# PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

# ARTICLE DETAILS

| TITLE (PROVISIONAL) | Modelling patient flows and resource use within a sexual health clinic through discrete event simulation to inform service redesign |
|---|---|
| AUTHORS | Mohiuddin, Syed; Gardiner, Rebecca; Crofts, Megan; Muir, Peter; Steer, Jonathan; Turner, Jonathan; Wheeler, Helen; Hollingworth, William; Horner, Paddy |

# VERSION 1 - REVIEW

| REVIEWER | Jingheng CAI<br>Sun Yat-sen University<br>P. R. CHINA |
|---|---|
| REVIEW RETURNED | 10-Feb-2020 |

| GENERAL COMMENTS | This manuscript developed a discrete-event simulation (DES) model to identify the bottlenecks in the current system and find possible routes for service improvement. In general, the manuscript proposed a computer simulation model to examine the use of an online service, however, the details were not presented well.

1. Page 9 line 1. The authors employed a Poisson distribution to sample "the arrival time" of patients. In general, Poisson distribution is used to sample "the number of arrival patients" during a specific time interval, whereas Exponential distribution is used to sample the "the gap of arrival time" between two patients. Some explanations of the proposed distributions are needed.

2. Page 11, second paragraph. One-way sensitive analyses were performed to assess the model robustness and some probability distributions with constant to verify the model outcomes. Unfortunately, the results were not presented. Given that the simulation model heavily depends on such perturbations, those results are very important.

3. Figure 1 is not clear. |
|---|---|

| REVIEWER | Emma Harding-Esch<br>LSHTM<br>Professional colleague of senior author, but do not know other authors. |
|---|---|
| REVIEW RETURNED | 07-Apr-2020 |

| **GENERAL COMMENTS** | This manuscript describes a discrete-event simulation method for assessing the impact on patient waiting times and costs of introducing an online STI/HIV testing service for asymptomatic patients, applied to a UK Sexual Health Clinic as a case study. |
|---|---|
| | Specific comments: |
| | Throughout: I would like to suggest that the term "self-testing" be replaced with "home-sampling", as "self-testing" implies the patient is conducting the test themselves: https://www.ncbi.nlm.nih.gov/pubmed/27811311. The model being described is on of home-sampling with others conducting the testing. |
| | Abstract: |
| | Spell out acronyms |
| | Page 2, line 13: "propose and evaluate through use of an example"? |
| | Page 2, line 22: Suggest clarify what "current system" means (SHS, in a specific service?) |
| | Page 2, line 28: I think it should be stated in the abstract methods that patient waiting times was the main outcome measure. |
| | Page 2, line 37: "longer waiting times" – compared to what? Perhaps this result should be removed from the abstract, with the introduction of the new service reduced average patient waiting time? |
| | Page 2, line 40: "reduced the cost … compared to the current system". |
| | Page 2, line 49: "could be beneficial" – in terms of patient waiting times and costs? Anything else? |
| | Methods: |
| | I think there is a checklist developed for the reporting of DES studies: https://www.tandfonline.com/doi/full/10.1080/17477778.2018.1442155 |
| | Table 1: |
| | - I wonder whether the table would look clearer by making gender columns, and the "bold headings" rows? |
| | - What does "interpreter used" mean? |
| | - Should "syphilis (heterosexual men only)" be moved to the "male" section rather than under "all"? |
| | Page 6, line 48: MSM already defined (page 5, line 32). |
| | Page 6, lines 51 & 53: Is this all patients, or just MSM? |
| | Page 6, lines 53 onwards: I'm not sure about this example to qualify the statement "For some patients, it is also necessary that they see a health adviser at their visit", since "If pelvic inflammatory disease is confirmed, she will then receive treatment and may be referred to a health adviser." I think the example could be removed, and the health adviser statement simply followed by: "Health advisers see patients with complex sexual health needs and…etc." This paragraph would benefit from Table 1 displaying the average proportion of patients that fall within each category, to gauge proportion of "complex" and "non-complex" patients in the SHC. |
| | Page 7, "Model formulation": would it be possible to alter the tense so everything is in the past tense? |
| | Table 2: |
| | - I do not understand footnote j "The cost estimates shown here are based on using a single resource". Would it be possible to note where the cost data come from in the footnote, as well as a note that this is staff time only? |
| | - I'm not clear from the table what the column "Proportion of patients from other service points" means. Would it be possible to clarify in the footnotes please? Should there be a value for "Patient registration at the reception" for this column? |

Page 10, line 22: Could you explain what the "cross-check" process involved?

