Should doctors use their judgment? How a System Dynamics model elicited knowledge in neonatal care services

Reda Lebcir.

Hertfordshire Business School. University of Hertfordshire. Hatfield AL10 9AB, UK

Email: M.R.Lebcir@Herts.ac.uk

Rifat Atun

T.H.Chan School of Public Health. Harvard University. Boston, MA 02115, USA.

Email: ratun@hsph.harvard.edu

Corresponding Author: Reda Lebcir

M.R.Lebcir@Herts.ac.uk

Abstract: Neonatal services, which provide care to babies born with medical complications, are under financial pressures in the United Kingdom (UK). The services are heavily regulated affecting efficiency. One possible solution is to reduce the Length of Stay (LoS) of babies and allow a degree of doctors' clinical judgment to make decisions.

The aim of this paper is to determine the impact of using clinical judgment on neonatal services performance. To achieve this aim, a System Dynamics (SD) model was built and validated in a UK neonatal unit. The model was used initially to evaluate the impact of LoS reduction on performance as clinical judgment policies were rejected by participants. The counterintuitive results led to a behavioural change and participants accepted clinical judgment policies, which the results indicated should lead to considerable performance improvement. Results' implications and SD's modelling ability to foster learning and alter participants' perceptions and behaviour are discussed.

Keywords: OR in Health; Neonatal Services; Behavioural OR; Simulation; System Dynamics.

Introduction

New born babies with health complications, or health risks due to premature birth (gestation period under 37 weeks), or low birth weight (under 2500g) need to be treated in special neonatal units (Demir et al, 2014). In the United Kingdom (UK), neonatal services are part of the National Health Service (NHS). These services are under pressure due to a sharp upward trend in demand as around 13% of births in the UK require neonatal care (Boyle et al, 2015; Asaduzzaman et al, 2009). The number of neonatal admissions in the NHS has increased from 61,372 in 2007 to 94,145 in 2013 and then to 110,997 in 2015 requiring a total of 1,157,091 care days (RCPCH, 2016; NDAU 2015, 2010). This increase is driven by higher birth rates and the developments in medical technologies, which allow treatment of more complex medical conditions even for babies born very prematurely (Boyle et al, 2015; Battin et al, 2012).

This high demand is impacting the functioning of neonatal units whose capacity has become chronically saturated. Data collected over a 6-month period in 2006-2007 indicate that neonatal units in the UK were closed to new admissions for an average of 24 days and 1 in 10 units exceeded their capacity for intensive care for more than 50 days during that period (Asaduzzaman et al, 2011). A recent document by BLISS, the prenatal charity in the UK, reported that a third of neonatal units had an occupancy rate of 80% with 9% reaching 100%, and 70% of intensive care units had an occupancy rate above the recommended safe level (BLISS, 2015).

There are no signs that this situation is going to improve as the fiscal austerity policies imposed by consecutive UK governments affected the funding allocated to the NHS. The 2015 BLISS report highlights a difficult situation regarding the level of neonatal resources as 64% of neonatal units do not have enough nurses and two thirds do not have enough medical staff to meet national standards on safe staffing levels (BLISS, 2015).

These challenges are exacerbated by the fact that care pathways and treatment procedures in neonatal services are heavily regulated and exclude any judgment on them by medical staff regardless of their experience and seniority. The regulation includes, for example, the definition of neonatal care categories, the medical conditions associated with each care category, the types of neonatal care units, and the different groups of clinical staff involved in neonatal care delivery (BAPM, 2010; DoH, 2009).

A new born baby admitted to a neonatal care unit is allocated to one of three care categories. These are, in ascending order of severity: (i) Special Care (SC), those who cannot be provided with care at home and need to be in hospital for breathing and heart rate checks; (ii) High Dependency Care (HDC), those who need continuous monitoring because they require breathing help or intravenous feeding, and (iii) Intensive Care (IC), those with very complex health problems needing, for example, breathing through mechanical ventilation.

Two areas of regulation in this multi care category structure have been found to create management complexities in neonatal units (Demir et al, 2014). First, the duration of the treatment (known as Length of Stay (LoS)) cannot be reduced regardless of the improvement in the clinical condition of the baby. Second, some treatment outcomes involve transfer between care categories. For example, some babies in the HDC care category do not leave the unit following treatment. They are, instead, moved to the SC care category or the IC care category. These cases can go through a number of treatment cycles between the three care categories spending the full LoS each time a transfer from one care category to another occurs. Consequently, the total time they spend in the unit includes all the LoS durations of the care categories they transit through until they leave the unit. This rigidity causes high occupancy rates and saturation of capacity and highlights the need for more streamlined care pathways to improve operational efficiency and cope with demand.

One possible way of achieving greater operational efficiency is to allow a degree of flexibility in neonatal treatment procedures by providing a space for medical staff to use their judgment. Clinical judgment is a long established feature of medical practice and doctors have always approached their profession by combining their knowledge, education, and experience when making decisions about the diagnosis of patients and the best course of treatment (Benner et al, 2009; Montgomery, 2006). For this reason, education and training methods have evolved significantly to enhance medical staff clinical judgment skills in order to improve the quality of medical practice (Fava et al, 2015; Lasater, 2007). Clinical judgment is defined as “the application of information based on actual observation of a patient combined with subjective and objective data that lead to a conclusion” (Mosby, 2009). In the context of neonatal services, clinical judgment can influence a number of decisions to streamline care pathways. Examples include reducing the LoS for those transferred from HDC to SC, increasing the number of babies discharged following HDC treatment instead of transferring them to SC, and moving more babies to less severe care categories (for example from IC to SC rather than to HDC). An evaluation of the impact of such policies could provide evidence on whether the use of clinical judgment is worthwhile in neonatal services.

System Dynamics (SD) methodology is suitable to evaluate the impact of clinical judgment on operational efficiency and performance of neonatal services. SD is underpinned by Systems Thinking principles and is appropriate to represent systems involving feedback processes, time delays, and non-linear relationships (Morecroft, 2015; Sterman, 2000). These characteristics are very common in health systems (Taylor and Dangerfield, 2005; Dangerfield, 1999) and this explains the recent upward trend of SD applications in healthcare (Chang et al, 2017; Brailsford et al; 2009). SD popularity stems also from its ability to capture the soft human behavioural aspects, which are very common in health contexts (Dangerfield, 2014; Taylor and Lane, 1998). SD combines qualitative and quantitative tools to evaluate and predict possible outcomes of policies and to facilitate organisational learning and knowledge elicitation among policy makers (Thompson et al, 2016; Andersen et al, 2007; Hämäläinen et al, 2013).

The aim of this paper is to evaluate the impact of reducing LoS on neonatal services performance and determine if allowing doctors to use their clinical knowledge and judgment has an effect on this performance. The paper's contribution is twofold: first, modelling of an important behavioural aspect of medical practice, which has been virtually untapped within the Operational Research (OR) discipline and, second, exploring this aspect in neonatal services, an area of healthcare significantly overlooked in past health OR related research. The paper starts with a literature review section followed by a description of the development and validation of the SD model. The following section focusses on the interventions and scenarios evaluated on the model and how the results led to group learning and shift in the consideration of policies. The paper concludes with a section discussing the implications of the results from academic and health policy perspectives.

