

<THE DISSERTATION TITLE>

By

Shuming Zhou

1611144

Supervised By

Nanlin Jin

A DISSERTATION

Submitted to

Xi’an Jiaotong-Liverpool University

in partial fulfillment of the requirements

for the degree of

MASTER OF RESEARCH

<THE DATE>

ABSTRACT

xxxx

DECLARATION

I hereby certify that this dissertation constitutes my own product, that where the language of others is set forth, quotation marks so indicate, and that appropriate credit is given where I have used the language, ideas, expressions or writings of another.

I declare that the dissertation describes original work that has not previously been presented for the award of any other degree of any institution.

Signed,

Shuming Zhou

ACKNOWLEDGEMENTS

Xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx.

**TABLE OF CONTENT**

[LIST OF TABLES vi](#_Toc95215279)

[LIST OF FIGURES vii](#_Toc95215280)

[Chapter 1. Introduction 1](#_Toc95215281)

[Chapter 2. Literature Review 2](#_Toc95215282)

[Chapter 3. Research Methodology 3](#_Toc95215283)

[Chapter 4. Results 4](#_Toc95215284)

[Chapter 5. Analysis 5](#_Toc95215285)

[Chapter 6. Discussion 6](#_Toc95215286)

[Chapter 7. Conclusions 7](#_Toc95215287)

[REFRENCES CITED 8](#_Toc95215288)

[Appendix A. Xxxxx 9](#_Toc95215289)

[A.1 Xxxxxxx 9](#_Toc95215290)

LIST OF TABLES

Page

[Table 1 Principle Topics in Pattlets 11](#_Toc61607958)

[Table 2 Full Pattern-Description Organization 13](#_Toc61607959)

{To update the LIST OF TABLES, right click on the entries, chose “Update Field”, choose “Update entire table” and click on OK}

LIST OF FIGURES

Page

[Figure 1. Navigational Data-Entity Hierarchy (typical) 14](#_Toc120440498)

[Figure 2. Converse Matching of Association Types in Data-Entity Relationships 15](#_Toc120440499)

{To update the LIST OF FIGURESS, right click on the entries, chose “Update Field”, choose “Update entire table” and click on OK}

# Introduction

研究背景

1. 研究领域——新冠疫情

自从2020以来，新冠疫情已经变成了一个全球性的传染病，有非常多和此相关的研究，然而实时能够预测疫情的未来发展能够给政府带来有力的建议更好地抗击疫情。根据xxx的数据，目前感染人数，比如全球感染人数，以及中国感染人数，以及全球死亡人数，突出疫情的 严重[图片](感染人数的图片)。COVID-19来源于SARS-CoV，它能够带来哪一些症状，比如呼吸困难之类的困难。而新冠疫情带来了xxx的社会问题，比如要戴口罩，这是一个新的习惯，比如全民疫苗，比如中国的红码绿码[图片](绿码的图片)，以及远程办公的流行。疫情也有很多的变种，比如德尔塔等等。而如果能够更好地预测新冠疫情的发展，就能够更好地制定政策并从中获取有益的教训。

然而世界各国有着完全不同的疫情制定的方向，政策的改变也会受到很多因素的影响，比如中国官方的动态清零，以及英国的herd immunity的方式都会影响后续疫情的发展。

因此本文的主题就是如何正确xxxxxx

1. 发展现状

那么对于新冠疫情的预测前期有xxx之类的工作，但是这些工作主要是对xxx的研究，采用的是xxx的机器学习的算法，

1. 研究对象

对象主要是新冠疫情感染人数，其中有政策和病毒两个主体，分为以下几个属性：

1. 病毒感染人数
2. 政策
3. 经济底线
4. 改变政策的具体模式

这四个是主要的属性，当然还有其他的属性也是研究的部分，但是并没有以上四个属性重要，故放置于后面的文章叙述

研究进展，指出存在的问题，已知-》未知

目前的研究虽然可以实现对传染人数的预测，但是其框架是存在问题的，并没有将经济代价考量纳入考核。比如上述提到的xxx机器学习，的参数并没有xxx

以及上述提到的Oxford的xxx，也并没有使用xxx

所以他们的研究有所欠缺，但是我的研究提供了一个新思路，我的研究目的为什么是有意义的

s the problem stated both in a general and in a specific way? Are the purpose and rationale of the project stated clearly? Are the advantages of the proposed work justified? Are the project aims and objectives clearly stated?