Results:
Page 10, line 35: As above, unsure about "trend towards longer patient waiting times" as a result…compared to what? Is it not that the provision of the new services reduces patient waiting times?
Page 10, line 42: Perhaps indicate why you have highlighted the walk-in patients' results?
Table 2 needs renumbering as Table 3 (and in the text as well). Suggest the final column heading indicates it is "with increased number of patients arriving to clinic", rather than this only being indicated in the footnote.
Would it be possible to have a table with same structure as Table 3, but for costs?

Discussion:
Page 12, line 22: "halt the rise in STIs and HIV"… infections? Incidence? Prevalence? Progression to reproductive sexual health sequelae? Please qualify what aspect of STI/HIV rise SHSs are being asked to halt.
Page 12, line 30: "improve resource use." What do you mean by "resources"?
Page 12, line 37: I think you need to be more specific in what you mean by "differ slightly" and "common features", in terms of what aspects specifically?
Page 13, line 13: "as this patient type usually favours face to face testing," Can you provide a reference to substantiate this?
Page 13, line 52: complements, rather than compliments.
Is there any literature on how well patients self-assess as being asymptomatic? What is the risk of patients not having a face to face consultation if they don't "triage" themselves appropriately/try to "game the system" for a faster service?

| REVIEWER | Andrew Salmon |
| --- | --- |
| | University of Exeter UK |
| REVIEW RETURNED | 09-Apr-2020 |

| GENERAL COMMENTS | • The range of examined scenarios seems rather limited and very much steered toward the idea of online testing, although other simple options (for example having separate queues for booked and walk-in patients) are not reported. It would be helpful to have some comment on these, including the costing of reconfiguring the web services, which is apparently already underway, compared to the cost of making these changes or hiring more HCPs.
• There is no discussion of how the results have been implemented, which might speak to point 1, but for publication in a journal such as the BMJ I would think is merited
• Although the specific configuration of this model might be new, there are examples in the literature of modelling STD clinics e.g Brailsford and Viana c.2010-11. Some discussion of the previous literature in this area feels appropriate.
Regarding the model itself
• Process times in simulation models are generally rather right skewed, however the authors appear to have used a normal distribution to model their process times, which is symmetrical, and may underestimate the proportion of shorter service times. A more |
| --- | --- |

| | usual choice is log normal, which generally works very well. Was there a specific justification for normal, and how was the uncertainty estimated?<br>• I found no mention of a run in or warm up period for the model, in order that it reaches an equilibrium condition. Otherwise the benefits of the different scenario changes might be underestimated. If this was not applied, it should be justified. The model run period of 4 weeks, especially with that in mind, feels rather short, even if a run in was performed.<br>• Although DES is a stochastic tool, no uncertainty in the main results is presented despite the fitting of uncertainty distributions. An analysis of the uncertainty in the results would be highly desirable. |
|---|---|

## VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

1. Page 9 line 1. The authors employed a Poisson distribution to sample "the arrival time" of patients. In general, Poisson distribution is used to sample "the number of arrival patients" during a specific time interval, whereas Exponential distribution is used to sample the "the gap of arrival time" between two patients. Some explanations of the proposed distributions are needed.

Response: We agree with the reviewer that it was not explained clearly. In SIMUL8, arrival data can be sampled through the Day Planner option wherein it is possible to use a distribution option such as Random (Poisson number/Exponential inter-arrival). We have now edited/added the following text on Page 8: "We randomly sampled (Poisson number/Exponential inter-arrival) the arrival time of patients at the clinic based on an estimated 11 patients arriving every hour on average. The number of patients arriving each hour was represented by the Poisson distribution, while the Exponential inter-arrival times were set for the hourly segments of each day."

2. Page 11, second paragraph. One-way sensitive analyses were performed to assess the model robustness and some probability distributions with constant to verify the model outcomes. Unfortunately, the results were not presented. Given that the simulation model heavily depends on such perturbations, those results are very important.