2. Literature Review

2.1 Modelling of LoS in neonatal services: Neonatal services in the UK are part of the NHS and tasked with providing care to babies born prematurely or at term but with medical complications. The care pathways in neonatal units involve admission, assessment of severity and allocation to a care category, treatment, and associated outcomes. The LoS is regulated for every care category and, as some of the treatment outcomes involve a transition from one care category to another (for example from SC to HDC), the total time spent in a unit can include, in some cases, a number of care category LoS durations.

Given the influence of LoS on the performance of neonatal units in terms, for example, of the number of babies treated in a neonatal unit per year or the number of those refused entry because of a unit capacity saturation (Demir et al, 2014), a number of models were developed to explore the factors affecting LoS and how the latter influences performance. Lee et al (2001) developed a model representing the drivers of LoS in US neonatal services following the introduction of new guidance regulating the services. A similar research aim was explored through statistical modelling to determine the most important factors affecting neonatal LoS in many countries (Adebandji et al, 2015; Manktelow et al, 2010; Yau et al, 2003). Other research investigated the relationship between neonatal LoS and clinical outcomes such as discharge, death, and transfer to other care wards (Boyle et al, 2015) or its role in informing decisions such as capacity requirements (Asaduzzaman et al, 2009). Despite the richness of this research, its scope was limited to the estimation of LoS and did not investigate how it affects the performance of neonatal units.

2.2 System Dynamics modelling in neonatal services: SD gained significant popularity and importance in health management in recent times. This is due to increased awareness that health contexts are dynamic, complex, and, very often, respond to policies and interventions in a counterintuitive way (Homer and Hirsch, 2006; Dangerfield, 1999). This created a need for more robust and evidence-based methodologies to inform policies in the health sector leading to a significant increase in the number of health related SD applications (Chang et al, 2017; Brailsford et al; 2009). SD represents the structure of healthcare contexts in the form of causality maps known as Causal Loop Diagrams (CLDs). Computer simulation models of the contexts are then developed using appropriate software (for example STELLA or Vensim) providing decision makers with a virtual environment to test the likely impact of policy interventions. Joint analysis of simulation results and CLDs has been found to stimulate decision makers' learning and understanding of the complex nature of health contexts and how these are affected by the interventions (Monks et al, 2016).

However, there are few SD models of neonatal services. Demir et al (2014) developed an SD model of a neonatal unit in the UK, which provides treatment to all care categories with a focus on the impact of LoS on some operational indicators such as the number of babies discharged, transferred to other hospital wards, and refused entry to the unit. In another study in Uganda, an SD model was developed to explore policy interventions aiming to reduce neonatal mortality (Rwashana-Semwanga et al, 2016). The findings suggested that the combination of two interventions, namely, the provision of clean free delivery kits and free transportation vouchers to health facilities during emergencies was the most effective in reducing neonatal mortality.

2.3 Group model building and knowledge elicitation in SD: Group model building has been an integral part of SD since its early days. The reasons for this include the reliance of SD modellers on the knowledge of participants in the organisational context of the study and the importance of developing participants' ownership so that findings from SD models are accepted and implemented (Rouwette et al, 2011; Vennix, 1999). This close relationship between SD modellers and participants is also important as SD tackles systems which behave in a counterintuitive manner. More often than not, SD models generate surprising results, which challenge participants' knowledge and change their mental models and sense making of the systems in which they operate (Hämäläinen et al, 2013; Thompson et al, 2016; Andersen et al, 2007).

These situations where participants are surprised and the ensuing changes in their mental models are referred to as learning instances. A wide range of studies found that this type of learning is universal and occurs regardless of the participants (university students, top management team in an organisation) or whether the SD model represents a real world context (an existing hospital) or a hypothetical one (a planned ophthalmology department which has not yet been built) (Thompson et al, 2016; Monks et al, 2016; Rouwette et al, 2011; Morecroft, 1992).

Despite the established track record of SD models in facilitating learning and eliciting knowledge, it seems that the health sector has not benefited from the evidence generated. Although SD health research has expanded significantly in recent years, it is still focused on the evaluation of interventions and policies. This is surprising given the complex nature of health contexts, the team based approach to healthcare delivery and management, and the potential of SD to enhance learning in health settings.

3. Model Development

The model was developed in a big neonatal unit located in a densely populated area in the UK. The unit was equipped to treat babies in all three care categories. Babies admitted to the unit are mainly from its local area with some transferred from other units in the region if these are not equipped to treat babies in the HDC and IC categories. The unit was chronically saturated due to its reputation for high quality of care, and a considerable number of babies were refused admission because of non-availability of resources to treat them.

A core team led by the head of the neonatal unit with the membership of a senior neonatal consultant, a senior academic in health systems, and an SD modeller was set up at the beginning of the project. The SD modeller, responsible for model building, was supported by members of the core team, who provided the necessary information and documents. Where input from other stakeholders (nurses, doctors, data managers) was required, access was facilitated by the head of the neonatal unit. The project lasted for 15 months and the core team had eight meetings to review the model building progress and agree on subsequent project activities. Additional meetings were organised on a quarterly basis to present the updated version of the model to the core team and stakeholders' representatives, which kept them engaged in the project. The icon-based user friendly interface of the SD STELLA software was new and attractive to the participants in these meetings. They enquired about several aspects of the model and provided the modeller with comments, suggestions, and encouragements to continue the model building process.

The SD modelling process started by developing the qualitative map (in the form of a CLD) of the neonatal unit structure. This was followed by the building of the SD simulation model portraying the unit care pathways including admission, initial clinical status check, assignment to a care category, treatment processes and outcomes, resources, and the rules governing the evolution of babies in the unit. Feedback and comments during the core team and quarterly review meetings led to the refinement and improvement of the CLD and the simulation model before they were approved by all participants.

The care pathways represented in the model follow the recommended national rules and standards (BAPM, 2010; DoH, 2009) and the professional consensus among the neonatal clinical community. Upon admission, a clinical check takes place and the baby is allocated to one of the care categories. The admission rate depends on the number of babies born per day, the fraction of these needing neonatal care, and the effect of cot availability on admission. The admission rate equation is as follows

$$ADR_t = BRN_t \times FNC \times ECAR \qquad (1)$$

Where

ADR_t: Admission rate per day [babies/day].

BRN_t: Number of babies born per day [babies/day].

FNC: Fraction of babies needing neonatal care [Dimensionless].

ECAR: Effect of cot availability on admission [Dimensionless].

The effect of cot availability on admission represents a managerial rule, which reduces the number of babies admitted as the number of cots occupied increases and the unit nears saturation. This rule is represented by the following decreasing non-linear function (See Figure 1)

$$ECAR = f(FRCOT_t) \qquad (2)$$

Where

$FRCOT_t$: Fraction of cots used [Dimensionless].

$FRCOT_t$ represents the fraction of the number of cots used to the total number of cots available in the unit.

Following allocation to a care category (SC, HDC, IC), the baby enters a treatment phase lasting for a period equal to the established care category LoS. Treatment outcomes can be death, discharge home, transfer to hospital wards outside the unit, transfer to other neonatal units, or transition to another care category. Regarding the latter outcome, transitions can take place (depending on the clinical state at the end of the treatment) from SC to HDC or IC, from HDC to SC or IC, and from IC to HDC or SC (see Figure 2 for a high level structure of the care pathways). An example of one of the stock and flow diagrams (HDC) is presented in Figure 3.