# Literature Review

Introduction of topic related research

研究原因：COVID-19对世界影响巨大，各国都收到了冲击, 对于环境，社会（<https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3644567>），经济(https://www.neaman.org.il/EN/Files/Global%20Economic%20Impact%20of%20COVID-19\_20200322163553.399.pdf)，教育(https://www.scirp.org/journal/paperinformation.aspx?paperid=103646)各个方面都有巨大影响。

不仅仅是国家，对个人而言也有着巨大的影响，不仅仅是生理的（https://www.sciencedirect.com/science/article/pii/S1936657420301461）也有心理（https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7165115/）的。因此这份研究的主题就是对于COVID-19而展开的。

对新冠疫情的类型分三大主体：对新冠疫情下的人民的研究(和上面相同)，对新冠疫情本体的研究，比如病毒本身的研究(https://www.sciencedirect.com/science/article/pii/S1567134820302537)医疗方面(https://www.sciencedirect.com/science/article/pii/S1521661620303181)以及对新冠疫情发展的预测(file:///Users/mac/Documents/GitHub/SAT405\_program/exercise/Prediction\_COVID\_19.pdf)，以及对政府的研究：政策的研究(https://www.nature.com/articles/s41562-020-0909-7)。（这三个）

对人民的研究主要是对疫情的心理和生理的评估，和本研究关系不大，而对病毒本身的研究以及医疗方面的研究涉及具体的医学和生物学知识，不在本文的探讨范围之内，其次就是对疫情的未来发展的趋势，这是对本文极其相关的点，具体也会在下文探讨。

对政府的研究主要是对NPI的研究，（https://www.sciencedirect.com/science/article/pii/S0048969721005982）NPI代表了具体的政策实施。当然也有对政策经济代价的研究，比如(https://www.medrxiv.org/content/10.1101/2020.03.26.20044552.abstract)，提到了中英两种模式下的经济代价。

再具体一些，对于这些研究，很多都是对单个因素的研究，比如单纯探讨疫情下人的心理生理状况，以及单独对npi的研究，其中也有很多是两个要素综合起来研究的，比如oxford的论文(file:///Users/mac/Documents/GitHub/SAT405\_program/405%E6%96%87%E7%AB%A0.pdf)，探讨病毒传播率以及NPI的政府政策，以及

<https://www.jmir.org/2020/9/e21419/>

这个文章主要是探讨政策和人民的态度。

当然最和本研究主题相关的，是政府政策以及病毒传染率这两个要素。对个人态度选择而言并不在本topic的论述范围之内。

通过对这两个要素的研究，我们可以获得政策和传染率之间的模拟过程，并预测将来政策的走向和疫情传染率的发展。这对于公共安全是有重大意义的，也能够为未来新的传染病提供研究方向，比如猴痘。

而对于中英两国不同的抗疫政策的研究也是具有重要意义的，各个国家之间互相借鉴抗疫的模式能够互相进步并获得对政策影响的更加深刻的认识，对全球疫情转好具有帮助

Theme2

目前对于疫情发展的趋势主要使用各种算法，对以下参数进行模拟。

首先是province/state, country/region, last update, confirmed, death and recovered cases的参数作为输入，应用support vector regression (SVR) [16] and polynomial regression (PR)的机器学习模型（file:///Users/mac/Documents/GitHub/SAT405\_program/exercise/Prediction\_COVID\_19.pdf）。

其次是Oxford的GA，对NPI，和病毒变异率以及传染率进行模拟（file:///Users/mac/Documents/GitHub/SAT405\_program/405%E6%96%87%E7%AB%A0.pdf）