Response: Page 9: "We added the provision of online-based self-sampling to existing clinic-based services. Based on findings from the literature,[3 6] we assumed that 50% of asymptomatic patients would opt for an STI or HIV self-sample. We then varied this assumption in sensitivity analyses by ±20%." Page 10: "Two one-way sensitivity analyses were performed to ensure varying parameter inputs had a viable and hypothesised impact upon the model outputs." The results of these two one-way sensitivity analyses are presented in Table 3 and also on Page 11: "Sensitivity analyses had the expected impact when we varied the provision of online-based self-sampling for 50% of asymptomatic patients by ±20% (Table 3)." We conducted white-box validation in which the internal working of each model component was tested. This method was used during the model development phase and includes unit-level checks. For example, in demand generation processes, the model was run to check that it is generating as many patients as expected. Anyway, to avoid confusion, we have edited/added the following text on Page 10: "We also tested the internal working of each model component during the model development phase. For example, in demand generation processes, the model was run to check if it was generating as many patients as expected."

3.  Figure 1 is not clear.

Response:  Figure 1 is clear to my computer which has a high screen resolution. Anyway, we have attempted to make Figure 1 clearer. We have also added some text in the figure caption for clarity.

Reviewer: 2

Specific comments:

Throughout: I would like to suggest that the term "self-testing" be replaced with "home-sampling", as "self-testing" implies the patient is conducting the test themselves:

https://www.ncbi.nlm.nih.gov/pubmed/27811311   The model being described is on of home-sampling with others conducting the testing.

Response:  Replaced "self-testing" with "self-sampling" as patients take their own sample(s) to be tested in the laboratory. The term "home-sampling" does not explicitly describe who (self or someone else) actually takes the sample(s) at home, and sampling may not necessarily take place at home. Although patients often self-sample when attending the clinic, we are explicit in stating that this is an online service and have chosen not to be restrictive about where the self-sampling takes place.

Abstract:

Spell out acronyms

Response:  Done

Page 2, line 13: "propose and evaluate through use of an example"?

Response:  Edited slightly

Page 2, line 22: Suggest clarify what "current system" means (SHS, in a specific service?)

Response:  Clarified

Page 2, line 28: I think it should be stated in the abstract methods that patient waiting times was the main outcome measure.

Response:  Stated

Page 2, line 37: "longer waiting times" – compared to what? Perhaps this result should be removed from the abstract, with the introduction of the new service reduced average patient waiting time?

Response:  Removed

Page 2, line 40: "reduced the cost … compared to the current system".

Response: Edited

Page 2, line 49: "could be beneficial" – in terms of patient waiting times and costs? Anything else?

Response: Edited

Methods:

Table 1:

- I wonder whether the table would look clearer by making gender columns, and the "bold headings" rows?

Response: Changed

- What does "interpreter used" mean?

Response: Clarified as "Language interpreter service by phone used"

- Should "syphilis (heterosexual men only)" be moved to the "male" section rather than under "all"?

Response: Moved

Page 6, line 48: MSM already defined (page 5, line 32).

Response: Edited

Page 6, lines 51 & 53: Is this all patients, or just MSM?

Response: MSM

Page 6, lines 53 onwards: I'm not sure about this example to qualify the statement "For some patients, it is also necessary that they see a health adviser at their visit", since "If pelvic inflammatory disease is confirmed, she will then receive treatment and may be referred to a health adviser." I think the example could be removed, and the health adviser statement simply followed by: "Health advisers see patients with complex sexual health needs and…etc." This paragraph would benefit from Table 1 displaying the average proportion of patients that fall within each category, to gauge proportion of "complex" and "non-complex" patients in the SHC.

Response: We decided to keep the example. Not all complex patients see a health adviser, the proportions relevant to our model are already provided in Table 2. We don't think it is necessary to provide these proportions as part of Table 1.

Page 7, "Model formulation": would it be possible to alter the tense so everything is in the past tense?

Response: Done

Table 2:

- I do not understand footnote j "The cost estimates shown here are based on using a single resource". Would it be possible to note where the cost data come from in the footnote, as well as a note that this is staff time only?

Response: Noted

- I'm not clear from the table what the column "Proportion of patients from other service points" means.

Response:  Added a footnote to make it clearer

Would it be possible to clarify in the footnotes please? Should there be a value for "Patient registration at the reception" for this column?

Response:  Clarified

Page 10, line 22: Could you explain what the "cross-check" process involved?

Response:  Explained on Page 10

Results:

Page 10, line 35: As above, unsure about "trend towards longer patient waiting times" as a result…compared to what? Is it not that the provision of the new services reduces patient waiting times?

Response:  Removed the text that caused confusion


Page 10, line 42: Perhaps indicate why you have highlighted the walk-in patients' results?

Response:  We indicated on Page 10 why we highlighted the walk-in patients: "We identified some bottlenecks under the current system, particularly within the consultation and treatment queues for male and female walk-in patients."


