



Towards generic modelling of hospital wards: Reuse and redevelopment of simple models

M.L. Penn, T. Monks, A.A. Kazmierska & M.R.A.R. Alkoheji

To cite this article: M.L. Penn, T. Monks, A.A. Kazmierska & M.R.A.R. Alkoheji (2019): Towards generic modelling of hospital wards: Reuse and redevelopment of simple models, Journal of Simulation, DOI: [10.1080/17477778.2019.1664264](https://doi.org/10.1080/17477778.2019.1664264)

To link to this article: <https://doi.org/10.1080/17477778.2019.1664264>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 23 Sep 2019.



Submit your article to this journal [↗](#)



Article views: 185



View related articles [↗](#)



View Crossmark data [↗](#)

Towards generic modelling of hospital wards: Reuse and redevelopment of simple models

M.L. Penn^a, T. Monks^b, A.A. Kazmierska^a and M.R.A.R. Alkoheji^a

^aSouthampton Business School, University of Southampton, Southampton, UK; ^bUniversity of Exeter Medical School, University of Exeter, Exeter, UK

ABSTRACT

Generic simulation models are designed to enable model reuse. We argue that there are two weaknesses within the generic simulation modelling literature. Firstly, that generic models sacrifice the relative simplicity of a bespoke simulation model for flexibility. Secondly, that generic models are published in conceptual form only. If researchers cannot access computer implementation of models, then there is little incentive or benefit to recode one over coding a simpler bespoke simulation model. We introduce an incremental approach to generic modelling in discrete-event simulation. We develop an archetype setting-specific generic model of a hospital ward. The archetype model is first developed and applied in a rehabilitation ward setting. Then a second team applies the model in a specialised intensive care setting. We report the successes, obstacles and redevelopment needed for reuse of the generic model along with how the results of these studies were used to inform healthcare delivery.

ARTICLE HISTORY

Received 22 October 2018
Revised 3 June 2019
Accepted 30 June 2019

KEYWORDS

Generic modelling; model reuse; discrete event simulation; model redevelopment

1. Introduction

The development, verification and validation of discrete-event simulation (DES) models are time-consuming and expensive. The *reuse of existing models* is often pointed to as a method to reduce this upfront cost (Kaylain et al., 2008; Robinson, Nance, Paul, Pidd, & Taylor, 2004). Model reuse can occur at different levels within a DES study: from the reuse of a conceptual model (Balci & Nance, 2008; Monks et al., 2017), to simulation components (Pidd & Carvalho, 2006), to an entire coded model (Robinson et al., 2004) to re-applying insights from abstract queuing models (Fletcher & Worthington, 2009). We adopted Fletcher and Worthington (2009) the definition of a *setting-specific generic model* to describe a DES model designed for reuse in a general hospital ward. There are several examples of setting-specific generic models in the health-care simulation literature (e.g., Di Mascolo & Gouin, 2013; Fletcher, Halsall, Huxham, & Worthington, 2007; Günal & Pidd, 2011; Sinreich & Marmor, 2004; Weerawat, Pichitlamken, & Subsombat, 2013). We argue that there are two weaknesses of the DES generic modelling literature. First, the models are often complex relative to a bespoke simulation model due to the need to be able to reconfigure the model to run in more than one setting. Second, the computer implementation of the generic models is unavailable. Complex models usually require more data, are harder to understand and can have a long run time.

Our approach advocates beginning with a simple archetypal generic model of a process and only adding further detail if needed. Application in multiple-settings can expose the weaknesses in the design of a generic model and the adaptations that are needed (Robinson et al., 2004). The requirement to recode the, possibly complex, generic model reduces the likelihood of opportunistic reuse and testing of a model by new modelling teams substantially (Monks & Meskarian, 2017). Our approach is to make computer implementations of the model findable, accessible and citable.

This study aims to address the re-application and re-development gaps in the DES generic modelling literature. We report two generic modelling studies in healthcare. The applied examples in question were two real and sequential studies about the configuration of hospital wards that were commissioned by the NHS (National Health Service) in the UK. The first model was designed by the first two authors who are experienced simulation modellers. The context is the transfer of patients from an acute hospital setting to a community hospital for rehabilitation. The second modelling study was executed by the last two authors who were novice simulation modellers and at the time were undertaking a master's degree in Operational Research. The second study tested the generic properties of the first model in a more specialised (intensive care) ward and identified additional flexibility needed in order to successfully reuse it. As both studies were

informing real decision-making, our work also contributes to the growing, but still limited, evidence about the impact of computer simulation in health (Brailsford & Vissers, 2011; Fone et al., 2003; Günal & Pidd, 2010; Monks et al., 2018; Crowe, Turner, Utley, & Fulop, 2017). In each case, we describe the use of the models and their results in practice and the challenges faced in reuse.

2. Study aims

Our study had five aims:

- To develop a base generic simulation model of a hospital general ward;
- To test the model in both general and highly specialised hospital wards;
- To identify adaptations needed to reuse the model in a more specialised setting;
- To identify challenges, obstacles and benefits of developing the generic models;
- To document the use or lack of use of these generic models in practice.

This paper begins with a description of the model development methods and then reports the two models in detail, before giving details of the case studies, including sample results and a narrative of the use of the models in practice. It then concludes with the implications of this work and suggestions for further work.

3. Methods

3.1. The general approach to model design

We grounded model design in real capacity planning problems faced in the NHS. We followed standard “include/exclude” approaches to set the model level of detail from the formal literature on conceptual modelling (Robinson, 2008a, 2008b). Overarching this process, were three generic modelling requirements that we based on generic hospital modelling literature (Güenal, 2012; Güenal & Pidd, 2011).

- The principle of keeping the model as *simple* as possible, to facilitate understanding and, where necessary, incremental adaptation, whilst providing sufficiently accurate results to convince decision-makers to use the results to support ward capacity planning;

- Incorporating enough *flexibility* in the model configuration to allow the model to be reused, for its intended purpose, across multiple hospitals;

- Ensuring the model was *intuitive* for users with limited simulation or programming training and for future adaptations by other simulation modellers.

