Institution: Queen's University Belfast		
Unit of Assessment: 10		
Title of case study: Data analytics driven modelling for healthcare management & government policy		
Period when the underpinning research was undertaken: 2004-2020		
Details of staff conducting the underpinning research from the submitting unit:		
Name(s):	**Role(s) (e.g. job title):**	**Period(s) employed by submitting HEI:**
Adele Marshall	**Professor, School of Mathematics and Physics**	**2001- current**
Period when the claimed impact occurred: 2019-2020		
Is this case study continued from a case study submitted in 2014? N		

1. Summary of the impact:

Effective healthcare provision relies on understanding how factors such as illness and age underpin a patient's progress through the healthcare system. A new suite of statistical models developed in Queen's University Belfast addresses this major challenge through novel front-end integration of data analytics. These models have provided government officials and healthcare managers the tools they needed to set policy and allocate resources in both normal and emergency circumstances:

(i) **Covid19 Alert level setting** (*Government policy*): Data analytics informed local and national coronavirus alert level setting.

(ii) **Staff deployment** (*Healthcare Management*)**:** Model-driven changes for shift times in Ulster Hospital's Emergency Department (ED) resulting in a 14% reduction in waiting times to under 60 minutes for an additional 10,000 patients in one year.

2. Underpinning research:

Fully understanding the origin of delays in triage and time-critical treatment for patients has been the major issue faced by healthcare systems worldwide for decades. For example, the problem of seasonal changes in NHS patient waiting times leads to regular national news headlines during the winter months. The origin of this is twofold:

1) **Extrinsic** – gaps in our knowledge of the factors driving prevalence/transmission of disease

2) **Intrinsic** – improper understanding of causes of delays in patient transit through hospitals

In 2020, this problem was thrust to the forefront of the public agenda. Inefficiencies and lack of reserve capacity have required the implementation of lockdowns and major disruptions to daily life to protect the NHS from surges in hospitalisations. These decisions have been informed by modelling disease transmission to best manage patients entering the healthcare system. Prof. Adele Marshall, working in the Mathematical Sciences Research Centre (MSRC) at Queen's University Belfast (QUB), has focused heavily over the last decade on implementing cutting edge data analytics at the front end of these models to increase their overall efficiency and efficacy.

Extrinsic: Social Network Analysis

Modelling disease transmission in a mobile population [R1]

Understanding how movement/travel can lead to the spread of diseases is critical for understanding how well healthcare systems can cope with evolving disease progression in a dynamic population. To date it has been difficult to test models on human populations as until very recently humans were not ubiquitously tested and geo-tagged for highly infectious disease. Livestock populations, however, provide an excellent platform upon which to develop and test the efficacy of models to predict disease progression. As they are continually tested and traced (*i.e.* farm to fork), the spread of disease in livestock herds is well documented. This allows models to be developed to identify the major drivers of this transmission.

Social Network Analysis (SNA) is ideally suited to interrogate such dynamic systems in a quantitative manner. By creating a map of the interactions that establish a network, SNA identifies the components that support and influence behaviours on this network. Key individuals and/or groups represent the nodes of the network. These nodes are connected by links that can be interpreted as social connections between the nodes. For example, in humans these are friendship, marital, family ties etc. SNA studies how the network structure influences such links.

In this context, Brown *et al.* **[R1]** have demonstrated the successful application of SNA to understanding the key motivators responsible for the spread of bovine Tuberculosis (bTB) in cattle in NI. By tracking cattle movements between different farms and geographical regions in NI, and pinpointing nodes with high betweenness centrality in such networks, this analysis has identified influential herds and markets as being extremely important to herd connectivity. This approach enabled the creation of the first epidemic meta-population model of bTB for NI, which builds upon a simple within-herd model using cattle movements (stratified by geographical area) and wildlife interactions. The use of geographically stratified movements and wildlife interactions, and the consideration of NI as a meta-population have increased the knowledge and understanding of bTB in NI. This model forms one of the *"basis components"* of the NI Covid19 Model (Lead Data Scientist, Strategic Investment board NI writing in **[S2]**).