Table 2 needs renumbering as Table 3 (and in the text as well). Suggest the final column heading indicates it is "with increased number of patients arriving to clinic", rather than this only being indicated in the footnote.

Would it be possible to have a table with same structure as Table 3, but for costs?

Response:  Thanks – Table 2 has now been renumbered as Table 3 throughout. We also changed the final column heading as suggested. We can only produce up to five figures and tables.


Discussion:

Page 12, line 22: "halt the rise in STIs and HIV"… infections? Incidence? Prevalence? Progression to reproductive sexual health sequelae? Please qualify what aspect of STI/HIV rise SHSs are being asked to halt.


Response:  We have chosen to use the terminology adopted in the recent PHE report. Current population prevalence and incidence is unknown but is likely to have risen. Page 12: "… halt the rise in sexually transmitted infections …"  Page 4: "The number of sexually transmitted infections (STIs) continues to rise in the UK with a 5% increase since 2018.1  This has resulted in Public Health England (PHE) concluding in 2019 that local and national services for the prevention, diagnosis, treatment and care of STIs need to be delivered to the general population as well as a focus on

groups with greater sexual health needs.1  A new recommendation in 2019 is that individuals having condom-less sex with new or casual partners should have an annual STI screen and HIV (human immunodeficiency virus) test.1  This new emphasis on the importance of providing sexual health services (SHSs) for the general population has come at a time when budgets for SHSs are being cut."

Page 12, line 30: "improve resource use." What do you mean by "resources"?

Response:  Edited

Page 12, line 37: I think you need to be more specific in what you mean by "differ slightly" and "common features", in terms of what aspects specifically?

Response:  Provided examples

Page 13, line 13: "as this patient type usually favours face to face testing," Can you provide a reference to substantiate this?

Response:  Edited and provided references on Page 13

Page 13, line 52: complements, rather than compliments.

Is there any literature on how well patients self-assess as being asymptomatic? What is the risk of patients not having a face to face consultation if they don't "triage" themselves appropriately/try to "game the system" for a faster service?

Response:  Changed to complements – thanks. We searched but are not aware of any literature to comment on "how well patients self-assess as being asymptomatic".

Reviewer: 3

• The range of examined scenarios seems rather limited and very much steered toward the idea of online testing, although other simple options (for example having separate queues for booked and walk-in patients) are not reported. It would be helpful to have some comment on these, including the costing of reconfiguring the web services, which is apparently already underway, compared to the cost of making these changes or hiring more HCPs.

Response:  Pages 4 and 5: "... we developed a computer simulation model to examine whether introducing an online service for STI and HIV self-sampling for asymptomatic patients in addition to existing clinic-based services would decrease patient waiting times and improve resource use compared with the current system without online-based self-sampling."  As such, the range of scenarios examined was mainly based on the idea of introducing an online service for STI and HIV self-sampling. We did discuss during the model development process about having a scenario whereby there will be two separate queues for booked and walk-in patients. For this, it would have required making a structural modification to the model and relevant data were not readily available to examine this scenario. The Unity clinic does not run with separate queues for booked and walk-in patients, and the clinic does not see added value of having a separate queue only for the booked patients – the simulation model did not reveal bottlenecks in the reception queue. The model structure

was confirmed to be highly representative of the real-world system by the clinical staff at various stages of the model development process. The usual approach taken in simulation studies is to create a representation of the real-world, and we represented the clinic how it currently runs in the real-world. We investigated several scenarios and potential improvements to the current system based on the objective of this study and identifying the bottlenecks within the current system.

Pages 9 and 10: "The average cost of managing each patient via online was estimated to be £6.42 (Appendix 1), which is based on testing provided in-house at the Unity SHC." The impact on patient waiting times and costs in the system under different scenarios are described in the main text on Pages 11 and 12. Online testing has been used within the English National Chlamydia Screening Programme for sexually active men and women aged under 25 years, which is provided in Bristol through the Unity SHC.

• There is no discussion of how the results have been implemented, which might speak to point 1, but for publication in a journal such as the BMJ I would think is merited.

Response:  Page 15: "The Unity SHC is in the process of implementing the introduction of an online service for STI and HIV self-sampling in conjunction with a new rapid STI service which can potentially provide the result of gonorrhoea and chlamydia NAATs within 4 hours from the time of collection. This is being formally evaluated by NIHR Applied Research Collaboration West (https://arc-w.nihr.ac.uk/). This will be the subject of a future publication. We propose to model the potential impact of near patient or point of care testing for STIs using the DES methodology described here and then formally evaluate how the model would reflect the observed outcomes with the new Unity rapid STI service incorporating an in-house online-based self-sampling service."