We note that these requirements are sometimes contradictory. Too much flexibility would come at the expense of both intuitive usability and simplicity and was likely to result in redundant functionality or “feature bloat”. Our approach, therefore, settled on creating an *archetype ward model* that could be used independently or easily adapted for more specialised wards. We believed our base model to be reusable across many general medical and surgical wards in different types of hospital. For example, it could be used to model a respiratory or rehabilitation wards or an acute medical unit. The base purpose of the model was to assist the planning of the number and configuration of beds required to minimise admission delays for heterogeneous patient populations both now and in the future. We identified the components of the problem that varied between wards (flexibility requirements) through discussions with our NHS collaborators, literature review and through our own experience of modelling to assist the NHS with ward capacity planning. We selected what we felt was a simple to learn, use and relatively inexpensive commercial simulation package for model implementation. We also limited the extent to which the user interacts with the simulation software by using an Excel interface.

A second team of modellers, supervised by one of the original team, reused the archetype model in a more specialist context: an intensive care unit. Where necessary we incrementally added to the level of detail and flexibility of the base model to produce sufficiently accurate results to support decision-making.

3.1.1. Access to the generic models

To facilitate reuse and refinement of the models we developed, we adopt an open science approach to this study (Taylor et al., 2017). All model code is findable, accessible, reusable and citable, via Zenodo (<https://zenodo.org/>), along with test data for model verification. The coded models are implemented in Simul8 Professional 2018. We document the model using the STRESS-DES reporting guidelines (Monks et al., 2018). The full STRESS-DES documentation is available in the online appendix.

3.2. Archetype ward model

3.2.1. Motivation and objectives

A common problem for managers in hospitals is planning ward bed capacity. Too few beds and there is a problem with *patient flow*: patients will experience significant admission delays to specialist care or may not be able to transfer from one ward to another either in the same or a different hospital. In extreme circumstances, patient may be admitted to wards

outside of their specialist needs. Too many beds and continuous patient flow comes at the expense of a hospital's need for value-for-money.

This is a classic queuing problem with a waiting time versus value-for-money trade-off. To support health-care planners, the generic ward model outputs the long-run average and distributions for admission waiting time and ward occupancy. Waiting metrics consist of the number of patients waiting for admission (queue length) and waiting time for those unable to enter the ward straight away and where relevant the number of patients who wait so long that they are treated elsewhere and never enter the ward. Ward occupancy metrics represent the utilisation of beds in the simulation. The model outputs both the average and maximum ward occupancy. The model also outputs a third type of output measure: transfers between beds within the ward. This was included due to single-sex accommodation requirements we discuss below.

3.2.2. Level of detail

To build a generic model for a hospital ward, we need to consider the features of a typical ward, as well as the variation in the arrival rate and length of stay. There is a requirement, introduced in 2011 (Department of Health and Social Care, 2011), that in the UK “all hospital accommodation is same-sex” (Care Quality Commission, 2015). This is also a requirement in other countries, for example, in South Africa (Bloem, 2015). This means that the division of the ward into a combination of same-sex multi-bed bays and single rooms is an important consideration. If only single rooms are used, then they will automatically be same-sex and it is only the capacity of the ward that needs to be considered. In general, it is more expensive to build and staff a ward of entirely single rooms (Spesvytseva, n.d.). There is also mixed evidence as to whether single rooms are better or worse environments for patients. Study findings range from reporting considerable patient benefits of single rooms (Wales NUS Estates, 2005), no clear effect of single rooms on safety

(Simon, Maben, Murrells, & Griffiths, 2016) and in some settings sharing a bay with other patients can actually assist patient recovery (Department of Health, 2008). Isles (2013) argues that a mix of single rooms and multi-bed bays facilitates infection control where needed, while allowing those who prefer having company to do so. Given this uncertainty, a generic simulation model needs to allow users to input different ward configurations, in terms of the number of single beds and the size of any bays used. The model must also take account of the same-sex requirement when assessing if there is space for the patient. Should this not be required the data can be entered as if all patients are of one gender.

3.2.3. Model logic

The logic of the model is shown in Figure 1. Patients arrive into the model at the point when the decision is made to admit them to the ward being modelled, the inter-arrival times are sampled from an empirical distribution. At this point, a patient is assigned a gender and length of stay sampled from the corresponding empirical distribution.

Patients then join a queue of individuals waiting for a space in the ward. In practice, this is a virtual as opposed to a physical queue. It is the list of patients waiting elsewhere in the hospital for transfer to the ward. From the queue, they move to a decision point, which attempts to allocate them to a bed on the ward. If a suitable space in the ward is found they move to a bed in the ward.

If there is no suitable space in the ward, the patient will be assumed to stay elsewhere in the hospital and be waiting for a space in the ward. A short waiting time is allocated before an admission is reattempted. Patients move from the queue in ‘first in first out’ order if more than one arrives or reaches the end of their short wait at the same time. For patients who wait for a space in the ward a proportion of the time that they wait may be removed from the length of stay (this is user-defined). It is also possible for the user to specify when patients have waited so long that they no longer require the ward (they have been

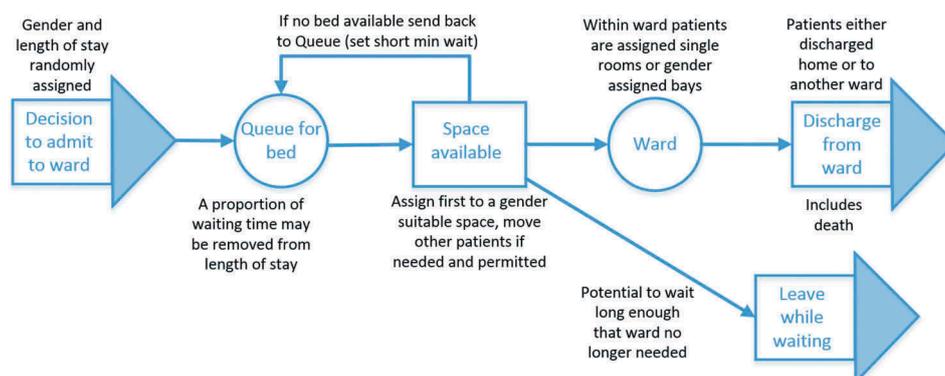


Figure 1. The structure of the single ward model.

treated elsewhere). When this happens patients will be sent out of the system.