以及<https://docserver.ingentaconnect.com/deliver/connect/tsp/15261492/v125n2/s17.pdf?expires=1665805258&id=0000&titleid=75011046&checksum=DF9ED9D5B75EBFC8D20205E39276A90D&host=https://www.ingentaconnect.comc>的采用启发式算法，主要应用高斯模型预测死亡率。

Theme3

但是经过上述分析发现以上的预测算法的参数是欠缺的，并未考虑政策的经济因素，根据上述提到的（file:///Users/mac/Documents/GitHub/SAT405\_program/exercise/2020.03.26.20044552v1.full.pdf）政策会导致经济代价的不同，而根据（https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19#C）事实，英国调整了政策用于开放经济，这说明政策不仅仅影响传染率，也会被经济代价所影响，而中英两国的经济政策的改变也有所不同（file:///Users/mac/Documents/GitHub/SAT405\_program/exercise/2020.03.26.20044552v1.full.pdf），因此，对未来政策变化的模拟和未来感染人数的确定需要结合当前政策，具体的模式，以及经济代价。

因此这里的研究主要是采取模拟的形式，运用类似GA的模型来体现政府政策的变化以及相应的经济代价和病毒传染率如何具体影响政策的变化。根据上述的讨论，这个方向具有研究的价值。

而对于猴痘病毒的预测，目前还是集中在神经网络来学习，缺少启发式算法的研究，这个研究能够给预测猴痘病毒的人群一点参考意义。<https://www.ejgm.co.uk/download/analyses-of-polynomial-neural-networks-for-prediction-of-the-prevalence-of-monkeypox-infections-in-12400.pdf>

# Research Methodology

具体会采用两种方式，一种是定性的方式进行模拟，另外一种是依靠定量的方式进行计算，下面将分别对于这两种方式进行详细的叙述。

## 定性研究

### 数据

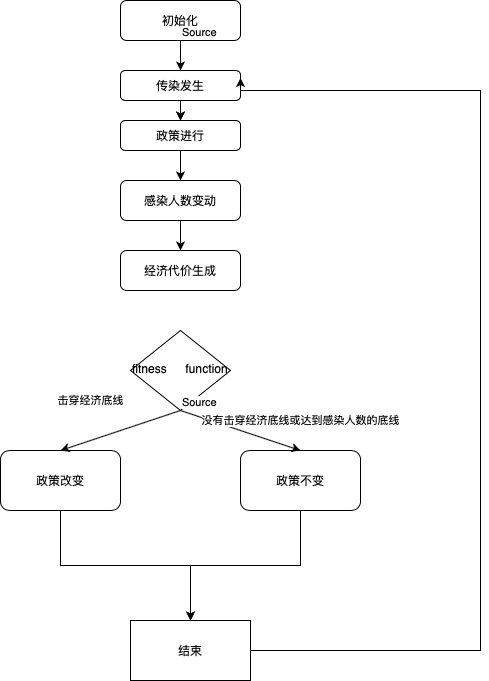
由于是定性研究，故而主要在于其趋势的发展是否和现实中的疫情发展趋势大致相同。主要有以下几个参数和指标：

|  |  |
| --- | --- |
| 每日新增感染人数 | 每日新增被感染的人数 |
| 当前政策 | 当前政府选择的政策 |
| 经济底线 | 每一项政策都会带来相应的经济代价，而政府会在经济代价过大的时候衡量是否要取消这一项政策 |
| 政策变化 | 当前政策会因为各种因素而发生改变，政策变化的fitness函数用来描述这种变化 |
| 被隔离人数 | 每日新增人数有一部分会被隔离 |

### 算法

首先每日会根据传染率由已经感染的人群传递给没有被感染的人群，政府的当前政策会对于当前的感染人数进行更新，而政府会根据本次政策带来的经济代价以及感染人数的变化进行政策的调整于是进入下一次的循环。