The treatment rate (the number of babies treated daily) depends on the LoS and the availability of resources. The treatment rate equation is as follows

$$TRT_{t,j} = Min(TLOS_{t,j}, TRES_{t,j}) \qquad (3)$$
$$j \in \{SC, HDC, IC\}$$

Where

$TRT_{t,j}$: Treatment rate for babies in care category j [babies/day].

$TLOS_{t,j}$: Treatment rate allowed by the LoS for babies in care category j [babies/day].

$TRES_{t,j}$: Treatment rate allowed by resources for babies in care category j [babies/day].

$TLOS_{t,j}$ reflects the fact that clinical regulations and established processes require that babies stay in the treatment phase for the whole LoS regardless of the level of resources in the neonatal unit. As a result, the $TLOS_{t,j}$ equation is as follows:

$$TLOS_{t,\,j} = \frac{NBAB_{t,\,j}}{LoS_j} \qquad (4)$$

$$j \in \{SC, HDC, IC\}$$

Where

$NBAB_{t,j}$: Number of babies in care category j needing treatment [babies].

LoS_j: Length of stay for babies in care category j [days].

$TRES_{t,j}$ represents the treatment rate allowed by resources and, as such, reflects the effect of resources availability on the daily treatment capacity. Resources required for treatment include nurses, doctors, and cots (BAPM, 2010). There are 3 categories of doctors: consultants, specialist registered, and senior officer, who can treat all categories but with different ratios of doctors to babies. Similarly, nurses are grouped into support nurses, who can treat SC babies only, non-specialist nurses, who can treat SC and HDC babies, and specialist nurses, who can treat babies in all the three care categories. As it is the case for doctors, the ratio of nurses to babies is also regulated for every category of nurses. There are two types of cots: (i) SC cots adequate for SC only and (ii) IC cots adequate for all categories.

As an example, the treatment rate allowed by resources for HDC babies is presented in equation (5).

$$TRES_{t,\,HDC} = Min(NUR_{t,\,HDC}, DOC_{t,\,HDC}, COT_{t,\,HDC}) \qquad (5)$$

Where

$TRES_{t,HDC}$: Treatment rate allowed by resources for HDC babies [babies/day].

$NUR_{t,HDC}$: Nurses treatment rate for HDC babies [babies/day].

$DOC_{t,HDC}$: Doctors treatment rate HDC babies [babies/day].

$COT_{t,HDC}$: Cots treatment rate HDC babies [babies/day].

$NUR_{t,HDC}$ represents the total number of HDC babies, which can be treated by the nurses every day. Similarly, $DOC_{t,HDC}$ and $COT_{t,HDC}$ represent the same number for doctors and cots respectively. The minimum condition in equation (5) reflects the fact that the treatment rate allowed by resources, which represents the daily treatment capacity, is determined by the least available resource in the unit.

As an illustration, $NUR_{t,HDC}$ is determined by the total number of nurses in the unit, the daily fraction of time allocated by nurses to treatment activities, the maximum number of babies a single nurse can treat, and the fraction of treatment capacity allocated to HDC babies. The equation for $NUR_{t,HDC}$ is as follows

$$NUR_{t,\,HDC} = \sum_{i=1}^{2} (A_i \times B_i \times C_{i,\,HDC} \times D_{t,\,HDC}) \quad (6)$$

$$i \in \{SPE, NSP\}$$

Where

A_i: Total number of nurses of category i in the unit [Nurses].

B_i: Daily fraction of time allocated by nurses of category i to treatment activities [fraction/day].

$C_{i,\,HDC}$: Ratio of nurses of category i to HDC babies [babies/nurse].

$D_{t,HDC}$: Fraction of nurses treatment activities allocated to HDC babies [dimensionless].

SPE: Specialist nurses.

NSP: Non specialist nurses.

The model captures the performance of the unit through a set of performance indicators representing the throughput from the unit (the number of babies leaving the unit following treatment), and the unit ability to cope with demand represented by the number of cases of refused admission due to unit capacity saturation.

4. Model parameters and validation

The data used to populate the model was collected from national regulatory sources and members of the project core team. National regulatory sources provided data such as doctors and nurses ratios to babies. The core team members made information available on the established LoS for each care category, the number of doctors, nurses, cots in the unit, the fraction of time allocated by doctors and nurses to treatment activities, and the rules governing admission to the unit.

Following inclusion of data in the model, validation tests (Sterman, 2000) were performed with the active involvement of the core team members and stakeholders' representatives. The qualitative map and the variables included in the model were confirmed in a validation workshop. The model equations were verified by the modeller and checked for dimensional consistency. A list of the variables included in the model was presented during the validation workshop to check that every variable was meaningful and that no dummy variables were included in the model. Extreme conditions tests (for example the death of all babies following treatment) were performed on the model and the results were consistent with expectations. The last validation test focused on checking the model ability to replicate past observed data. This test covered a number of variables and the model predictions were very close to real world observations. As an example, the results of this test regarding the cumulative number of SC babies discharged are presented in Figure 4.

5. Scenario Analysis and Results

5.1 Selection of simulation scenarios

At the project start, the core team agreed that its aim is to evaluate the impact of reducing LoS on the unit performance. Therefore, the initial set of scenarios excluded any reference to clinical judgment (Table 1). Following model validation, a workshop involving the core team members and stakeholders' representatives was held to present and discuss the simulation results.

As the simulation generated counterintuitive results and following their explanation with the help of a CLD (more of this later), the workshop participants considered the possibility of allowing doctors to use their clinical judgment if this did not alter significantly the current pathways. Three policies were identified and agreed by the participants. These are as follows:

Policy 1: Reduce the number of babies transferred from HDC to SC by 50%. These are instead discharged home, transferred to other units, or transferred to other wards in the hospital with an equal proportion for each outcome.

Policy 2: In addition to policy 1, the number of babies transferred from IC to SC is reduced by 50%. They are instead discharged home, transferred to other units, or transferred to other wards in the hospital with an equal proportion for each outcome.

Policy 3: In addition to policy 2, the number of babies transferred from IC to HDC is reduced by 25%. These are instead discharged home, transferred to other units, or transferred to other wards in the hospital with an equal proportion for each outcome.

The simulation model was adjusted to represent these 3 policies (Table 2) and the initial set of scenarios (Table 1) was simulated under each of the three policies.

5.2 Simulation results

As described earlier, the neonatal unit performance is measured by its throughput and ability to cope with demand. Throughput refers to the number of babies leaving the unit following successful treatment and is represented through the performance indicators "Cumulative Number of Babies Discharged Home (BDH)", "Cumulative Number of Babies Transferred to Other Units (BTO)", and "Cumulative Number of Babies Transferred to other Wards in the Hospital (BWH)". The unit ability to cope with demand is represented by the "Cumulative Number of Babies Refused Entry (BRE)". Performance is considered positive if throughput (BDH, BTO, BWH) is high and BRE is low.

5.2.1 No use of clinical judgment (business as usual)

The model was run for 1 year and the results from the initial set of scenarios (Table 1) came as a great surprise to the workshop participants, who were expecting significant performance improvement (Table 3: columns "business as usual"). Instead, the model showed that, under many scenarios, performance will drop as reflected by the decreasing values of BTO and the increasing values of BRE (the most important performance indicator for the participants). This created confusion and disbelief in the workshop and many participants commented that *"the model must be wrong"* and "*how we are doing the right thing and getting the opposite results to expectations*". The modeller, who facilitated the workshop, explained that health contexts are a good example of dynamic complex systems and these tend to behave in a counterintuitive manner (Homer and Hirsch, 2006).