Patients then remain on the ward for their sampled length of stay before being discharged from the ward.

3.2.4. Model inputs

The model includes two controllable ward design variables: the *number of beds* in a hospital ward and how they are divided into a mixture of *single rooms and multi-bedbays*.

The variation in the arrival of patients to the ward and their subsequent length of stay on a ward is handled through sampling from empirical distribution functions. The model allows multiple patient types. Each of which can be assigned its own length of stay distribution. This for example, could be age bands within the population or patients with different care needs. Health-care planners, therefore, have the flexibility to use the model with data of differing levels of detail. In cases where data are highly limited, planners can use the model with a single homogeneous patient population. However, if more detailed data are available, planners can plan across multiple populations by varying arrival and length stay distributions based on predictions of future demand.

3.2.5. Simul8 implementation

In a generic model the flexibility to model different systems should come from making changes to the input data (Di Mascolo & Gouin, 2013; Fletcher & Worthington, 2009). For our model the model configuration and input data for the ward (discussed above) are specified in a separate Microsoft Excel 2016 spreadsheet. Our target audience are individuals who work in health-care service delivery, they are likely to be familiar with Excel and therefore be more comfortable with entering data in a familiar environment.

Within the spreadsheet patients can be divided into up to five different groups, these could represent different types of condition, age groups, or any other division that effects either the arrival rate or the amount of time required on the ward (length of stay), for the particular ward being considered. In some cases a clinically relevant division into groups will be apparent, for others, statistical techniques such as classification and regression tree (CART) analysis could be used, an example of how this can work in a health-care setting can be found in Harper et al. (2003). The arrival rates and empirical data are all entered via the user interface. The interface allows users to enter raw numbers from their data and then transforms these into the format required by the model. This is all intended to increase the user-friendliness of the model for those already familiar with Excel and allow adaptations to allow for some differences between wards without requiring

a simulation expert to make changes to the model. The interface also includes features to assist the user in setting up scenarios involving changes to patients' lengths of stay and/or demand for the ward.

The data contained in the interface is loaded into a model in Simul8 Professional 2018 (<https://www.simul8.com/>), which is discrete-event simulation specialist software. The arrows containing text in Figure 1 represent the entry and exit points of the model, the circles the queues and the rectangle the decision point. The actions discussed on the diagram are generated using a combination of Simul8's labels (patient attributes) and, its bespoke programming language, visual logic. The ward is modelled as a queue where patients "wait" to be discharged, the number of patients in the ward is controlled by keeping track of the number of patients in each bay/in single beds through the implementation of a spreadsheet to which they are added when admitted and removed from when discharged.

As any multi-bed bays must be single sex the algorithm below is used to assign each patient to a bay within the ward:

- (1) If there are spaces in at least one bay containing patients of the same gender as the patient to be assigned, then assign the patient to the one with the highest proportion of its beds occupied. This is to allow the possibility of the least full bays becoming empty and potentially switching gender. Bay index number is used as a tie break if required.
- (2) If the patient was not assigned a bed in step 1 and there are any empty bays assign the patient to the bay with the highest index number.
- (3) If the patient was not assigned a bed in step 1 or 2 and there are any spaces in single rooms available assign the patient to a single room.
- (4) If the patient was not assigned a bed in step 1, 2 or 3 and there is a patient of the opposite gender in a single room and it is possible to move that patient to an appropriately gendered bay move that patient and assign the original patient to the single room.
- (5) Else assign a short waiting time to the patient and send them back to the queue.

The user interface, Simul8 model and a brief user guide are available from DOI 10.5281/zenodo.1468287 (Penn & Monks, 2018). The full STRESS-DES documentation for the model is contained in the online appendix.

The model was validated using the applied example set out below by comparing the outputs when run with current admissions with the results for the current wards. Detailed verification of the logic was conducted by creating data sets that would require the full range of

outcomes and running through the code step by step. It was further validated through demonstrations of the model and results to our contacts at the hospital.

3.3. Redevelopment of specialised version: Intensive care unit model

3.3.1. Motivation

We went on to test and reuse the archetype model in a more specialised ward environment: capacity planning in intensive care units (ICU). ICU capacity planning requires similar experimentation functionality to that provided in the archetype ward model. However, there are differences in how patients are treated whilst waiting for a space and the lengths of stay are considerably shorter. This section will explain how the model was redeveloped to address these differences.

ICUs are where the most critically ill patients are cared for, the equipment required is complex and expensive and high staffing ratios are required. The annual running cost of an ICU bed in the UK is approximately £361,000 (Griffiths, Jones, Read, & Williams, 2010). It is important to balance the need to have enough ICU beds to avoid delays in caring for critically ill patients, with concern that running ICU beds that are not used would be poor use of limited resources (Zhu, Hoon, & Teow, 2012). The quality of intensive care has been the subject of national focus due to the consequences of inadequate capacity, particularly for elective surgery cancellations and emergency patient transfers (Costa, Shahani, & Harper, 2003).

As for the archetype model, there is variability in both the arrival rates and lengths of stay of patients, with concern about reducing the extent to which patients stay longer than medically required. The significant differences are that ICUs are exempt from the same-sex accommodation rule, and that ICU patients can only wait limited time for a bed to become available before they must be transferred elsewhere or have their operation cancelled.

In an ICU the groups into which patients naturally divide are based on the level of care they require and whether they will require more than one level of care during their stay.

3.3.2. Adaptions to the conceptual model

We have started with the generic ward model discussed above and redeveloped it to become a generic ICU model. We implement a user interface based on that discussed above, but with more data required for the additional features of the model. Figure 2 illustrates the conceptual model. The basic flow model is identical apart from the replacement of “ward” with “ICU”. However, the comments about what is happening at each stage are different; the differences are underlined for emphasis.

Table 1 provides a comparison of the two models, providing full details of how the original model has been adapted.