具体的流程图以下



## 定量研究

### 数据

#### 政策表的数据

政策的数据主要来自xxx，其中的数据结构太多了，可以详情见Appendix，具体我们需要的列有以下：

|  |  |
| --- | --- |
| type | 政策的名称 |
| |  | | --- | | date\_start | | 政策实际实施的时间 |
| |  | | --- | | date\_end | | 政策实际结束的时间 |

以下是挑选出的各个政策

|  |  |
| --- | --- |
| newFirstEpisodesBySpecimenDate | 第一次被感染的人数 |
| cumFirstEpisodesBySpecimenDate | 累计到今天的首次感染 |
| New Task Force, Bureau or Administrative Configuration | 强制措施 |
| Anti-Disinformation Measures | 反造谣 |
| Closure and Regulation of Schools | 关闭学校 |
| COVID-19 Vaccines | 疫苗 |
| Curfew | 宵禁 |
| Declaration of Emergency | 紧急宣告 |
| External Border Restrictions | 外部限制 |
| Health Monitoring | 健康监测 |
| Health Resources | 医疗资源 |
| Health Testing | 健康检测 |
| Hygiene | 卫生 |
| Internal Border Restrictions | 内部边界封锁 |
| Lockdown | 关闭 |
| Other Policy Not Listed Above | 其他政策 |
| Public Awareness Measures | 公众意识 |
| Quarantine | 隔离 |
| Restriction and Regulation of Businesses | 商业限制 |
| Restriction and Regulation of Government Services | 政府规定 |
| Restrictions of Mass Gatherings | 聚集限制 |
| Social Distancing | 社交距离 |

#### 每日感染人数数据

主要来自xxx，但是其中的列过多，这里仅仅挑出最重要的部分：

|  |  |
| --- | --- |
| date | 当天日期 |
| new\_cases | 当天新增感染 |
| total\_deaths | 总死亡 |
| new\_deaths | 新死亡 |
| reproduction\_rate | 传播R值 |

### 数据预处理

政策和cases两张表需要合并，作为输入提供给机器学习算法所以以下是具体如何合并的过程。

具体的做法是先将政策那一张表中所有的policy全部提取出来，然后作为列添加到cases的列中。

其次根据policy的时间将对应的cases所在时间的政策置为1，这代表在这段时间中采取了这个政策。

### 算法简述

主要适用python库中的sklearn来进行线性的预测，主要使用LinearRegression,SGDRegressor这两个。

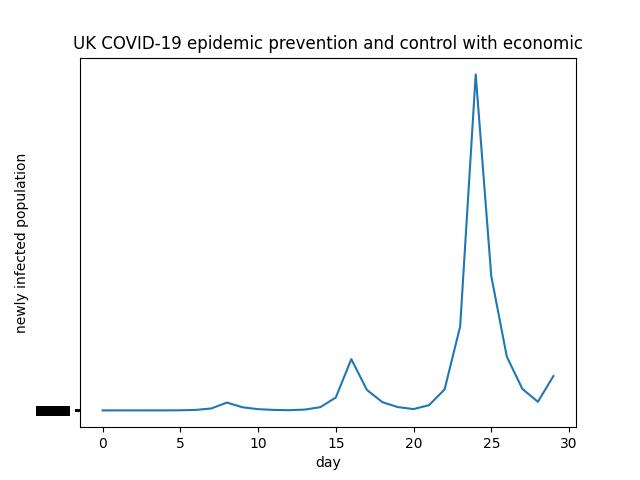
The regression approaches are trained, tested and used for prediction on real data using the the mentioned column.