To overcome the confusion and enhance confidence in the model, the modeller used a high level CLD (Figure 5) representing babies in the HDC category to explain the unexpected results (HDC was selected just as an example to illustrate the unit's complexity). In the CLD, balancing loops B1 and B2 represent the processes of admission and treatment of HDC cases whereas loops B3 to B6 represent their movement towards post-treatment outcomes (B3 to B6 are balancing loops because they reflect "depletion" processes from HDC once the treatment is completed). Reinforcing loops R1 and R2 portray cycles of HDC babies transferred to SC (R1) and IC (R2) post HDC treatment who enter another treatment phase in these categories and then go back to HDC.

Reduction of LoS increases HDC treatment rate. This gives more power to loop B2 and triggers two conflicting processes. First, loops B5 and B6 become stronger leading to a faster movement of babies out of the unit, hence improving throughput. Second, loop B2 power is transmitted to B3 and R1, which increases the number of cases going through the cycle HDC to SC and back to HDC, and to loops B4 and R2 with the same consequences for the cycle HDC to IC and back to HDC. Dominance of the loops portraying the inner cycles in the unit (B3, B4, R1, and R2) increases the number of babies stuck inside the unit. This leads to higher occupation of cots and, therefore, in more babies being refused entry to the unit.

5.2.2 Use of clinical judgment

Simulation results of policies where clinical judgment is allowed (policy 1, 2, and 3) suggest that BDH decreases as clinical judgment is allowed and this trend is valid for all scenarios. However, a reverse trend is observed with respect to BTO and BWH, which increase as doctors are given the opportunity to exercise judgment. Regarding BRE, the shift to the new policies has a dramatic effect. The decline is moderate under policy 1 but becomes steep under policies 2 and 3 and this is valid for all scenarios as shown in Figure 6. The average reduction is 30%, 61%, and 72% under policies 1, 2, and 3 respectively. The most significant reduction is achieved under scenario 4 reaching 46%, 64%, and 80% under policies 1, 2, and 3 respectively. As this scenario assumes reducing LoS for HDC by 3 days, it highlights the importance of HDC (the intermediate care category) as a key driver of the unit saturation.

6. Discussion and Conclusion

This paper focusses on an important area of health management, namely, the ability of doctors to use their clinical judgment, an important aspect of medical practice (Benner et al, 2009; Montgomery, 2006). Equally important is the medical context of the study as neonatal services attracted virtually no attention from the OR community despite the increased demand for neonatal care, the vulnerability of the patients, and the high treatment cost (BLISS, 2015; Demir et al, 2014; Asaduzzaman et al, 2009). This research is pertinent and comes at a time when the NHS faces unprecedented challenges to provide high quality care with less resources (Lafond and Charlesworth, 2016).

The use of the SD methodology to address the research aims led to some interesting and surprising findings, which altered the collective understanding of the problem by the project stakeholders and participants and widened the scope of policies evaluated beyond what was agreed at the project outset. Furthermore, this research provides a vivid example on how SD can generate new understanding, challenges established assumptions, enhances team learning, and guides policy making towards new directions The project shows how exposure to and interaction with SD modelling can lead to significant shifts in positions, a key feature of behavioural OR (Hämäläinen et al, 2013).

The results indicate that improving the performance of neonatal services requires a combination of LoS reduction and reconfiguration of care pathways by allowing doctors to use their clinical judgment. This is a good example on how the implementation of a single policy, however logical it may seem, does not lead to performance improvement and the objective can only be achieved if a number of policies are deployed simultaneously. This is a known characteristic of dynamically complex systems (such as health systems), which behave in a counterintuitive manner and defeat what may appear to be logical and intuitive interventions (Morecroft, 2015).

From a care delivery perspective, the implementation of policies such as the ones discussed in this research, should not impact negatively on clinical outcomes and quality of care. The clinical teams (doctors, nurses) involved in neonatal treatment are highly qualified and experienced and, therefore, could be trusted to make decisions involving a degree of subjectivity, which is a core aspect of clinical judgment (Mosby, 2009).

This research provides further evidence regarding the usefulness of SD to inform policy making. The model allowed rigorous evaluation of alternative policies and provided clear information regarding their expected outcomes. As such, the model prevented implementation of costly changes, which would have been in vain. The results show that success is only achieved if the rigidity of the care pathways is relaxed to a certain extent. Neonatal services policy makers need to find ways to reconciliate between the need for a regulated care process to protect vulnerable patients and trusting doctors to change their behaviour and exercise clinical judgment.

Additionally, this research is an excellent example of SD's ability to facilitate engagement with participants in real world contexts and change their behaviour and understanding of the systems in which they operate. The results of the initial LoS reduction policy were met with scepticism, disbelief, indeed hostility. As various SD tools were used to generate and explain the results, shock and resistance changed to a realisation that neonatal services are far more complex than what was originally assumed. This led to a more positive attitude when participants were asked to come up with suggestions for innovative policies. Once these policies were simulated and results determined, the initial scepticism about SD turned into adoption of the results and the methodology. All participants in the project emerged from this modelling exercise with a much improved understanding of the complexity of neonatal care services, and the importance of communication and dialogue between the management and clinical staff to implement the changes suggested by the model findings. This gives credit to the ability of OR based projects to lead to situations where perceptions and views are altered, collective behaviour changed, and organisational learning fostered (Monks et al, 2016).

This study can be expanded in a number of ways. The SD model could be widened to evaluate the impact of doctors' clinical judgment on a neonatal network not just a single unit. Another possibility is to investigate the process and the factors facilitating the implementation of the clinical judgment policies in a heavily regulated context. It would also be important to determine the impact of these policies on quality of care and readmissions.

This research demonstrates the potential of OR techniques to capture the important 'soft behavioural variables', which characterise the health sector. It also provides evidence of the ability of OR to engage health policy makers, enhance learning, and provide evidence for designing policies. This can only be welcomed in a sector where improving efficiency and quality of care are expected to be the dominant challenges in the future.

Statement of declaration of interest: There is no interest or conflict of interest to declare.

References

Adebandji A, Adeyemi S and Gyamfi M (2015). Empirical analysis of factors associated with neonatal length of stay in Sunyani, Ghana. Journal of Public Health and Epidemiology 7: 59-64.

Andersen D F, Vennix J A M Richardson G P and Rouwette E A J A (2007). Group Model Building: Problem Structuring, Policy Simulation and Decision Support. Journal of the Operational Research Society 58: 691-694.

Asaduzzaman M D, Chaussalet T J and Robertson N J (2009). A loss network model with overflow for capacity planning of a neonatal unit. Annals of Operations Research 178: 67–76.

Asaduzzaman M D, and Chaussalet T J (2011). An overflow loss network model for capacity planning of a perinatal network. Journal of the Royal Statistical Society A 174: 403-417.

Battin M R, Knight D B Kuschel C A and Howie R N (2012). Improvement in mortality of very low birthweight infants and the changing pattern of neonatal mortality: The 50 year experience of one perinatal centre. Journal of Paediatrics and Child Health 48: 596-599.

Benner P, Tanner C A and Chesla C A (2009). Expertise in nursing practice. Caring, clinical judgment, and ethics. 2nd Eds. Springer: New York.

BLISS (2015). Bliss baby report 2015: Hanging in the balance. Available at http://www.bliss.org.uk/babyreport. Last accessed 25 September 2017.