Patients will go through the queuing loop until the maximum waiting time for their patient type. At the end of this time, they will be transferred elsewhere. In reality, patients would be transferred earlier

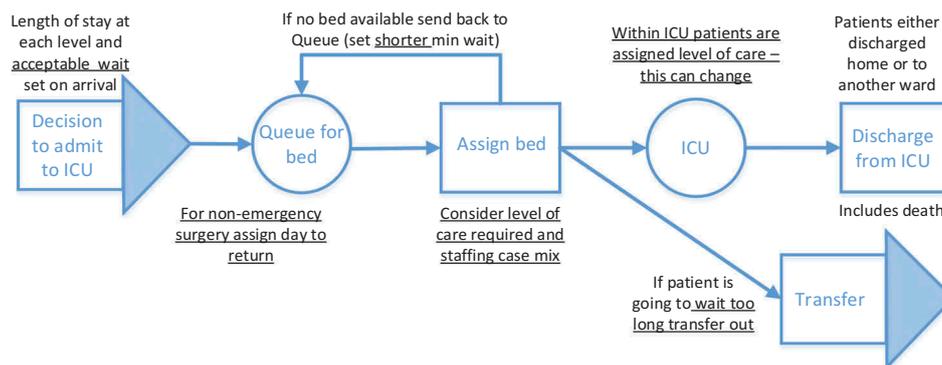


Figure 2. The model structure adaptations for ICU.

Table 1. Summary of differences between the models.

Feature	Generic ward model	Generic ICU model
Time represented	Daily	Hourly – as the lengths of stay are shorter and waiting time more crucial
Arrival process	Single distribution	Time-dependent; as fewer patients arrive at night
Waiting times	All patients loop until they enter ward or no longer need it	Different max waiting times for different groups assigned on entering the system
Time between admission attempts	The same for all patients	Short for emergencies, several days for surgical patients to allow for rebooking.
Bed types	Bays to be single gender	Only level of care considered, with limit added on number with highest care level
Input data	Divided by age and gender	Divided by level of care required

if it seemed unlikely that a bed would become available, however, the important result for decision-making is the number transferred. We, therefore, do not model the transfer process in detail.

3.3.3. Simul8 implementation

The ICU model is an incremental adaption of the archetype model. As for the archetype ward model the data input is using empirical distributions based on hospital data, along with settings to control experimentation. All of these are entered via an Excel interface. They are then read into the adapted version of the original Simul8 model.

This version of the Simul8 model does not need to include the same gender assignment algorithm given in section 3.2.5, however the routing back of those patients who cannot be assigned beds in the ward straight away is more complex. The model was verified using the applied example given in Section 4.2, it was validated by comparison with the data analysis and the opinion of those working in the system.

Both the Excel user interface and Simul8 model are available online at DOI 10.5281/zenodo.1468314 (Penn, Monks, Kazmierska, & Alkoheji, 2018) the full STRESS-DES documentation is also included in the online appendix.

4. Applied examples

4.1. Designing a new community rehabilitation ward

4.1.1. Background

An NHS clinical commissioning group (CCG) and community trust approached us for analytical support for the business case for combining two rehabilitation wards. The two wards were part of a community trust and cared for patients that were transferred from a separate large acute hospital. There were substantial waiting times (in the UK, delayed transfers of care) for these wards and the creation of a new-combined ward provided an opportunity to review capacity requirements. We worked with the local commissioning manager and the nurse responsible for coordinating the rehabilitation beds. The former was responsible for the business case overall and was able to provide the required data, while the latter works in the existing wards and was therefore invaluable in validating the model.

4.1.2. Data sources

Anonymised data from the two existing wards including empirical distributions for the arrival rate of patients and their lengths of stay, by age group and gender have been brought together to create a patient profile for the new ward. A distribution for the lengths of stay with that part of the stay considered

“excess bed days” removed has been set-up, to allow scenarios considering removing part of this aspect of patients stay to be created. Public data on population projections, by age and gender, allows us to consider the potential increase in demand as the numbers using the ward increases.

4.1.3. Illustrative model results

In the study the archetype model was used for large-scale search experimentation (see Hoad, Monks, & O’Brien, 2017). The solution space for total numbers of beds and configurations of those beds into bays, proportions of the population change by age group affecting demand and reducing the excess bed days was explored. In total, we conducted 280 simulation experiments.

Table 2 is an example of the results for different bed configurations, with the model run for 1 year (approximately 650 arrivals) following a 1 year warm-up period for 100 replications. The number of replications was selected so that the half width of the confidence intervals for all performance measures would be less than 5% of the mean using Simul8’s inbuilt facility for selecting the number of replications, which is based on research by Hoad, Robinson, and Davis (2010). This set of results shows the impact of changing the mix of single rooms and bays of four beds. As the number of single beds increases to around 10 the average waiting time and queue length for entering the ward go down, but the number of transfers occurring increases. As the number of single beds increases beyond 14 the number of transfers drops, as single rooms create greater flexibility in finding suitable accommodation for patients.

Similar sets of scenarios were run with different bay sizes and also with possible changes to the case mix. The later were developed based on input from the charge nurse responsible for the current wards, based on groups of patients who cannot usually be accommodated in the wards but who would be expected to benefit from such care.

Table 2. Sample results for a range of 50-bed configurations.

Experiment Input			Experiment Results			
Size of Bays	Number of Bays	Number of Singles	Avg. Wait (days)	Avg. Queue	Avg. % Occupancy	Number of Transfers
0	0	50	0.6	1	87.5	0
4	12	2	2.4	5	87.4	48
4	11	6	1.5	3	87.5	97
4	10	10	1.0	2	87.5	117
4	9	14	0.7	1	87.5	123
4	8	18	0.6	1	87.5	115
4	7	22	0.6	1	87.5	101
4	6	26	0.6	1	87.5	87
4	5	30	0.6	1	87.5	71

4.1.4. Use of the results in practice

The results of the model demonstrated that creating a new ward with the combined bed numbers of the two original wards would not be sustainable. This is due to the increasing elderly population in the region, which will increase demand. The NHS used the model results as the basis for a £16m (\$21m; €18m) business case development of the wards. The CCG have estimated that by 2022 the saving to the local health economy of the redesigned wards at £3.2m per annum.

