMRI)

Pros: Good Spatial Resolution (3D)

Cons: Slow BOLD transient response

# EEG and fMRI integration based models of brain disorders



# EEG and fMRI integration based models of brain disorders

- **Collaborative partners:** Intelligent Systems Research Centre, University of Ulster, UK, Nanyang Technological University, Singapore
- **Funding:** Northern Ireland Department for Education and Learning
- **Approach:** Probabilistic clustering, cluster analysis, functional analysis, convolution, SVM, ELM, factor analysis, latent class model (LCM)

# EEG and fMRI integration based models of brain disorders

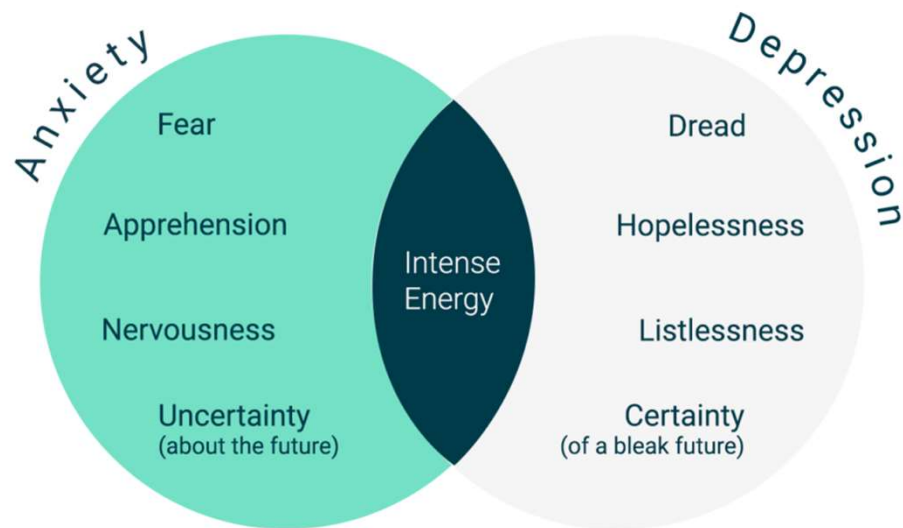
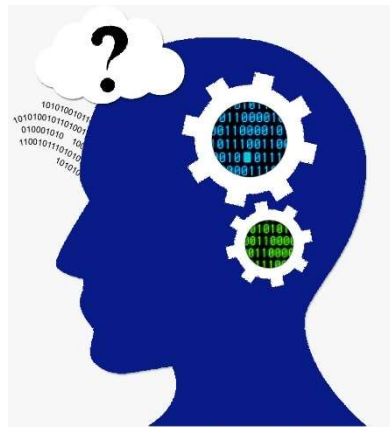
- **More info:**

- Garg G, Prasad G, Garg L, Coyle D (2011). [Gaussian Mixture Models for Brain Activation Detection from fMRI Data](#), [International Journal of Bioelectromagnetism](#). 13(4):255-260.

- Garg G, Girijesh P, Damien C (2013). [Gaussian Mixture Model-based noise reduction in resting state fMRI data](#). [Journal of neuroscience methods](#). 215(1):71-77.



# Predicting Neurological Disorder via Social Media



**15 Million** Adults will suffer the symptoms of major depression, social anxiety, or both, in any given year.

**Nearly 60%** of those diagnosed with depression have a co-occurring anxiety disorder.

# Having anxiety and depression...

Depression: Just lay in bed all day and do nothing. You're life is worthless anyway.

Me: Okay.

Anxiety: What the hell are you doing? You need to study or you'll fail all your classes, drop out of school, and end up living on the street with no friends!!

Depression: No stay here with me.

Me: ????

# Predicting Neurological Disorder via Social Media

- **Collaborative partners:** Jiwaji University Gwalior



**L-Università  
ta' Malta**

- **Approach:** CES-D screening test, Social media analytics, Major Depressive Disorder (MDD) classifier, Probabilistic clustering, cluster analysis, functional analysis, convolution, SVM, ELM, factor analysis, latent class model (LCM)

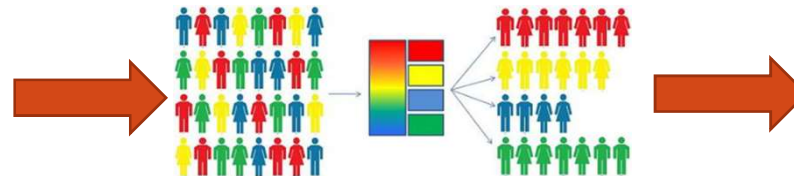
# Web-based tools for Missing data handling in medical questionnaires