• Although the specific configuration of this model might be new, there are examples in the literature of modelling STD clinics e.g Brailsford and Viana c.2010-11. Some discussion of the previous literature in this area feels appropriate.

Response:  We have now added the following text on Page 14: "Viana et al.23 developed a composite model in which a DES model of a hospital outpatient department that treats Chlamydia patients was combined with a system dynamics model of the infection process in the community. Their model shows how the prevalence of Chlamydia at a community level affects operational level decisions made in a hospital outpatient department."

Regarding the model itself

• Process times in simulation models are generally rather right skewed, however the authors appear to have used a normal distribution to model their process times, which is symmetrical, and may underestimate the proportion of shorter service times. A more usual choice is log normal, which generally works very well. Was there a specific justification for normal, and how was the uncertainty estimated?

Response:  We agree with the reviewer that a more appropriate choice would indeed have been to use the Log Normal distribution for service time. However, there is some evidence in the literature about the use of a Normal distribution to represent activity time where little data was available.22 Page 8: "Service time data (e.g. duration of consultation, examination and treatment) were estimated by 3 clinical consultants (MC, PH and HW) because these data were not readily available from the EPR.  We used probability distributions to model the variation in service times (Table 2), as these vary depending on the workloads, number of staff, patient factors and demands from the rest of the clinic." When using SIMUL8 software, it is advisable to use the Average distribution in simulations where little data is available to justify use of parametrised statistical distributions. With this, we only needed to input our estimate of the average (mean) of the time. A Log Normal distribution requires two parameter values (mean and standard deviation), but we could not obtain credible standard deviation

for using this distribution. An Average distribution in SIMUL8 is actually a Normal distribution with the standard deviation set to 25% of the mean. Given this large standard deviation, its use is more appropriate when a large variability is desired. Anyway, we have now added the following text on Page 14 as part of the limitations: "Service time data were estimated by 3 clinical consultants (MC, PH and HW) since these data were not readily available from the EPR. Although it would have been more appropriate to use the Log Normal distribution to sample the service time data, we used the Average distribution in simulations as we only needed to input our estimate of the average (mean) of the time and a large variability was desired. This was mitigated by using the same distributions for the analysis of all scenarios."

• I found no mention of a run in or warm up period for the model, in order that it reaches an equilibrium condition. Otherwise the benefits of the different scenario changes might be underestimated. If this was not applied, it should be justified. The model run period of 4 weeks, especially with that in mind, feels rather short, even if a run in was performed.

Response: The warm-up time is the time that the simulation runs before starting to collect results. This allows to get into conditions that are typical of normal running conditions in the system we are simulating. However, it is not always necessary to incorporate a warm-up time. An example of such a situation would be when we are simulating a shop that starts and ends the day empty. A warm-up time would be required when simulating a week in a factory, when the factory starts on Monday morning in the exact same state as it finished on Friday evening. Anyway, we added the following text on Page 10: "The Unity SHC that we modelled starts and ends the day empty, hence it was not necessary to incorporate a warm-up time in the model."

There is no fixed rule about how many runs one should do, it varies depending on lots of factors in a simulation. We simulated more than 2,000 patients over the model run period of 4 weeks and performed a trial of 1,000 model runs, meaning we virtually simulated more than 2,000,000 patients for each scenario to produce the comparative outputs. We feel this was a reasonable model run time period and is unlikely to alter the overall conclusions which were drawn from the model. We ran a scenario with a period of 8 weeks and compared it to the same scenario with a period of 4 weeks – the difference in outputs varied only by a fraction.

• Although DES is a stochastic tool, no uncertainty in the main results is presented despite the fitting of uncertainty distributions. An analysis of the uncertainty in the results would be highly desirable.

Response: We have now added 95% confidence intervals in Table 3.