The search experimentation demonstrated that there were multiple “optimal” splits of single and multi-bay options. The model suggested that similar waiting times and numbers of transfers could be achieved with fewer single beds and larger bays, than is being put forward in the business case. The NHS opted for the same proportion of single beds as the existing wards and with bays of three beds, as they know that this works well from a clinical perspective.

We note that the reporting of the results from the simple model was not as straightforward as we expected. After the results of the study were reported we were contacted by a senior manager at the community hospital with a query over the model’s validity. The issue related to a forecasting component of the study where we examined queue lengths given different projections of the adult population. Our analysis demonstrated the classic queuing versus traffic intensity trade-off: in scenarios where ward occupancy was very high, small increases in the number of beds led to large reductions in queues. However, the mental model of the health-care manager (and his team) led to the expectation that if you add five more beds to a ward then the queue for the ward would reduce by five patients. This is the so-called fallacy of planning capacity by average and failing to account for stochasticity. It is also worth noting that it was not a straightforward process to correct this misperception. We conducted a series of educational interventions including demonstrations of even simpler simulation models with and without stochastic behaviour; detailed written explanations and additional analysis of the model to generate probabilities of different queues lengths.

4.2. The effect of delayed discharges on ICU capacity requirements

4.2.1. Background

A regional adult critical care network approached us regarding the possibility of creating a generic model that could be used to explore changes to the number of beds, lengths of stay or arrival rates for any of its ICUs. For this modelling, we worked predominantly with the network manager, but ICU visits were also

conducted, and sample results were presented to a selection of those working across the network.

4.2.2. Data sources

Anonymised data from all the ICUs in the network have been used to generate empirical distributions for the arrival rate of patients and their lengths of stay, for each of the levels of need. A distribution for the lengths of stay with that part of the duration considered “discharge delays” removed has been set-up, to allow scenarios considering removing part of this aspect of patients stay to be created. Public data on population projections, by age and gender, allows us to consider the potential increase in demand as the numbers using the ward increases.

4.2.3. Illustrative model results

The specialised model was again used for search experimentation (see Hoard et al., 2017). Factors considered were, reducing discharge delays, increasing population and changes to bed numbers. Table 3 provides a sample of the results when considering removing the part of a patient’s stay that is considered a “delayed discharge” for one of the 18 hospitals in the network.

This illustrates how reducing the delay in discharging patients could allow the number of beds to be reduced. Given the significant cost of running a single ICU bed, this allows comparison of the costs and benefits of reducing delayed discharges.

The decision to consider removing delays over 24 h and over 4 h is based on these being important cut-offs in the data recording. The adjustments are made in the data processing so other time periods could be considered without any changes to the model.

These results are a small sample from the scenarios explored, which included consideration of a range of numbers of beds available for each of the periods of delay being removed. Removing different proportions of the delays were also considered, giving an indication of the impact of different levels of success in reducing delayed discharges. Comparing these results for different hospitals in the network can assist in identifying where to target resources to make the biggest improvements.

Table 3. Sample results for reducing delayed discharges.

	Base	Remove delays over 24 hours	Remove delays over 4 hours
Number of beds available	25	24	23
Occupancy of unit (average percentage)	76.4%	76.7%	76.7%
Emergency Transfers (average number)	30	31	31
Electives cancelled (average percentage)	20.7%	20.1%	21.2%

4.2.4. Use of the results in practice

The original analysis considered two units from the ICU network. The network manager is continuing to use the model to conduct similar analysis for the other units and is using the results in ongoing discussions about reducing the delay in discharging patients.

Other ICU networks are aware of the model and are considering using it with their data. The network manager for whom this analysis was originally undertaken has a copy of Simul8 Professional and the Excel interface is allowing him to use the model with very limited support. However, lack of access to Simul8 is limiting the ability of other networks to access the model.

5. Discussion

5.1. Summary

Our study has designed an archetype setting-specific generic model of a hospital ward that can be used in capacity planning. We grounded this model in a real and pressing problem facing health-care planners in the UK. Although simple, the model was highly effective in supporting decisions and has real tangible benefits for the NHS. We tested the limits of the archetype model and its ease of adaptation through a second team of (novice) modellers who reused the original model in a more specialised ICU setting. We found that most of the flexibility offered by the archetype was needed when reused in the ICU setting. However, there were some key differences. The most prominent was the need to include a more detailed arrival and admission processes.

Redeveloping the archetype model rather than starting from scratch, saved time particularly in terms of planning how to operate a user interface to allow scenarios to be generated outside of the Simul8 software. It has also reduced the time taken to validate the model, as part of this process had already been undertaken. As the second model was worked on by a different team, there was some time required for understanding the original model and how it needed to be adapted to the new problem. This process was assisted by documentation of the first model and access to the original modeller when required. Thus, demonstrating that modifying a sufficiently documented generic model can have significant advantages over creating a bespoke model. In deciding whether to modify a generic model or create a bespoke model the extent of the differences between the modelling requirements and therefore the changes required will remain a significant consideration.

The second model could in itself be considered a generic model for an ICU setting and be reused for similar problems, rather than going back to the more basic model.

5.2. Implications for generic modelling

We found that reuse of our original model led to new learning about what flexibility might be incorporated into an archetype model. The adaptations that the second team of modellers made could arguably have been incorporated into the original model's flexibility. For example, the model could accept inputs for multiple time-dependent arrival distributions. We justified our decision for a single arrival input to navigate the simplicity and flexibility trade-off (Günal, 2012). Our aim was to allow novice simulation users, perhaps with limited data, to more easily reuse the model and reduce the potential for mistakes in its use. We also note that to reduce the complexity of the ICU model to only that which was needed, the ability to consider gender has been removed from it, so future users will need to select which of the models most closely fits the problem that they are considering.