**Funding body:** Nanyang Institute of Technology in Health & Medicine (NITHM), Singapore, University of Malta, Malta

Medical questionnaires with missing data

SN	Third follow-up							
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Patient 1	1	2	3	4	5	6	7	8
Patient 2	1	2	3	4	5	6	7	8
Patient 3	1	2	3	4	5	6	7	8
Patient 4	1	2	3	4	5	6	7	8
Patient 5	1	2	3	4	5	6	7	8
Patient 6	1	2	3	4	5	6	7	8
Patient 7	1	2	3	4	5	6	7	8
Patient 8	1	2	3	4	5	6	7	8
Patient 9	1	2	3	4	5	6	7	8
Patient 10	1	2	3	4	5	6	7	8
Patient 11	1	2	3	4	5	6	7	8
Patient 12	1	2	3	4	5	6	7	8
Patient 13	1	2	3	4	5	6	7	8
Patient 14	1	2	3	4	5	6	7	8
Patient 15	1	2	3	4	5	6	7	8
Patient 16	1	2	3	4	5	6	7	8

CP based collaborative filtering for missing data imputation



Completed medical questionnaires

SN	Third follow-up							
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Patient 1	1	2	3	4	5	6	7	8
Patient 2	1	2	3	4	5	6	7	8
Patient 3	1	2	3	4	5	6	7	8
Patient 4	1	2	3	4	5	6	7	8
Patient 5	1	2	3	4	5	6	7	8
Patient 6	1	2	3	4	5	6	7	8
Patient 7	1	2	3	4	5	6	7	8
Patient 8	1	2	3	4	5	6	7	8
Patient 9	1	2	3	4	5	6	7	8
Patient 10	1	2	3	4	5	6	7	8
Patient 11	1	2	3	4	5	6	7	8
Patient 12	1	2	3	4	5	6	7	8
Patient 13	1	2	3	4	5	6	7	8
Patient 14	1	2	3	4	5	6	7	8
Patient 15	1	2	3	4	5	6	7	8
Patient 16	1	2	3	4	5	6	7	8

# Web-based tools for Missing data handling in medical questionnaires

## Collaborative partners:

*Lalit Garg, Justin Dauwels<sup>1</sup>, Arul Earnest<sup>2,3</sup>, Leong Khai Pang<sup>3</sup>*



<sup>1</sup>Nanyang Technological University, Singapore



<sup>2</sup>Duke-NUS Graduate Medical School, Singapore



<sup>3</sup>Tan Tock Seng Hospital (TTSH), Singapore

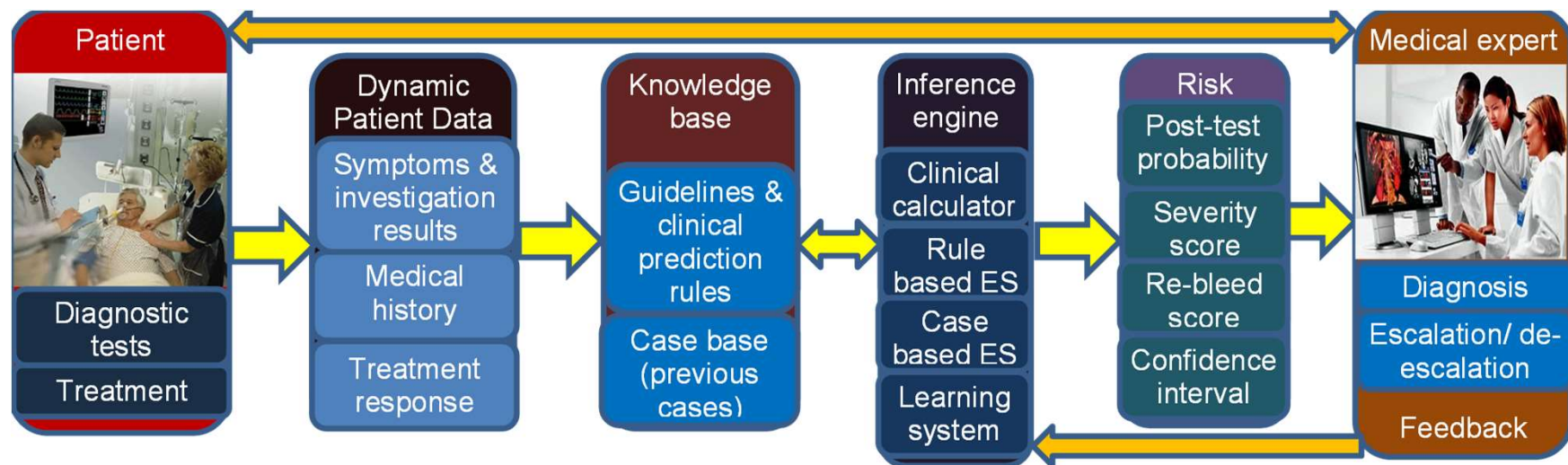
# More info...