Boyle E M, Johnson S Manktelow B Seaton S E Draper E S Smith L K Dorling J Marlow N Petrou S and Field D J (2015). Neonatal outcomes and delivery of care for infants born late preterm or moderately preterm: a prospective population-based study. Archives of Disease in Childhood - Fetal and Neonatal Edition 100: F479-F485

Brailsford S C, Harper P R Patel B and Pitt M (2009). An analysis of the academic literature on simulation and modeling in healthcare. Journal of Simulation 3: 130-140.

British Association of Perinatal Medicine (BAPM) (2010). Service standards for hospitals providing neonatal care, 3rd edition. Available at https://www.bapm.org/resources/service-standards-hospitals-providing-neonatal-care-3rd-edition-2010. Last accessed 27 July 2018.

Chang, A Y, Ogbuoji O Atun R and Verguet S (2017). Dynamic modelling approaches to characterize the functioning of health systems: A systematic review of the literature. Social Science & Medicine 194: 160-167.

Dangerfield B C (1999). System dynamics applications to European health care issues. Journal of the Operational Research Society 50: 345–353.

Dangerfield B C (2014). Systems thinking and system dynamics: a primer. In: Brailsford et al (Eds). Discrete event simulation and system dynamics for management decision making. Wiley: Chichester, pp 26-51.

Demir E, Lebcir R and Adeyemi S (2014). Modelling Length of Stay and patients flows: Methodological case studies form the UK neonatal services. Journal of the Operational Research Society 65: 532-545.

Department of Health (DoH) (2009). Toolkit for High Quality Neonatal Services. Department of Health: London.

Fava G A, Guidi J Rafanelli C and Sonino N (2015). The clinical inadequacy of evidence based medicine and the need for a conceptual framework based on clinical judgment. Psychotherapy and Psychosomatics 84: 1-3.

Hämäläinen R P, Luoma J and Saarinen E (2013). On the importance of behavioral operational research: The case of understanding and communicating about dynamic systems. European Journal of Operational Research 228: 623-634.

Homer J B, and Hirsch G B (2006). System dynamics modelling for public health: background and opportunities. American Journal of Public Health 96: 452-458.

Lafond S and Charlesworth A (2016). A perfect storm: an impossible climate for NHS providers' finances? The Health Foundation. Available at https://www.health.org.uk/publications/a-perfect-storm-an-impossible-climate-for-nhs-providers%E2%80%99-finances. Last accessed 8 August 2019.

Lasater K (2007). High fidelity simulation and the development of clinical judgment: Students' experiences. Journal of Nursing Education 46: 269-276.

Lee A H, Ng A S and Yau K K (2001). Determinants of maternity length of stay: A gamma mixture risk-adjusted model. Health Care Management Science 4: 249–255.

Manktelow B, Draper E S Field C and Field D (2010). Estimates of length of neonatal stay for very premature babies in the UK. Archives of Disease in Childhood-Fetal and Neonatal Edition 95: F288:F292

Monks T, Robinson S and Kotiadis K (2016). Can involving clients in simulation studies help them solve their future problems? A transfer of learning experiment. European Journal of Operational Research 249: 919-930.

Montgomery K (2006). How doctors think: clinical judgment and the practice of medicine. Oxford University Press: Oxford.

Morecroft J D W (1992). Executive knowledge, models, and learning. European Journal of Operational Research 59: 9-27.

Morecroft J D W (2015). Strategic Modelling and Business Dynamics. A feedback systems approach. (2nd eds) Wiley: Chichester.

Mosby (2009). Clinical Judgment. Mosby's Medical Dictionary. (8th eds). Elsevier Health Sciences: London.

Neonatal Data Analysis Unit (NDAU) (2010). Report: Care days and reason for admission of term babies. Imperial College London, UK, available at http://www.imperial.ac.uk/medicine/research/researchthemes/reprodscience/paediatrics/neonatalmedicine/ndau/. Last accessed 20 December 2011.

Neonatal Data Analysis Unit (NDAU) (2015). 2015 report. Imperial College London, UK.

Rouwette E A J A, Korzilius H Vennix J A M and Jacobs E (2011). Modeling as persuasion: the impact of group model building on attitudes and behaviour. System Dynamics Review 27: 1-21.

Royal College of Paediatrics and Child Health (RCPCH) (2016). National neonatal audit programme 2016 annual report on 2015 data. Available at https://www.rcpch.ac.uk/improving-child-health/quality-improvement-and-clinical-audit/national-neonatal-audit-programme-nnap. Last accessed on 25 September 2017.

Rwashana-Semwanga A, Nakubulwa S, and Adam T (2016). Applying a system dynamics modelling approach to explore policy options for improving neo natal health in Uganda. Health Research Policy and Systems. DOI 10.1186/s12961-016-0101-8

Sterman J D (2000). Business dynamics. Systems thinking and modeling for a complex world. Mc-Graw Hill: Singapore.

Taylor K, and Dangerfield B C (2005), Modelling the feedback effects of reconfiguring health services. Journal of the Operational Research Society 56: 659-675

Taylor K, and Lane D C (1998). Simulation applied to health services: Opportunities for applying the System Dynamics approach. Journal of Health Services Research and Policy 3:226-232.

Thompson J P, Howick S and Belton V (2016). Critical learning incidents in system dynamics modelling engagements. European Journal of Operational Research 249: 945-958.

Vennix J A M (1999). Group model building: tackling messy problems. System Dynamics Review 15: 379-401.

Yau K, Lee A and Ng A (2003). Finite mixture regression model with random effects: Application to neonatal hospital length of stay. Computational Statistics and Data Analysis 41: 359–366.

Table 1: Initial set of the scenarios simulated on the model

Scenario	Description
Scenario 0	Baseline. LoS (days): SC 11.75; HDC 6.41; IC 10.21.
Scenario 1	Reduce LoS by 1 day for SC Babies
Scenario 2	Reduce LoS by 3 days for SC Babies
Scenario 3	Reduce LoS by 1 day for HDC Babies
Scenario 4	Reduce LoS by 3 days for HDC Babies
Scenario 5	Reduce LoS by 1 day for IC Babies
Scenario 6	Reduce LoS by 3 days for IC Babies
Scenario 7	Reduce LoS by 1 day for SC, HDC, and IC Babies
Scenario 8	Reduce LoS by 3 days for SC, HDC, and IC Babies

SC: Special Care

HDC: High Dependency Care

IC: Intensive Care

Table 2: Fraction of babies for all post- treatment outcomes

(Eg. SC Discharge: Fraction of babies in the SC care category discharged following treatment)

Treatment Outcome	**BAU**	**Policy 1**	**Policy 2**	**Policy 3**
SC Category				
SC Discharge	0.384	0.384	0.384	0.384
SC Transfer to other wards	0.214	0.214	0.214	0.214
SC Transfer to other hospitals	0.181	0.181	0.181	0.181
SC Death	0.000	0.000	0.000	0.000
SC to HDC	0.121	0.121	0.121	0.121
SC to IC	0.100	0.100	0.100	0.100
HDC Category				
HDC Discharge	0.009	0.133	0.133	0.133
HDC Transfer to other wards	0.006	0.131	0.131	0.131
HDC Transfer to other hospitals	0.054	0.179	0.179	0.179
HDC Death	0.000	0.000	0.000	0.000
HDC to SC	0.748	0.374	0.374	0.374
HDC to IC	0.183	0.183	0.183	0.183
IC Category				
IC Discharge	0.022	0.022	0.089	0.110
IC Transfer to wards	0.036	0.036	0.099	0.120
IC Transfer to other hospitals	0.210	0.210	0.277	0.297
IC Death	0.088	0.088	0.088	0.088
IC to SC	0.394	0.394	0.197	0.197
IC to HDC	0.250	0.250	0.250	0.188

BAU: Business As Usual (No Clinical Judgment)

SC: Special Care

HDC: High Dependency Care

IC: Intensive Care

Table 3: Simulation results for policies excluding and allowing clinical judgment.