In **[R1]**, the efficacy of interventions such as "*movement restrictions"* (or 'lockdowns' in the Covid19 vernacular) and the threat posed by "*superspreader*" events and how best to deal with them *i.e. "node removal"* (or 'self-isolation') have been quantitatively characterised. That these terms have entered the lexicon in 2020 highlights the prescience, and fundamental importance, of this work to society in general. To clarify, while it could be argued that this emergence is somewhat phenomenological, the research itself is based on long-established problems faced by our society.

Intrinsic: Coxian phase-type distributions models

Patient flow through healthcare systems [R2-R5]

For the last two decades, researchers at QUB have been spearheading the quest to find ways to interrogate the enormous dynamic data sets that healthcare systems produce. The aim has been to extract common identifiers across distinct groups that can directly inform how to manage their efficient transport through the healthcare system while simultaneously optimising the deployment of finite healthcare resources.

A phase-type probability distribution results from a system of inter-related processes occurring in sequence (phases). A sub-set of these distributions (Coxian phase-type) describe 'duration until an event occurs' in terms of a process consisting of a sequence of latent phases. This sequence ends with an absorbing state (a final phase from which there are no further transitions). Here, transitions through these phases correspond to the stages in patient treatment *i.e.* diagnosis, assessment, rehabilitation and long-stay care, eventually reaching the absorbing state (corresponding to departure from hospital) either through discharge, transfer or death.

Coxian phase-type distributions have been successfully used by Marshall *et al.* to suitably represent the skewed stochastic nature of patient duration of stay in hospital **[R2]**. The use of such models in complex dynamically changing healthcare settings inspired the development of the discrete conditional phase-type model which facilitates the integration of a data analytics approach at the front-end of the model. This acts as a mechanism for identifying and incorporating key influential patient characteristics which are normally too computationally difficult to include **[R3]**. Such discrete conditional phase-type approach models the survival time or length of hospitalisation. Moreover, it has also been applied to represent the patient activity in Emergency Departments (EDs) and has exposed key variables influencing waiting time and bottlenecks in the system. When the patient characteristics are only a few, the Coxian Phase-type regression models have provided an understanding of the relationship between patient attributes, overcrowding, and length of stay in EDs **[R4]**.

The data analytics front-end of the discrete conditional model can take many forms. For example, a Bayesian network has been shown to fit well and has the added benefit of visualising the variables in a network of the inter-related nodes and edges **[R3]**. The complexity extends when multiple patient data are updated over time, and the characteristics in the data analytics model require dynamic updating. This has been accommodated by QUB researchers by integrating into the conditional phase-type model, dynamic Bayesian networks and survival trees **[R5]**. These have had impact in their application to Covid19 symptom data and has proved successful as acting as an early warning indicator of the disease pandemic changes over time and forms another one of the *"basis components"* of the NI Covid19 Model **[S2]**.

3. References to the research

R1. Brown, E., **Marshall, A. H.**, Mitchell, H. J., & Bryne, A. W. (2019). "Cattle movements in Northern Ireland form a robust network: implications for disease management". *Preventive veterinary medicine*, 170, 104740. https://doi.org/10.1016/j.prevetmed.2019.104740

R2. Marshall A. H. and McClean S. I. (2004), "Using Coxian Phase-Type Distributions to Identify Patient Characteristics for Duration of Stay in Hospital". *Health Care Management Science Journal*, 7, pp. 285-289. https://doi.org/10.1007/s10729-004-7537-z

R3. Marshall A. H., Burns M. L., Shaw B. (2007), "Patient activity in hospital using discrete conditional phase-type (DC-Ph) models". In Skiadas (Ed), *Recent Advances in Stochastic Modelling and Data Analysis*, pp. 154-161. https://doi.org/10.1142/9789812709691_0019