- **Garg L**, Dauwels J, Earnest A, Pang L (2013) Tensor based methods for handling missing data in quality-of-life questionnaires. IEEE Journal of Biomedical and Health Informatics. 18(5):1571 - 1580.
- Asif MT, Srinivasan K, **Garg L**, Dauwels J, Jaillet P (2013) Low-dimensional Models for Missing Data Imputation in Road Networks, ICASSP 2013, accepted.  
[http://web.mit.edu/jaillet/www/general/missingdata\\_final.pdf](http://web.mit.edu/jaillet/www/general/missingdata_final.pdf).
- <http://lalitgarg.weebly.com/missingdatahandlingproject.html>

# More info...

- Dauwels J, **Garg L**, Earnest A, Pang LK (2012). Tensor Factorizations for Missing Data Imputation in Medical Questionnaires, The 37th International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Kyoto, Japan, March 25 - 30, 2012.
- Dauwels J, **Garg L**, Earnest A, Pang LK (2011). Handling Missing Data in Medical Questionnaires Using Tensor Decompositions. The Eighth International Conference on Information, Communications, and Signal Processing (ICICS 2011). Singapore 13-16 December, 2011.

# MDSS for managing acute upper gastrointestinal bleeding

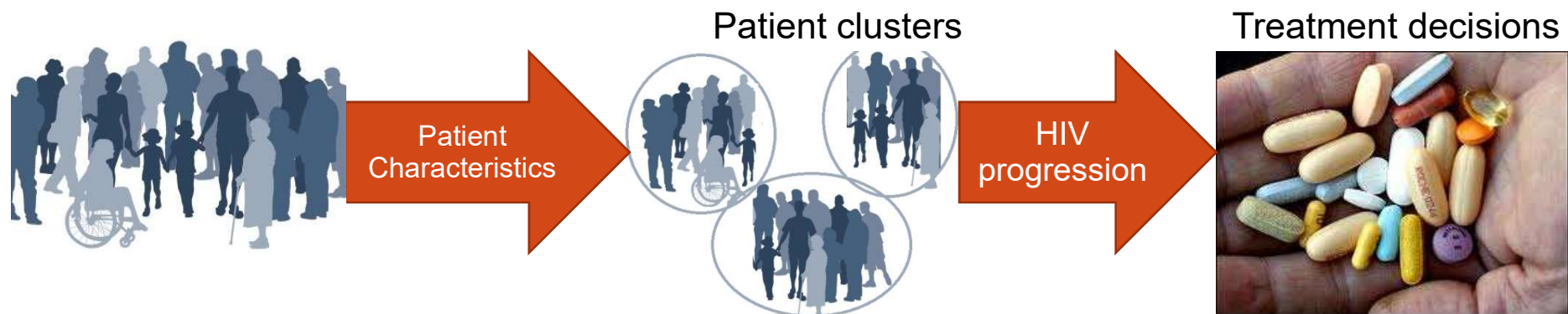




# MDSS for managing acute upper gastrointestinal bleeding

- **Collaborative partners:** Nanyang Technological University and Tan Tock Seng Hospital, Singapore.
- **Data:** Tan Tock Seng Hospital, Singapore.
- **Approach:** Pattern analysis and matching, Machine learning, rule based systems.