		BDH				BTO		
Scenario	**BAU**	**Policy 1**	**Policy 2**	**Policy 3**	**BAU**	**Policy 1**	**Policy 2**	**Policy 3**
Scenario 0	230	226	213	212	187	182	194	197
Scenario 1	234	224	213	212	182	181	193	197
Scenario 2	235	226	214	212	182	182	194	197
Scenario 3	234	225	212	212	182	181	193	197
Scenario 4	233	226	213	212	181	182	194	197
Scenario 5	233	224	213	212	181	181	194	196
Scenario 6	234	225	213	211	181	181	194	196
Scenario 7	236	226	213	212	182	182	194	197
Scenario 8	240	226	214	212	185	182	195	197
		BWH				**BRE**		
Scenario	**BAU**	**Policy 1**	**Policy 2**	**Policy 3**	**BAU**	**Policy 1**	**Policy 2**	**Policy 3**
Scenario 0	136	154	156	156	18	11	8	5
Scenario 1	138	153	156	156	18	15	8	4
Scenario 2	139	154	156	156	17	11	6	5
Scenario 3	138	153	155	156	18	14	9	5
Scenario 4	138	154	156	156	20	11	7	4
Scenario 5	138	153	156	156	21	14	6	6
Scenario 6	138	153	156	155	20	13	7	6
Scenario 7	139	154	156	156	16	11	7	5
Scenario 8	139	154	157	156	12	10	4	4

BAU: Business as Usual (No Clinical Judgment)

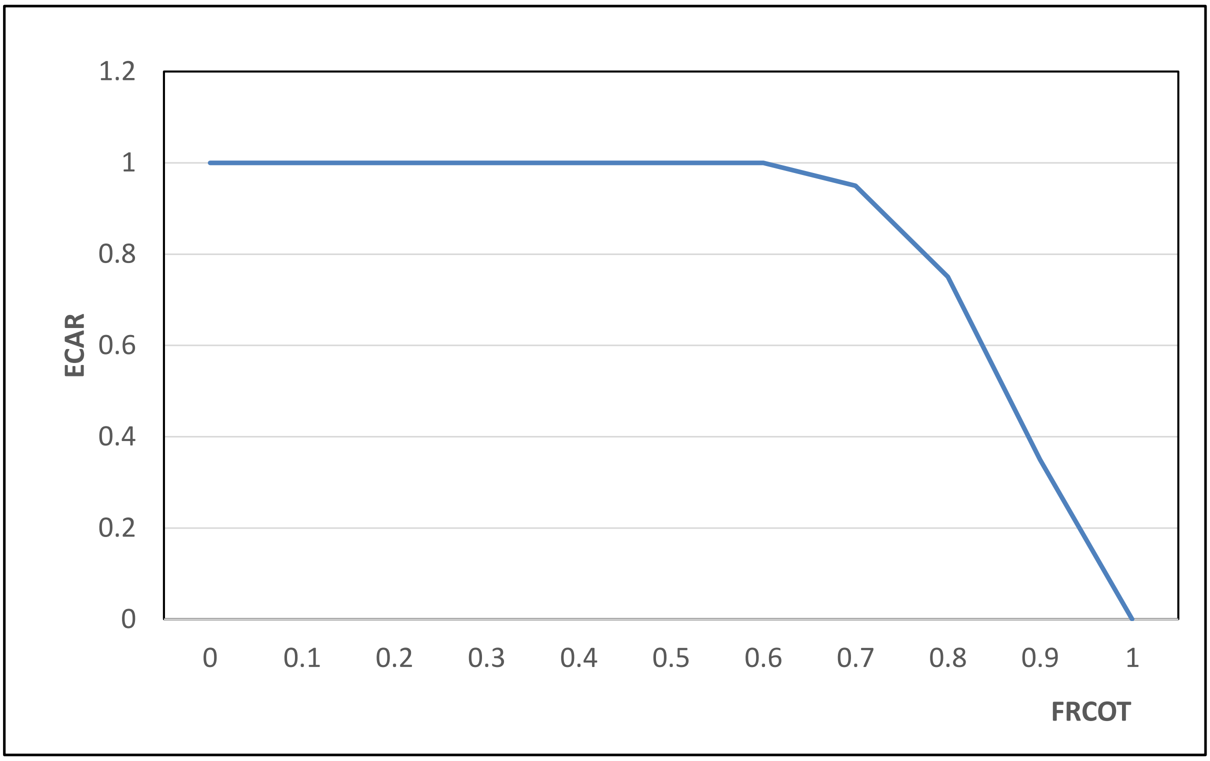
BDH: Cumulative Number of Babies Discharged Home (Babies)

BTO: Cumulative Number of Babies Transferred to Other Units (Babies)

BWH: Cumulative Number of Babies Transferred to other Wards in the Hospital (Babies)

BRE: Cumulative Number of Babies Refused Entry (Babies)

Figure 1: Graphical representation of the effect of cot availability on admission (ECAR)