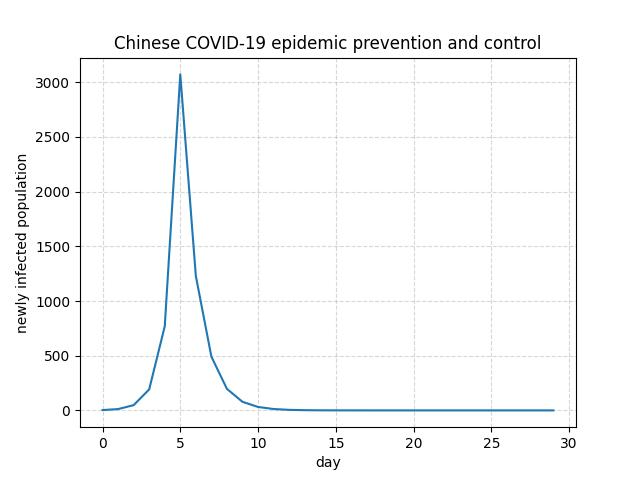
真实的数据将会被切分为训练集和测试集，均使用mean\_squared\_error和R2来测试训练集的准确度并会输出图像，来查看是否和真实数据匹配作为和上述模拟过程的支撑。

# Results

定性模拟

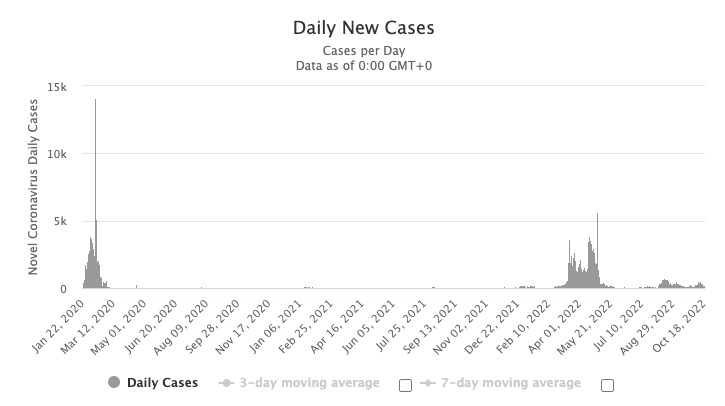
通过上述定性的模拟，得到结果如下



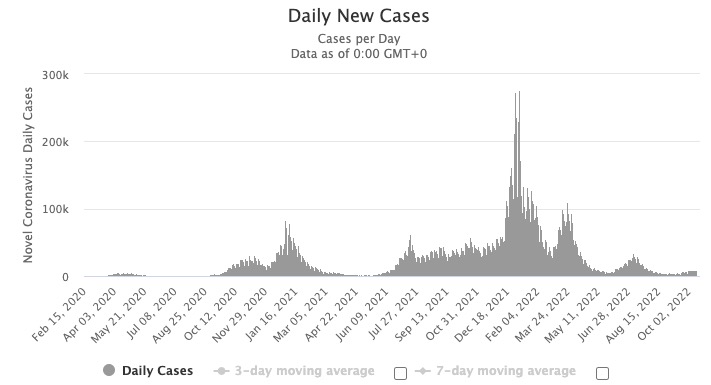


以下是现实的趋势

Chinese

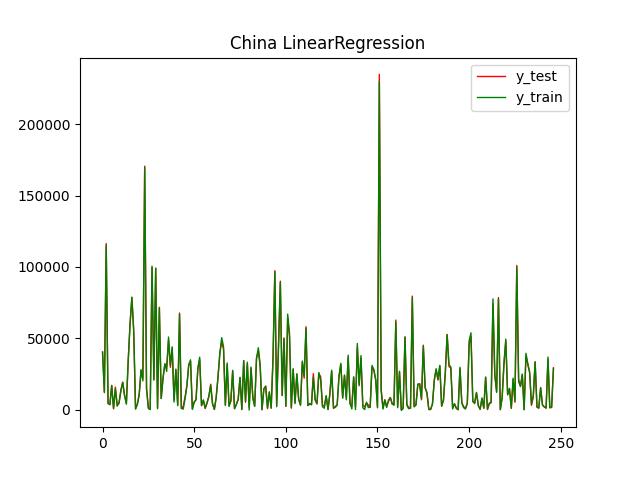


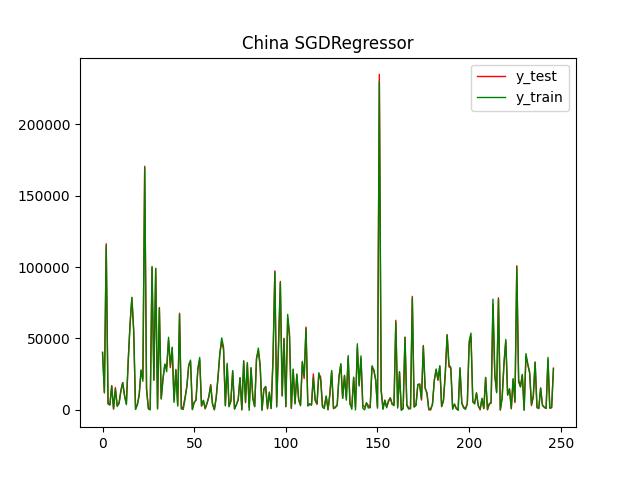
UK

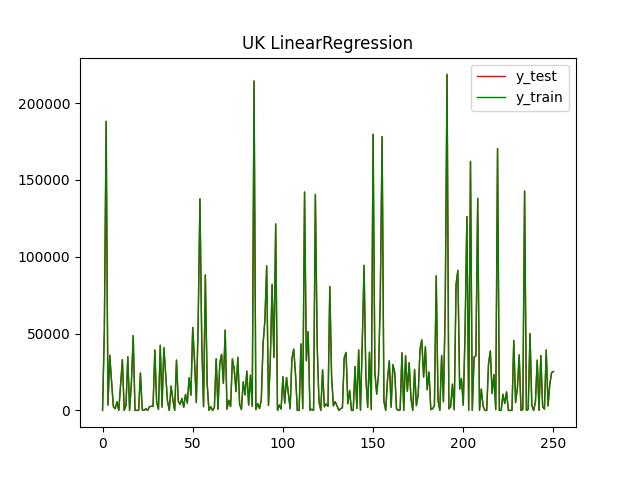


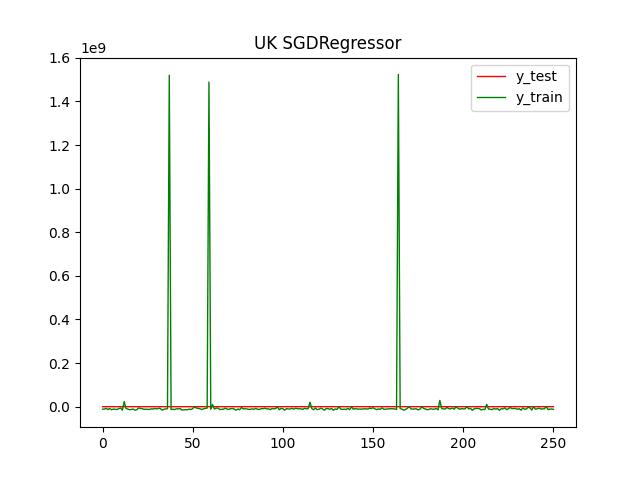
定量计算

定量计算得到的各项数据如下









很明显地可以看出，这其中SG明显不如普通的LinearRegression，而其中的其他的值对比如下

英国

|  |  |
| --- | --- |
| LinearRegression |  |
| mean\_squared\_error | 6.956805139013924e-20 |
| r2\_score | 1.0 |

|  |  |
| --- | --- |
| SGDRegressor |  |
| mean\_squared\_error | 25739583591725.77 |
| r2\_score | -2013312.1793092396 |

中国

|  |  |
| --- | --- |
| LinearRegression |  |
| mean\_squared\_error | 4.22123953376573e-24 |
| r2\_score | 1.0 |

|  |  |
| --- | --- |
| SGDRegressor |  |
| mean\_squared\_error | 7097213485448.217 |
| r2\_score | -555132.8231911982 |

# Analysis

Xxxxx

Xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx.

# Discussion

Xxxxx

Xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx.

# Conclusions

Xxxxx

Xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx.

REFRENCES CITED

[1] xxx

APPENDICES

###### Xxxxx

Xxxxxxx

Xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx.