Strengthening the Reporting of Empirical Simulation Studies (STRESS)

Discrete-event simulation guidelines STRESS-DES

Section/Subsection	Item	Recommendation	How our paper addresses the recommendation
1. Objectives			
Purpose of the model	1.1	Explain the background and objectives for the model.	The paper states this in the background section.
Model Outputs	1.2	Define all quantitative performance measures that are reported, using equations where necessary. Specify how and when they are calculated during the model run along with how any measures of error such as confidence intervals are calculated.	The quantitative measures are mentioned (without equations) and include confidence intervals. The performance measures are collected after each run (300) and the average with 95% confidence intervals are presented in the results section of the paper. The performance measures reported are the mean time to treatment for each of the main treatment options and the percentage of patients that receive their first treatment within 62 days of the point of suspicion.
Experimentation Aims	1.3	If the model has been used for experimentation, state the objectives that it was used to investigate. a.) Scenario based analysis – Provide a name and description for each scenario, providing a rationale for the choice of scenarios and ensure that item 2.3 (below) is completed. b.) Design of experiments – Provide details of the overall design of the experiments with reference to performance measures and their parameters (provide further details in *data* below). c.) Simulation Optimisation – (if appropriate) Provide full details of what is to be optimised, the parameters that were included and the algorithm(s) that	The paper includes a description of the scenarios considered with a name and description for each. The scenarios were suggested by one of the lung consultants as they examine what is needed for the 62-day target to be met by 95% of the patients on the lung cancer diagnostic pathway under different testing and reporting strategies.

		was be used. Where possible provide a citation of the algorithm(s).		
2. Logic				
Base model overview diagram	2.1	Describe the base model using appropriate diagrams and description. This could include one or more process flow, activity cycle or equivalent diagrams sufficient to describe the model to readers. Avoid complicated diagrams in the main text. The goal is to describe the breadth and depth of the model with respect to the system being studied.		The models for the current pathway and the single cancer pathway both have flow diagrams included in the text.
Base model logic	2.2	Give details of the base model logic. Give additional model logic details sufficient to communicate to the reader how the model works.		A description of the base models is included in the method section of the paper and follows the diagnostic pathway a patient currently follows. The pathway process was provided by experts within each hospital to ensure that the models accurately represent the current diagnostic pathway.
Scenario logic	2.3	Give details of the logical difference between the base case model and scenarios (if any). This could be incorporated as text or where differences are substantial could be incorporated in the same manner as 2.2.		The scenario differences are explained in the text. The scenarios consider different frequencies of diagnostic tests and reporting strategies. The structure of the model does not change, only values of some of the input parameter.
Algorithms	2.4	Provide further detail on any algorithms in the model that (for example) mimic complex or manual processes in the real world (i.e. scheduling of arrivals/appointments/operations/maintenance, operation of a conveyor system, machine breakdowns, etc.). Sufficient detail should be included (or referred to in other published work) for the algorithms to be reproducible. Pseudo-code may be used to describe an algorithm.		The interarrival times and time in an activity are modelled using well-known distributions estimated from the NHS data provided. The distributions used include exponential, triangular, Gamma, Log Normal and Weibull. The models with the distributions can be made available on request.
Components	2.5	2.5.1 Entities	Give details of all entities within the simulation including a description of their role	The model considers urgent and non-urgent suspected cancer patients on their

		in the model and a description of all their attributes.	diagnostic pathway from the point of suspicion through to their first treatment.
	2.5.2 Activities	Describe the activities that entities engage in within the model. Provide details of entity routing into and out of the activity.	The activities reflect the clinics and diagnostic tests that each patient goes through during their diagnostic pathway. For example, patients attend outpatient appointments at the start of their pathway and towards the end when their diagnostic test results are explained, and their treatment discussed. The diagnostic tests are also included in the model. The activities which make up the current and proposed pathways are explained in the method section of the paper and use flow diagrams for ease of use.
	2.5.3 Resources	List all the resources included within the model and which activities make use of them.	The resources include: Lung consultants, oncologists, clinic appointments, slots for each of the diagnostic tests, MDT meeting slots, nurses, radiographers, and radiologists. In the simulation model, most activities have resources attached – these are set within the software.
	2.5.4 Queues	Give details of the assumed queuing discipline used in the model (e.g. First in First Out, Last in First Out, prioritisation, etc.). Where one or more queues have a different discipline from the rest, provide a list of queues, indicating the queuing discipline used for each.	The queuing discipline is First in First out. Most of the queues are unlimited in terms of capacity constraint apart from the queue between the MDT and decision to treat in the single cancer pathway which is limited to 3 days as specified in the

			If reneging, balking, or jockeying occur, etc., provide details of the rules. Detail any delays or capacity constraints on the queues.	National Optimal Lung Cancer Pathway.
		2.5.5 Entry/Exit Points	Give details of the model boundaries i.e. all arrival and exit points of entities. Detail the arrival mechanism (e.g. 'thinning' to mimic a non-homogenous Poisson process or balking)	Separate arrival points for urgent suspected cancer and non-urgent suspected cancer referrals. There are separate exit points for each of the treatment points (chemotherapy, chemoradiotherapy, radiotherapy, surgery, palliative care) as well as for active monitoring patients. There are also exit points for patients that are downgraded and no longer require treatment.
3. **Data**				
Data sources	3.1	List and detail all data sources. Sources may include: • Interviews with stakeholders, • Samples of routinely collected data, • Prospectively collected samples for the purpose of the simulation study, • Public domain data published in either academic or organisational literature. Provide, where possible, the link and DOI to the data or reference to published literature. All data source descriptions should include details of the sample size, sample date ranges and use within the study.		Referral data provided by the information team in the NHS. Diagnostic Test data provided by the NHS. Interviews and email conversations with stakeholders. The referrals data set relates to approximately 1,200 patients attending two Welsh hospitals between January 2018 and November 2019. The diagnostic test data set relates to the 1,928 tests that were conducted for the patients described in the referrals data set.
Pre-processing	3.2	Provide details of any data manipulation that has taken place before its use in the simulation, e.g. interpolation to account for missing data or the removal of outliers.		The data has not been manipulated to remove outliers or to account for any missing data.
Input parameters	3.3	List all input variables in the model. Provide a description of their use and include parameter values. For stochastic inputs provide details of any continuous, discrete, or empirical		The list of input parameters can be found in a separate document.

		distributions used along with all associated parameters. Give details of all time dependent parameters and correlation. Clearly state: • Base case data • Data use in experimentation, where different from the base case. • Where optimisation or design of experiments has been used, state the range of values that parameters can take. Where theoretical distributions are used, state how these were selected and prioritised above other candidate distributions.	The list of input parameters relates to the base model. For the scenario models, most parameters remain the same. The parameters that relate to arranging or reporting diagnostic tests change to fixed values specified in the scenario descriptions.
Assumptions	3.4	Where data or knowledge of the real system is unavailable what assumptions are included in the model? This might include parameter values, distributions, or routing logic within the model.	In the case of the activities where time studies were not available, estimates were obtained using expert opinion. For example, the time taken to discuss a patient at an MDT meeting.
4. **Experimentation**			
Initialisation	4.1	Report if the system modelled is terminating or non-terminating. State if a warm-up period has been used, its length and the analysis method used to select it. For terminating systems state the stopping condition. State what if any initial model conditions have been included, e.g., pre-loaded queues and activities. Report whether initialisation of these variables is deterministic or stochastic.	The model is terminating and ends after a simulated time of 692 days which matches the length of time specified in the referrals and diagnostic test data sets. The model does not include a warm-up period. The model does not use any preloaded queues or activities.
Run length	4.2	Detail the run length of the simulation model and time units.	692 days – matching the number of days covered in the NHS data set. The time unit is days.
Estimation approach	4.3	State the method used to account for the stochasticity: For example, two common methods are multiple replications or batch means. Where multiple replications have been used, state the number of replications and for batch means, indicate the batch length and	The model ran for 300 iterations (replications) and modelled the diagnostic pathways of 660 patients (Prince Charles Hospital) and

		whether the batch means procedure is standard, spaced or overlapping. For both procedures provide a justification for the methods used and the number of replications/size of batches.	511 patients (Royal Glamorgan Hospital) in each run.
5. **Implementation**			
Software or programming language	5.1	State the operating system and version and build number.	
		State the name, version and build number of commercial or open source DES software that the model is implemented in.	SIMUL8 (26, 3788)
		State the name and version of general-purpose programming languages used (e.g. Python 3.5).	Not used
		Where frameworks and libraries have been used provide all details including version numbers.	Not used
Random sampling	5.2	State the algorithm used to generate random samples in the software/programming language used e.g. Mersenne Twister.	Not known
		If common random numbers are used, state how seeds (or random number streams) are distributed among sampling processes.	Base Random Number Set 1 used in SIMUL8.
Model execution	5.3	State the event processing mechanism used e.g. three phase, event, activity, process interaction.	
		Note that in some commercial software the event processing mechanism may not be published. In these cases, authors should adhere to item 5.1 software recommendations.	SIMUL8
		State all priority rules included if entities/activities compete for resources.	No priority rules included.
		If the model is parallel, distributed and/or use grid or cloud computing, etc., state and preferably reference the technology used. For parallel and distributed simulations, the time management algorithms used. If the HLA is used then state the version of the standard, which run-time infrastructure (and version), and any supporting documents (FOMs, etc.)	Not applicable – model was run on a standalone laptop.
System Specification	5.4	State the model run time and specification of hardware used. This is particularly important for large scale models that require substantial computing power. For parallel, distributed and/or use grid or cloud computing, etc. state	Approximately 1 minute for 300 runs on a standalone laptop with Intel (i7) processor.

		the details of all systems used in the implementation (processors, network, etc.)	
6. Code Access			
Computer Model Sharing Statement	6.1	Describe how someone could obtain the model described in the paper, the simulation software, and any other associated software (or hardware) needed to reproduce the results. Provide, where possible, the link and DOIs to these.	The models can be made available on request. The person requesting the model would need the latest professional version of SIMUL8.

Article Information: http://dx.doi.org/10.21037/tlcr-20-919