R4. Boyle L. M., **Marshall A. H.**, Mackay M. (2020), "Coxian Phase-Type Regression Models for Understanding the Relationship Between Patient Attributes, Overcrowding, and Length of Stay in Hospital Emergency Departments". In Belanger V, *et al. Health Care Systems Engineering* (HCSE 2019), Springer. pp. 53-64. https://doi.org/10.1007/978-3-030-39694-7_5

R5. Gordon A., **Marshall A. H.**, & Zenga M. (2017), "Predicting elderly patient length of stay in hospital and community care using a series of conditional Coxian phase-type distributions, further conditioned on a survival tree prediction using CCPh distributions and a survival tree". *Health Care Management Science*, 21, pp 269-280. https://doi.org/10.1007/s10729-017-9411-9

4. Details of the impact

The research in Section 2 underpinned the creation of a suite of statistical data analytics models by the Mathematical Sciences Research Centre (MSRC), which have been applied to 1) model disease prevalence and spread in the general population (extrinsic), and 2) model hospital resource demands via capturing correlations in patient waiting time, disease progression, and survival (intrinsic). This research directly led to the development of the Northern Ireland Covid19 model, which has significantly impacted government policy. Furthermore, the hospital resource models have been implemented and verified as effective.

The application of these models has measurably improved the management of hospital and healthcare resources in Northern Ireland (NI) and the rest of the UK.

Social Network Analysis and Coxian flow model informed government policy

Modelling for Emergency planning in the Covid19 pandemic ((i) in Summary of Impact**)**

The research outlined in **[R1-R5]** has led to Prof Marshall being invited to be part of the Northern Ireland 'Covid19 Modelling Group' (CMG) to work directly with the Chief Medical Officer (CMO) of Northern Ireland who states: "*The previous research work conducted by Adele and her team …. has impacted by informing new policy in measuring the effectiveness of Non-Pharmaceutical Interventions (NPIs) including social distancing measures, modelling the pandemic, the hospital admissions and length of stay and early alerts in the symptoms models which all are components that feed into the Stochastic SIR Covid19 Northern Ireland model"* **[S1]**.

The efficacy of Prof Marshall's models in policy planning as part of the CMG saw additional input requested for "*the Four Nations Analysts group for which she is a member working alongside Department of Health analyst colleagues across all four nations since August 2020"*. This work informed best practice for International Travel Corridors and Travel Advice methodology in the context of Green, Amber and Red countries.

Importantly, the CMO outlines how "*Prof Marshall brings her previous modelling research from the Queen's group and Northern Ireland Covid19 model outputs to the Alert Level Weekly Updates by the Joint Biosecurity Centre for UK."* **[S1]**. The Joint Biosecurity Centre (JBC), set up under the auspices of the Department of Health & Social Care, is a UK Government body that provides evidence-based, objective analysis to inform local and national decision-making in response to Covid19 outbreaks. In particular, it advises the 4 Chief Medical Officers of the UK and the Medical Director of NHS England of any changes to the alert level for the risk to public of Covid19, with the CMOs and then advising ministers. As outputs **[R1]** and **[R5]** are "*basis components*" in the *Northern Ireland Covid19 model* **[S2],** this is direct evidence of contributing to Alert Level setting.

The Head of National Alerting and Assessment for JBC emphasises the direct nature of the contributions of Prof Marshall's models to this work: *"the Alert Level update is read every week by all four Chief Medical Officers and directly informs their decision making on the COVID-19 pandemic. Decisions on the national response to COVID-19 over the last 12 months have been built directly on the evidence from you and your colleagues"* and strengthens this position by stating that "*Your work ensures that Northern Ireland's position is properly understood; and your engagement with our other participants contributes to the quality of the whole report."* **[S3]**

These unequivocal statements from the CMO for NI and the Head of National Alerting and Assessment for JBC demonstrate how crucially important the underpinning research performed by Prof Marshall over the last two decades at QUB has been in establishing a regionally leading, and nationally important, modelling capability that has contributed directly the UK's preparedness to tackle epidemics, pandemics and healthcare management.