ECAR: Effect of cot availability on admission

FRCOT: Fraction of cots used

Figure 2: : High level structure of the care pathway in the neonatal unit

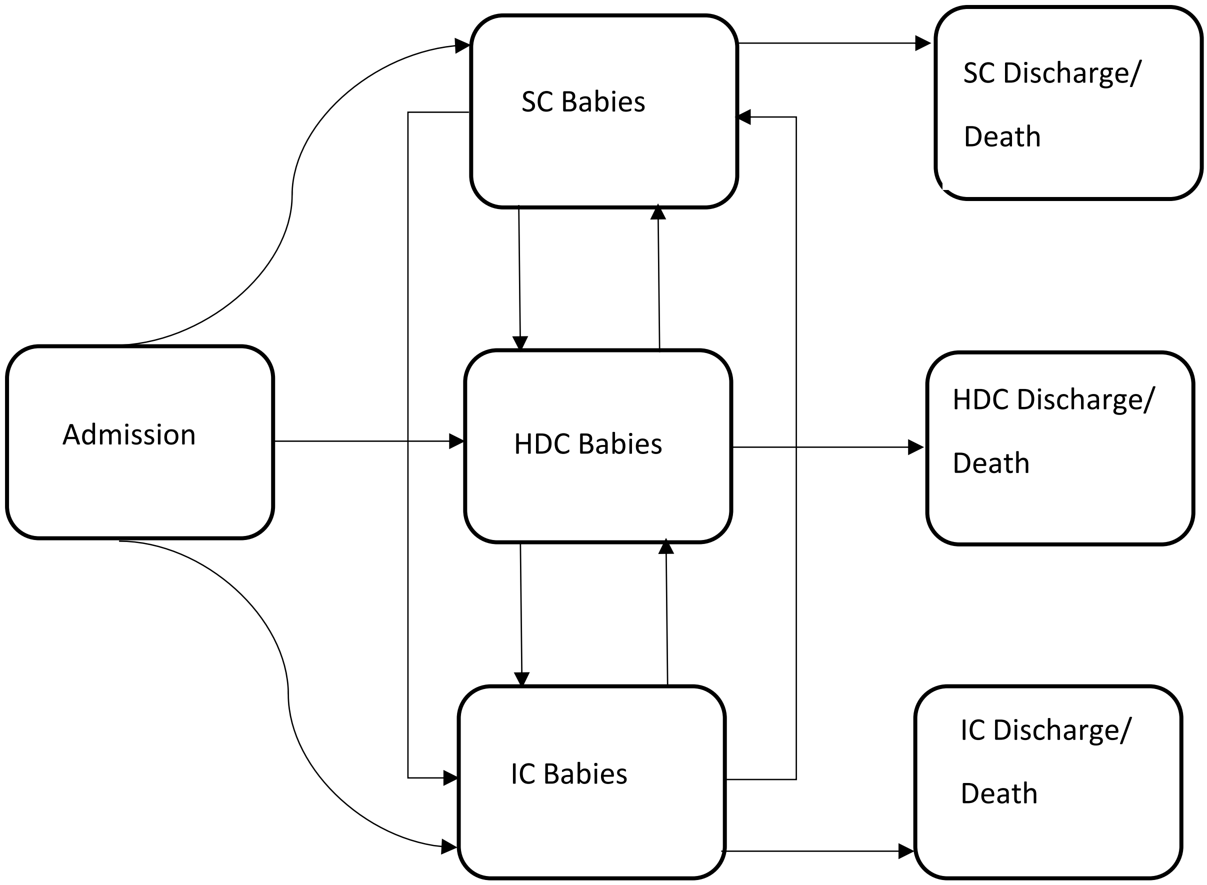


Figure 3: Stock and Flow Diagram of the High Dependency Care (HDC) care category.

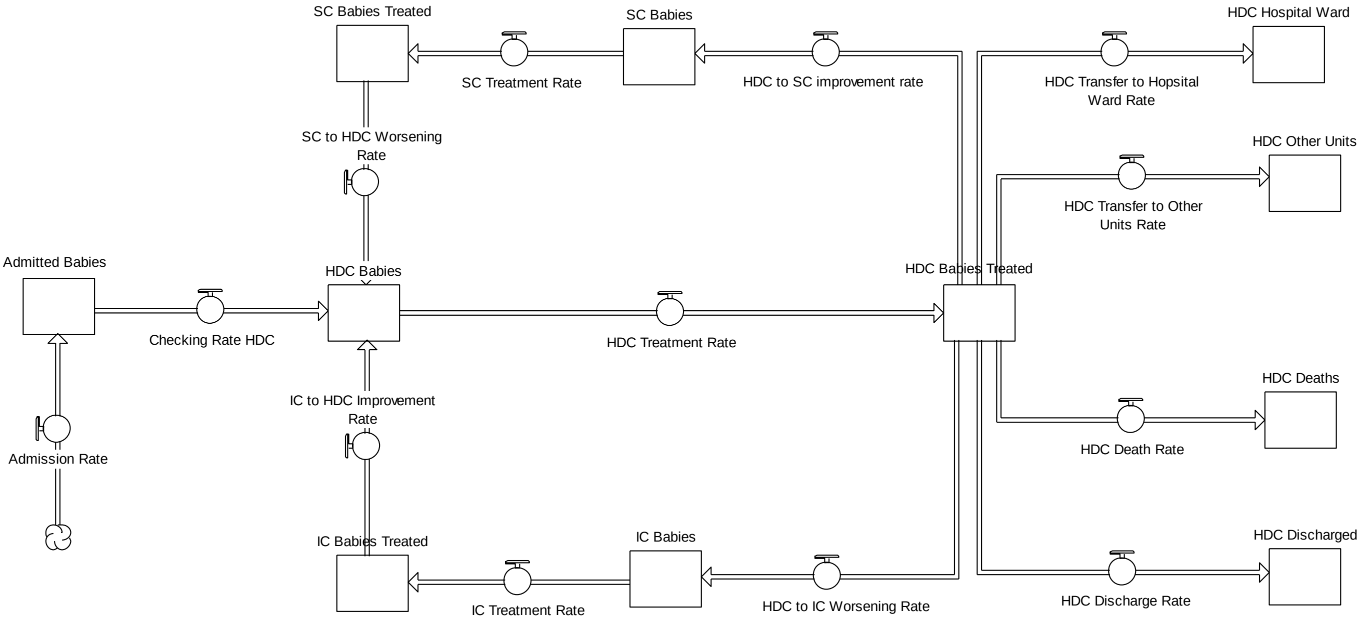


Figure 4: Real versus simulated results for the variable “Cumulative Special Care (SC) babies discharged”

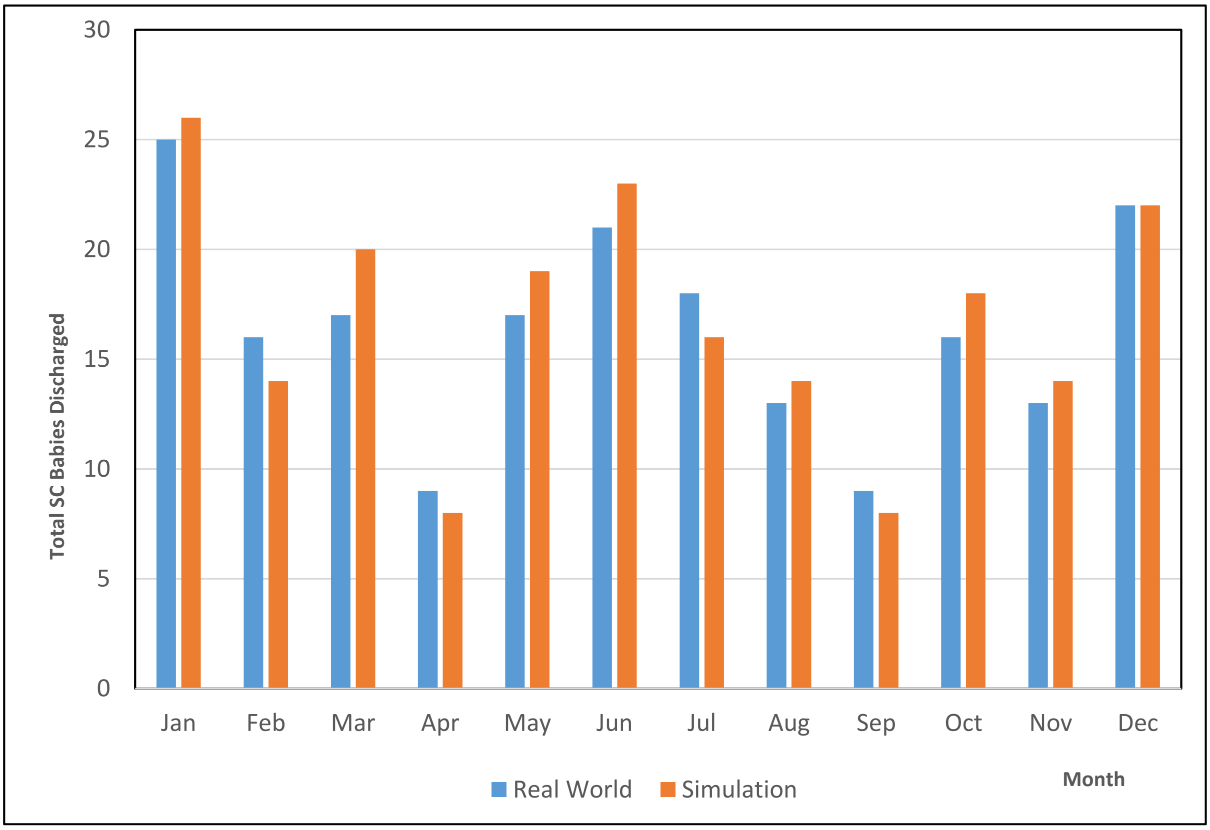


Figure 5: Causal Loop Diagram (CLD) for the policy of no use of clinical judgment in HDC