# HIV-disease progression modelling



# HIV-disease progression modelling

- **Collaborative partners:** University of Ulster, UK and University of Cagliari, Italy.
- **Approach:** Phase type survival tree analysis, survival analysis, Markov process model, Bayesian Analysis
- **Data:** Istituto Superiore di Sanità, Roma, Italy

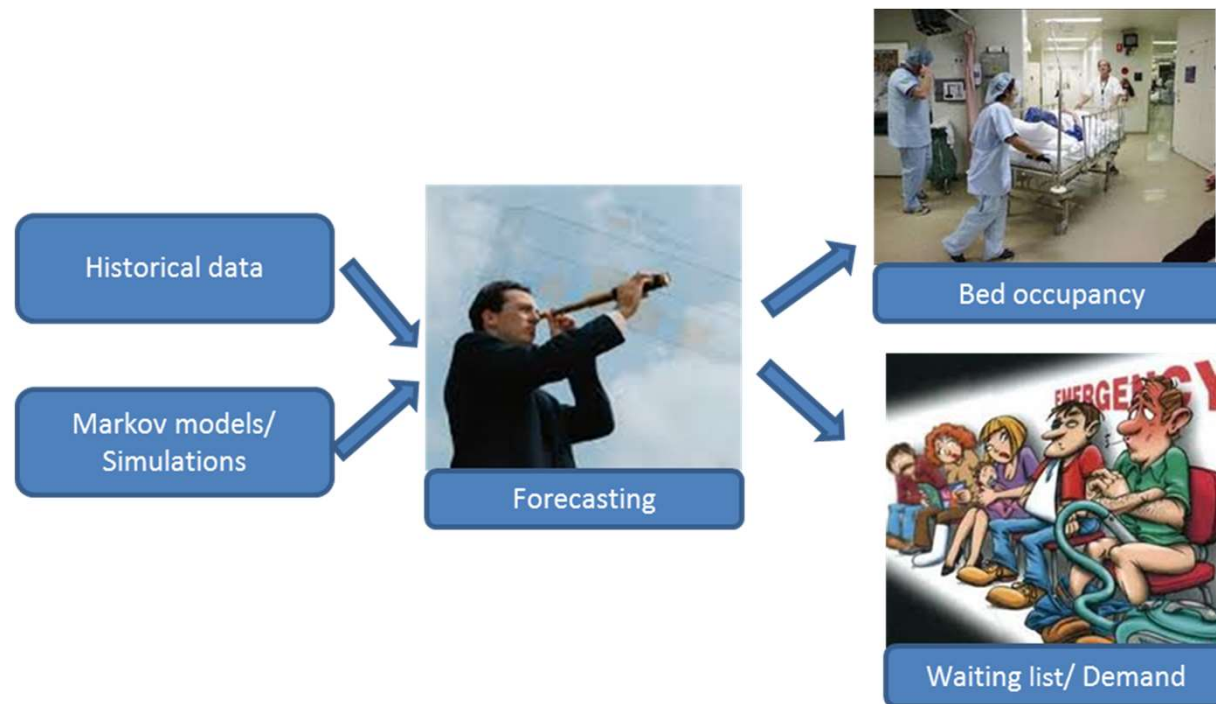
# HIV-disease progression modelling

- **More info:**

- Garg, L., Masala G., McClean S.I., Micocci M., Cannas G. (2012). Using phase type distributions for modelling HIV disease progression, Computer-Based Medical Systems (CBMS), 2012 25th International Symposium on, 20-22 June 2012. doi: 10.1109/CBMS.2012.6266408.

- Garg L, McClean SI, Meenan BJ, Millard PH (2011). Phase-type survival trees and mixed distribution survival trees for clustering patients' hospital length of stay. INFORMATICA. 22(1): 57-72.

# Hospital bed occupancy and requirements forecasting



# Hospital bed occupancy and requirements forecasting

- **Collaborative partners:** Nanyang Technological University and Tan Tock Seng Hospital, Singapore.
- **Approach:** Markov modelling, reinforcement learning
- **Data:** Tan Tock Seng Hospital, Singapore.

# Hospital bed occupancy and requirements forecasting

- **More info:**

- Garg L, McClean SI, Meenan BJ, Millard PH (2010). A non-homogeneous discrete time Markov model for admission scheduling and resource planning in a care system. *Health Care Management Science*. 13(2):155–169.

- Garg L, McClean SI, Meenan BJ, Millard PH (2009). Non-homogeneous Markov Models for Sequential Pattern Mining of Healthcare Data. *IMA journal Management Mathematics*. 20(4): 327-344.

- Garg L, McClean SI, Meenan BJ, Barton M, Fullerton K (2012). Intelligent patient management and resource planning for complex, heterogeneous and stochastic healthcare systems. In press. *IEEE Transactions on Systems, Man, and Cybernetics--Part A: Systems and Humans*.