In addition to this, the Lead Data Scientist in the Data Analytics Research & Exploitation (DARE) of the Strategic Investment Board (SIB) of NI, provides minute detail of how Professor Marshall's work maps directly onto the statements above. SIB is a key stakeholder in major government initiatives and their Lead Data Scientist's assertion that "*it is the models that I have developed with the underpinning research and models by Prof. Marshall and her team, that has made the models developed by the Northern Ireland COVID-19 Modelling Group the most accurate and realistic and most trusted by the Chief Scientific Advisor."* **[S2]** once again identifies the research conducted in QUB as being the central tenet used to inform local **[S4, S5, S6]**, and strong contributor to national **[S3],** government policy throughout the pandemic.

Strong public interest in this work saw the Belfast Telegraph profiling the CMG, including Prof. Marshall, under the title "*Behind-the-scenes team navigating Northern Ireland out of crisis"* **[S6]** again emphasising the influence of this research underpinning these decisions right across society.

Application of Coxian Phase-type Distribution Models in Healthcare Management

South Eastern H&SC Trust (Ulster hospital) ((ii) in Summary of Impact**)**

Previous research in discrete conditional phase-type models **[R1, R2]** has led to the formulation of a predictive ambulatory activity model for the South Eastern Health and Social Care Trust (SEHSCT) currently in use in the Ulster Hospital. By gaining a better understanding of the ED activity and expected workload, from the ambulatory activity model, the SEHSCT has changed the shift patterns of staff in the ED with the inclusion of a senior medical shift in the late evening shift 20:00-03:00. "*This has resulted in a reduction in the amount of patients who have to wait longer than 60 minutes before being seen by a doctor yielding an improvement of 14% overall with 80% of patients being seen by a doctor compared to 70% seen in the previous year. This is equivalent to significantly shortening waiting times to less than one hour and reducing any potential adverse effects of prolonged waiting for approximately 10,000 patients this year. The predictive model is achieving greater than 85% accuracy within a 4-hour timeslot. For core working hours, of 0800-2000, the accuracy against pre-covid activity was consistently above 92%"* **[S7].** The Ulster Hospital also report improved management of their resources, better control on their spend (in particular on temporary staffing), streamlining of the patient journey, and the release of doctors to care for critically unwell patients. A second model taken from [**R4**] identifies the key factors that are used to predict whether a patient can be treated through ambulatory care or whether they will need to be admitted to hospital. The hospital reports that the model is used on a daily basis to manage the flow of respiratory patients through the ED and as a result *"reduces hospital admissions by approximately 4745 patients per year"* **[S7]**.

5. Sources to corroborate the impact

[S1] Testimonial from the Chief Medical Officer for Northern Ireland

[S2] Testimonial from the Lead Data Scientist, Strategic Investment Board, Northern Ireland

[S3] Testimonial from the Head of National Alerting & Assessment, Joint Biosecurity Centre, UK Government.

[S4] Department of Health, "Minister highlights key findings of NI Covid-19 modelling", April 2020
https://www.health-ni.gov.uk/news/minister-highlights-key-findings-ni-covid-19-modelling

[S5] Department of Health, "Modelling data shows social distancing works", May 2020, https://www.health-ni.gov.uk/news/modelling-data-shows-social-distancing-works

[S6] Belfast Telegraph, "Behind-the-scenes team navigating Northern Ireland out of crisis", May 2020, https://www.belfasttelegraph.co.uk/news/northern-ireland/behind-the-scenes-team-navigating-northern-ireland-out-of-crisis-39209672.html

[S7] Testimonial from the Assistant Director for Unscheduled Care, South Eastern Health Social Care Trust (SEHSCT)