Fletcher and Worthington (2009) argued that users of generic models should not be required to invest in specific software. There are several popular open-source DES tools and programming languages now available (Dagkakis & Heavey, 2016). For example, there are Python-based open source DES tools available Ciw (Palmer, Knight, Harper, & Hawa, 2018) and SimPy (Team SimPy, 2018). However, the current drawback of open science simulation tools is that setup and deployment of models requires the right programming and simulation expertise (Dagkakis & Heavey, 2016). Although not insurmountable, at this time this poses a significant barrier for an organisation such as the NHS to quickly pick up and use the models. Our computer implementation of the models reported here use Simul8 Professional: a relatively low-cost commercial simulation package with a strong ethos of usability. Our decision was based on our positive experience of working with novice simulation users, both students and within the NHS, using the software. We demonstrated that it was successful when reused with the regional critical care network (who held a Simul8 license), but spread to the national network, who would need to invest in licenses, now seems unlikely in a financially struggling NHS.

5.3. Evidence of simulation improving healthcare planning

Although a full implementation study was out of the scope of our work, we can point to several contributions that our work has made to the limited evidence of simulation improving health service delivery. Firstly, the enthusiastic adoption of the model by health-care planners and care providers in the region. Second, the predicted multi-million-pound

saving to the regions health economy over the next 5 years. Lastly, the education we provided to the leaders of these health systems about the dynamics of queuing systems.

The educational impact of the model was not an original aim of the study. We, therefore, did not have a substantive educational intervention to hand such as that from SimLean (Robinson, Radnor, Burgess, & Worthington, 2012); however, we did have simple queuing models that could be deployed from working in acute hospitals on similar problems (Monks & Meskarian, 2017). It is interesting to note that the queuing education we provided was to explain the results of the (simple) archetype simulation model. Our experience agrees with rigorous implementation research conducted outside a simulation context, but still within an Operational Research study (Crowe et al., 2017); namely that to be implemented results need to be accessible to policymakers and practitioners. Given the reuse emphasis of such studies, this may suggest that the design of generic models should include a significant education package.

5.4. Future work

The open approach of our work means that both models are freely available. They can be accessed by researchers or health-care planners who wish to use them with different data, or to redevelop them for other similar problems. We will keep track of the number of times that they are accessed and would be very interested in hearing from anyone who is making use of them, in full or as inspiration for new models.

We have discussed the limitations of software availability for the use of simulation models in the NHS. We are actively exploring open-source simulation software to increase the flexibility of our future generic models and encourage other modellers to consider this issue. There is scope to explore the use of generic modelling in a wide range of applications. While our research has focused on the healthcare context, there is no reason why such generic modelling should not bring similar benefits in a range other domains. We hope our incremental approach to development, reuse of our own model, and the publication of our coded models, will assist others in creating and implementing generic models in other areas of simulation.

As the creation and publication of generic models increases it would be beneficial to develop a metric to define the level of generality of models in a standardised manner. This might also include standardisation of documentation of models in a similar way to the STRESS-DES guidelines (Monks et al., 2018), specifically focussed on the scope of the model to adapt to different system features and performance measures.

Acknowledgments

This article presents independent research funded by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) Wessex. The views expressed in this publication are those of the author(s) and not necessarily those of the National Health Service, the NIHR, University of Southampton or the Department of Health and Social Care.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

M.L. Penn  <http://orcid.org/0000-0002-7269-7981>

T. Monks  <http://orcid.org/0000-0003-2631-4481>

References

- Balci, O., & Nance, N. E. (2008). Accomplishing reuse with a simulation conceptual model. *Proceedings of the 2008 Winter Simulation Conference* (pp. 959–965). Piscataway, New Jersey.
- Bloem, C. (2015). *Improving hospital bed utilisation through simulation and optimisation in South African Hospitals*. Pretoria: University of Pretoria.
- Brailsford, S., & Vissers, J. (2011). OR in healthcare: A European perspective. *European Journal of Operational Research*, 212(2), 223–234.
- Care Quality Commission. (2015). *CQC Brief guide: Assessment of same-sex accommodation*. Newcastle: Author.
- Costa, A. R. S., Shahani, A., & Harper, P. D. S. V. (2003). Mathematical modelling and simulation for planning critical care capacity. *Anaesthesia*, 58(4), 320–327.
- Crowe, S., Turner, S., Utley, M., & Fulop, N. J. (2017). Improving the production of applied health research findings: Insights from a qualitative study of operations research. *Implementation Science*, 12(1), 112.
- Dagkakis, G., & Heavey, C. (2016). A review of open source discrete event simulation software for operations research. *Journal of Simulation*, 10(3), 193–206.
- Department of Health. (2008). *Ward layouts with single rooms and space for flexibility*. Norwich: The Stationary Office.
- Department of Health and Social Care (2011). Letter: Eliminating mixed sex accommodation – Declaration exercise. London: Department of Health and Social Care. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/215773/dh_124233.pdf
- Di Mascolo, M., & Gouin, A. (2013). A generic simulation model to assess the performance of sterilization services in health establishments. *Health Care Management Science*, 16(1), 45–61.
- Fletcher, A., Halsall, D., Huxham, S., & Worthington, D. (2007). The DH accident and emergency department model: A national generic model used locally. *Journal of the Operational Research Society*, 58(12), 1554–1562.
- Fletcher, A., & Worthington, D. (2009). What is a 'generic' hospital model? - a comparison of 'generic' and