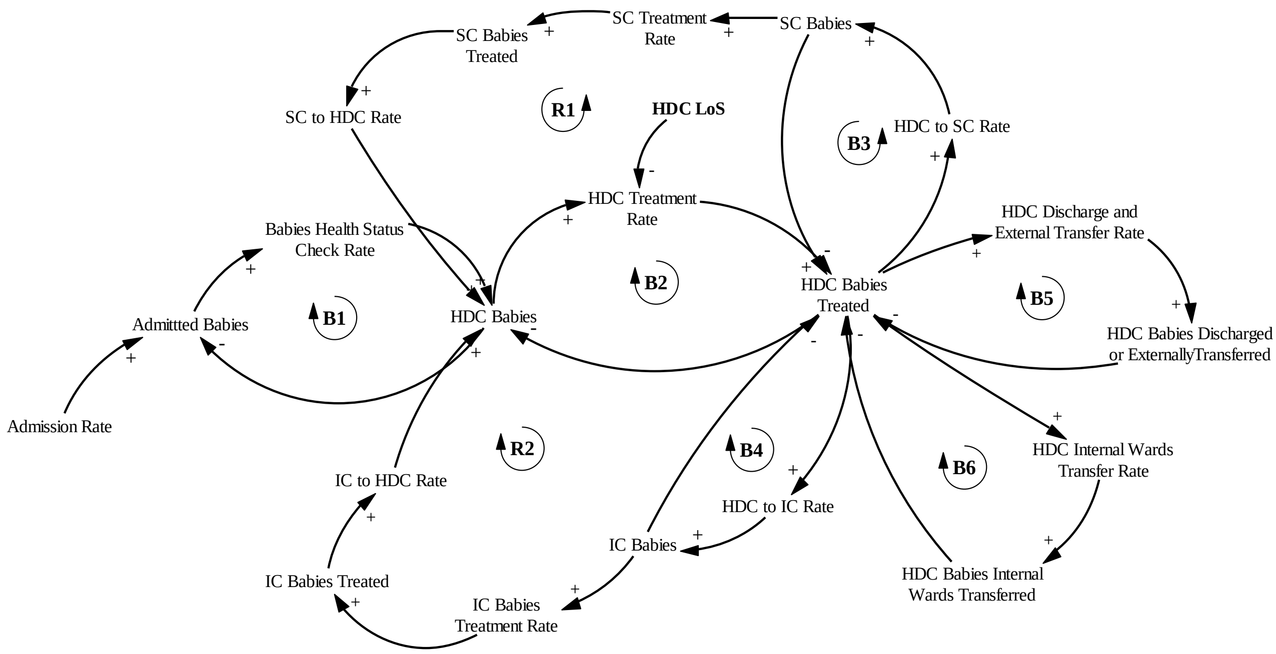
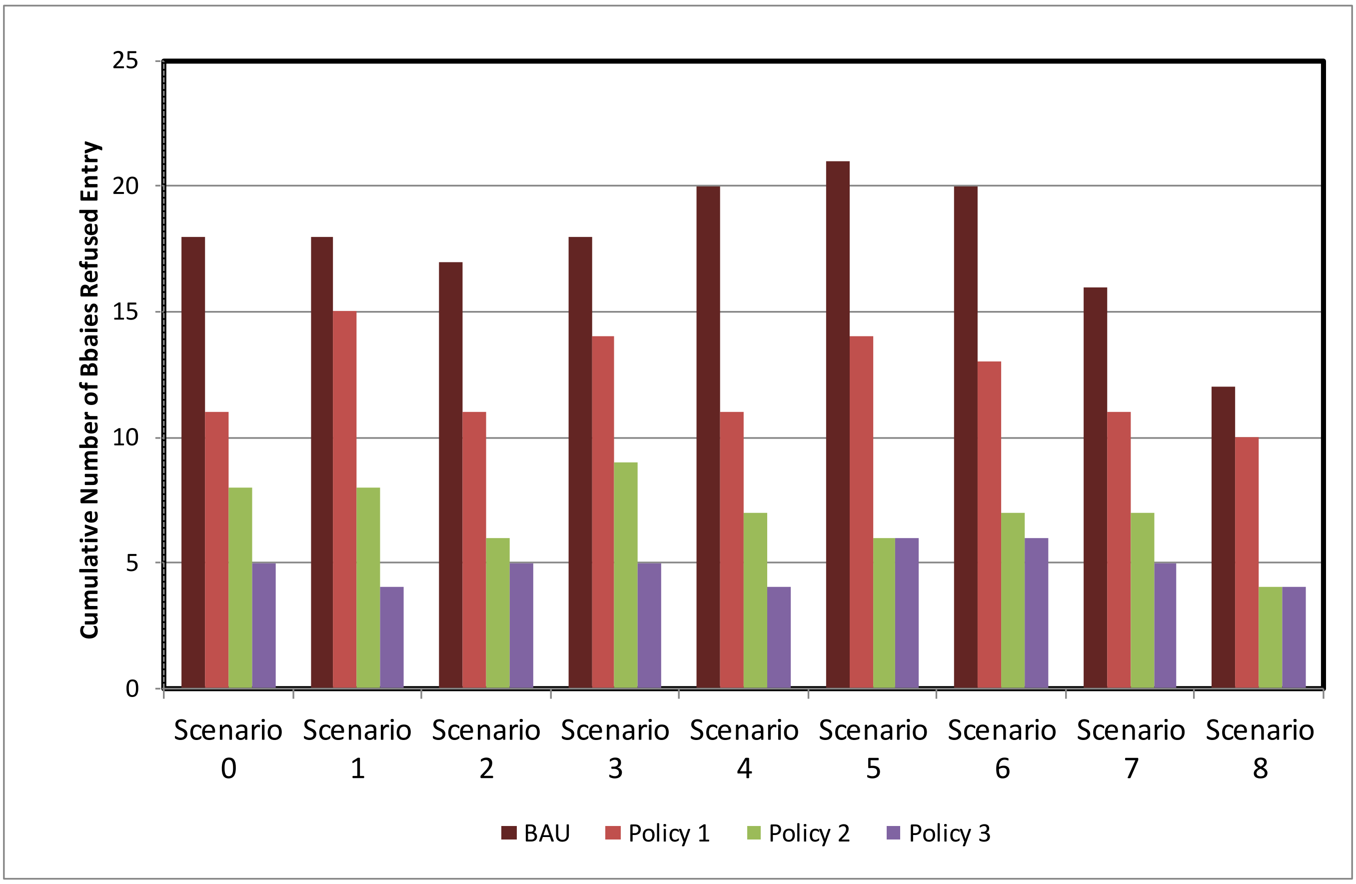


Figure 6: Cumulative number of babies refused entry for policies involving the use and non-use of clinical judgment



BAU: Business as Usual (No Clinical Judgment)