- 'specific' hospital models of emergency patient flows. *Healthcare Management Science*, 12, 374–391.
- Fone, D., Hollinghurst, S., Temple, M., Round, A., Lester, N., Weightman, A., ... Palmer, S. (2003). Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *Journal of Public Health Medicine*, 25(4), 325–335.
- Griffiths, J. D., Jones, M., Read, M. S., & Williams, J. E. (2010). A simulation model of bed-occupancy in a critical care unit. *Journal of Simulation*, 4(1), 52–59.
- Günal, M., & Pidd, M. (2011). DGHPsim: Generic simulation of hospital performance. *Transactions on Modeling and Simulation (TOMACS)*, 21(4). doi:10.1145/2000494.2000496
- Günal, M. M. (2012). A guide for building hospital simulation models. *Health Systems*, 1(1), 17–25.
- Günal, M. M., & Pidd, M. (2010). Discrete event simulation for performance modelling in health care: A review of the literature. *Journal of Simulation*, 4(1), 42–51.
- Harper, P., Sayyad, M. G., de Senna, V., Shahani, A. K., Yajnik, C. S., & Shelgikar, K. M. (2003). A systems modelling approach for the prevention and treatment of diabetic retinopathy. *European Journal of Operational Research*, 150(1), 81–91.
- Hoad, K., Monks, T., & O'Brien, F. (2017). The use of search experimentation in discrete-event simulation practice. *Journal of the Operational Research Society*, 66(7), 1155–1168.
- Hoad, K., Robinson, S., & Davis, R. (2010). Automated selection of the number of replications for a discrete-event simulation. *Journal of the Operational Research Society*, 61(11), 1632–1644.
- Isles, C. (2013). No. In: H. Pennington & C. Isles. Should hospitals provide all patients with single rooms? *British Medical Journal*. doi:10.1136/bmj.f5695.
- Kaylain, A., Mollaghesemi, M., Cope, D., Favez, S., Rabadi, G., & Steele, M. (2008). A generic environment for modelling future launch operations –GEM-FLO: A success story in generic modelling. *Journal of the Operational Research Society*, 59(10), 1312–1320.
- Monks, T., Currie, C. S. M., Onggo, B. S., Robinson, S., Kunc, M., & Taylor, S. J. E. (2018). Strengthening the reporting of empirical simulation studies: Introducing the STRESS guidelines. *Journal of Simulation*, 1–13. doi:10.1080/17477778.2018.1442155
- Monks, T., & Meskarian, R. (2017, December 3–6). *Using simulation to help hospitals reduce emergency department waiting times: Examples and impact*. Paper presented at the 2017 Winter Simulation Conference (WSC), Las Vegas.
- Monks, T., van der Zee, D. J., Lahr, M. M. H., Allen, M., Pearn, K., James, M. A., & Luijckx, G.-J. (2017). A framework to accelerate simulation studies of hyperacute stroke systems. *Operations Research for Health Care*, 15, 57–67.
- Palmer, G. I., Knight, V. A., Harper, P. R., & Hawa, A. L. (2018). Ciw: An open-source discrete event simulation library. *Journal of Simulation*, 1–15. doi:10.1080/17477778.2018.1473909
- Penn, M., Monks, T., Kazmierska, A., & Alkoheji, M. R. A. R. (2018, October 22). ICU capacity model (Version v1). *Zenodo*. doi:10.5281/zenodo.1468315
- Penn, M. L., & Monks, T. (2018, October 21). Generic ward configuration simulation model (Version v1). *Zenodo*. doi:10.5281/zenodo.1468288
- Pidd, M., & Carvalho, A. (2006). Simulation software: Not the same yesterday, today or forever. *Journal of Simulation*, 1(1), 7–20.
- Robinson, S. (2008a). Conceptual modelling for simulation Part I: Definitions and requirements. *Journal of the Operational Research Society*, 59(3), 278–293.
- Robinson, S. (2008b). Conceptual modelling for simulation Part II: A Framework for conceptual modelling. *The Journal of the Operational Research Society*, 59(3), 291–304.
- Robinson, S., Nance, R., . E., Paul, R. J., Pidd, M., & Taylor, S. J. E. (2004). Simulation model reuse: Definitions, benefits and obstacles. *Simulation Modelling Practice and Theory*, 12(7–8), Simulation model reuse: definitions, benefits and obstacles, 479–494.
- Robinson, S., Radnor, Z. J., Burgess, N., & Worthington, C. (2012). SimLean: Utilising simulation in the implementation of lean in healthcare. *European Journal of Operational Research*, 219(1), 188–197.
- Simon, M., Maben, J., Murrells, T., & Griffiths, P. (2016). Is single room hospital accommodation associated with differences in healthcare-associated infection, falls, pressure ulcers or medication errors? A natural experiment with non-equivalent controls. *Journal of Health, Services Research and Policy*, 21(3), 147–155.
- Sinreich, D., & Marmor, Y. N. (2004). A simple and intuitive simulation tool for analyzing emergency department operations *Proceedings of the 2004 Winter Simulation Conference* (pp. 1994–2002). Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Spesvytseva, S. S. (n.d.). Hospital single rooms safer than multi-bed wards. *British Science Association: online*. Retrieved from <https://www.britishtscienceassociation.org/news/hospital-single-rooms-safer-than-multi-bed-wards>
- Taylor, S. J. E., Anagnostou, A., Fabiyi, A., Currie, C., Monks, T., Barbera, R., & Becker, B. (2017, December 3–6). *Open science: Approaches and benefits for modeling & simulation*. Paper presented at the 2017 Winter Simulation Conference (WSC), Las Vegas.
- Team SimPy. (2018). Overview — SimPy 3.0.11 documentation. Retrieved from <https://simpy.readthedocs.io/en/latest/>
- Wales NUS Estates. (2005). *Ward layouts with single rooms and space for flexibility*. Harrogate: NUS Estates.
- Weerawat, W., Pichitlamken, J., & Subsombat, P. (2013). A generic discrete-event simulation model for outpatient clinics in a large public hospital. *Journal of Healthcare Engineering*, 4(2), 285–306.
- Zhu, Z., Hoon, B. T. H., & Teow, K. (2012). Estimating ICU bed capacity using discrete event simulation. *International Journal of Health Care Quality Assurance*, 25(2), 134–144.

Appendix

Strengthening the Reporting of Empirical Simulation Studies (STRESS)

This table reports the models against the STRESS-DES guidelines, indicating where information is given in the paper or providing the details.

Section/Subsection	Item	Ward Model	ICU Model
1. Objectives			
Purpose of the model	1.1	See sections 3.2.1 and 4.1.1	See sections 3.3.1 and 4.2.1
Model Outputs	1.2	Data recorded for each patient/day during the model run and totals/averages reported at the end of the run: <ul style="list-style-type: none"> • Average waiting time – all patients and for those who wait for a bed • Maximum waiting time • Average time spent on the ward • Average number on the ward • Maximum number on the ward • Percentage occupancy of ward • Number of transfers between bays/single beds • Number who leave without getting into the ward 	Data recorded for each patient/day during the model run and totals/averages reported at the end of the run: <ul style="list-style-type: none"> • Percentage of Beds Occupied • Average Length of Stay • Total Patients Treated • Total Patients Admitted • Total Patients who waited • Total Patients Transferred to another hospital • Total Cancelled Surgeries • Percentage Cancelled Surgeries • Percentage of time at full capacity • Maximum number of beds used • The percentage of the time that each possible number of beds are in use. • The percentages are calculated at the end of each run using summations that are added to for each relevant patient through a simulation run.
Experimentation Aims	1.3	See section 4.1.3	See section 4.2.3
2. Logic			
Base model overview diagram	2.1	See Figure 1	See Figure 2
Base model logic	2.2	See sections 3.2.3 and 3.2.5	See sections 3.3.2 and 3.3.3 for adaptations to Ward Model
Scenario logic	2.3	Scenarios are created by changing the input data not the model.	Scenarios are created by changing the input data not the model.
Algorithms	2.4	See section 3.2.5	No additional
Components	2.5	The entities within the model are the patients who move through the system the attributes recorded for them are: <ul style="list-style-type: none"> • Gender • Group (for age groupings or other groupings) • Length of Stay • Max wait (now long they can wait for) • Start wait (simulation time at start of waiting period) • To Ward (for label based routing) • Bay No (when they are assigned to a bay) • Wait for assignment • Waiting time <p>Assign to ward, see section 3.2.5</p> <p>The beds in the ward.</p> <p>Patients who have completed their minimum waiting time are considered for entry to the ward in order of arrival to the queue.</p> <p>See section 3.2.3</p> <p>Patients enter at the point where a decision is made that they should transfer to the ward. Patients are sampled from the empirical distribution of patient arrivals.</p> <p>Patients exit when either they have waited too long or they have completed their stay in the ward.</p>	The entities within the model are the patients who move through the system the attributes recorded for them are: <ul style="list-style-type: none"> • Group (for groupings within patients) • LoopRoute (for waiting routing) • Length of Stay • Mixed (for label based routing) • NewElec (if they are an elective patient) • New EM (if they are an emergency patient) • QValue (used in calculations) • Space (for label based routing) • Space (for label based routing) • Start_wait (simulation time at start of waiting period) • Type (for the type of care needed) • Weight (used to assess staffing levels) <p>Assign bed see section 3.3.2</p> <p>The beds in the ICU. Nurse numbers are considered implicitly in the limit on the number of level 3 patients.</p> <p>Patients who have completed their minimum waiting time are considered for entry to the ward in order of arrival to the queue.</p> <p>See section 3.3.2</p> <p>Patients enter at the point where a decision is made that they should transfer to ICU. Patients arrivals are sampled from the empirical distribution of patient arrivals.</p> <p>Patients exit when either they are transferred out of the system because they have waited too long or they have completed their stay in the ICU.</p>
3. Data			
Data sources	3.1	• Interviews with stakeholders	• Interviews with stakeholders

(Continued)

(Continued).

Section/Subsection	Item	Ward Model	ICU Model
		<ul style="list-style-type: none"> • Samples of routinely collected data – 21 months of data; 970 patients using the existing wards including admissions, excess bed days, length of stay, gender and age • Population projections provided by Hampshire County Council see: http://www3.hants.gov.uk/factsandfigures/population-statistics/pop-estimates.htm 	<ul style="list-style-type: none"> • Samples of routinely collected data – 2 years of data covering 25,000 patients across the test network, including admission and discharge dates, delay in discharge, level of care (with dates of transfers between levels), type of patient (emergency/elective) and patient demographics • Population projections provided by Hampshire County Council see: http://www3.hants.gov.uk/factsandfigures/population-statistics/pop-estimates.htm
Pre-processing	3.2	Empirical distributions are used.	Empirical distributions are used.
Input parameters	3.3	Empirical distributions used throughout see section 3.2.4 and data input file available online for full list of input variables.	Empirical distributions used throughout see section 3.2.4 and data input file available online for full list of input variables.
Assumptions	3.4	That the patients arriving will follow the same arrival and LoS patterns for patients of each age group and the same proportion of patients will use the ward in relation to the age groups in the general population. Seasonal variations are not considered.	That the patients arriving will follow the same arrival and LoS patterns for patients of each age group and the same proportion of patients will use the ward in relation to the age groups in the general population. Seasonal variations are not considered.
4. Experimentation			
Initialisation	4.1	The model is empty on initialisation. A warm up period of 1 year is used. After 1 year the occupancy of the ward has stabilised. Model is non-terminating.	The model is empty on initialisation. A warm up period of 1 month is used. After 1 month the occupancy of the unit has stabilised. Model is non-terminating.
Run length	4.2	See 4.1.3	1 year run length, time units modelled are hours
Estimation approach	4.3	See 4.1.3 – 100 independent replications were used. This number was based on the advice of Simul8's trial calculator for a precision of 5% on the reported performance measures.	680 independent replications were used. This number was based on the advice of Simul8's trial calculator for a precision of 5% on the reported performance measures. This high number of reps was needed for the number of transfers, which was generally a low number.
5. Implementation			
Software or programming language	5.1	Simul8 Professional 2018 was used.	Simul8 Professional 2018 was used, this model has been run on a number of different computers.
Random sampling	5.2	Commercial software was used see 5.1.	Commercial software was used see 5.1.
Model execution	5.3	Commercial software was used see 5.1. Patients who have completed their minimum waiting time are considered for entry to the ward in order of arrival to the queue.	Commercial software was used see 5.1. Patients who have completed their minimum waiting time are considered for entry to the ward in order of arrival to the queue.
System Specification	5.4	Simul8 Professional 2018 was used. Run on a Dell Latitude E7440 with an Intel(R) Core™ i7-4600U CPU @2.10G Hz 2.70 GHz processor, 16.0 GB of RAM and a 64-bit operating system. The running time varied considerably depending on the scenario that was being run.	Simul8 Professional 2018 was used, this model has been run on a number of different computers including a Dell Latitude E7440 with an Intel(R) Core™ i7-4600U CPU @2.10 GHz 2.70 GHz processor, 16.0 GB of RAM and a 64-bit operating system. The running time varied considerably depending on the scenario that was being run.
6. Code Access			
Computer Model Sharing Statement	6.1	The model is available from: DOI 10.5281/zenodo.1468287	The model is available from: DOI 10.5281/zenodo.1468314