## VERSION 2 – REVIEW

| REVIEWER | CAI Jingheng<br>Sun Yat-sen University<br>P.R.CHINA |
|---|---|
| REVIEW RETURNED | 28-Apr-2020 |

| GENERAL COMMENTS | My comments are well addressed. Thank you for the response. |
|---|---|

| **REVIEWER** | Emma Harding-Esch |
| | LSHTM, UK |
| | I am a member of professional bodies with the senior author but do |
| | not know the other authors. |
| **REVIEW RETURNED** | 07-May-2020 |

| **GENERAL COMMENTS** | I am happy with the response to reviewer comments. |

| **REVIEWER** | Andrew Salmon |
| | University of Exeter UK |
| **REVIEW RETURNED** | 15-May-2020 |

| **GENERAL COMMENTS** | 2020_05_15 |
| | |
| | First of all apologies for the late review due to a combination of annual leave and work on COVID related projects. Thanks to the co-authors for their patience, and their consideration of my previous comments. |
| | |
| | Having read the paper again I think there are a number of things that can be made more explicit. It is good to see OR models being put forward to a clinical audience, given that decision making within the health service does considerably under-use the technology and expertise that is available to it through collaborations such as the CLAHRC/ARC |
| | |
| | Most people who read the BMJ have probably never seen a SIMUL8 model diagram. Whilst figure 1 is reasonably interpretable visually, there are some features that could be clarified, which is perhaps what my colleague was suggesting in the previous round of comments. Could the authors confirm and make explicit that the four service points marked 'consultation (MW, FB etc)' act as a funnel through which the patients are directed into the array of purpose specific service points which is the next step in the model? If so, why not simply extract them from the queue using a label based system? If they do spend more than 0 time at this point, how is this adjusted with the consultation specific service point that immediately follows? |
| | |
| | It seems from the model that essentially this clinic is composed of two/four separate work streams using a combined lab facility. That is to say, the booked and walk in patients are separately staffed and could operate as separate entities. Or do staff cross over between stations depending on demand? |
| | |
| | I understand the interaction of exponential/Poisson as regards arrival times, but at the same time there is only one queue for the patient arrivals. Since a booked service doesn't really lend itself to exponential distribution modelling, could the authors provide more detail about how the entry to the mdoel operates. In particular, since the patients who turn up at the start of the day hoping to be seen are |

allocated slots that are evenly distributed throughout the day, then would they not find it more convenient to come back nearer the time? Furthermore, some patients presumably do not get slots? How are these accounted for, since under the current model it seems to be assumed that they do not attend again?

The lab seems to be a particular bottleneck, but if online capacity continues to increase, how do the authors envisage the lab service coping? What next for the service when this capacity is full?

I understand that the authors had very little data on service times and how the provision of the average distribution in simul8 helped the difficulty, but that still leaves the problem that the service times are unknown, and I found little in Maull et al that persuaded me that this use of a normal distribution is a way round the problem, rather, only a demonstration that in Maull et al the normal did not fit their data! To allow Simul8 to fix the variance doesn't seem to entirely solve the problem in the sense that a decision is still made as to what the variance is, and in the normal case the variance is entirely independent of the mean. I agree that a one parameter distribution such as exponential would likely not be a good choice. A chi square could be tried, as they have some skew at lower df, and could be programmed in terms of its gamma formulation if a chi2 is not available explicitly. Also, Simul8 licences are expensive, and health trusts as a rule do not have them. Open source methods such as Jaamsim or through SimPy are available, but require coding knowledge. So could the authors at least suggest what reproducers of this model might do in such a situation? Or comment on the desirability of this data being collected? Good models need good data.

## VERSION 2 – AUTHOR RESPONSE

Reviewer: 3

•      Most people who read the BMJ have probably never seen a SIMUL8 model diagram. Whilst figure 1 is reasonably interpretable visually, there are some features that could be clarified, which is perhaps what my colleague was suggesting in the previous round of comments. Could the authors confirm and make explicit that the four service points marked 'consultation (MW, FB etc)' act as a funnel through which the patients are directed into the array of purpose specific service points which is the next step in the model? If so, why not simply extract them from the queue using a label based system? If they do spend more than 0 time at this point, how is this adjusted with the consultation specific service point that immediately follows?

Response:

In modelling, there are multiple ways to solve a problem. We found it easier to model and visualise the patient flow the way we did. For clarity, we have added the following text as part of the caption of Figure 1 on Page 20:

"The four service points labelled as CONSULTATION MB, CONSULTATION MW, CONSULTATION FB and CONSULTATION FW are dummy activities through which patients are only directed to the next points of care. This was achieved by setting the process time of these activities to zero using a Fixed distribution."

•        It seems from the model that essentially this clinic is composed of two/four separate work streams using a combined lab facility. That is to say, the booked and walk in patients are separately staffed and could operate as separate entities. Or do staff cross over between stations depending on demand?

Response:

Two nurses work between the laboratories depending on demand. Also, there is only one health adviser who serves patients from both booked and walk-in clinics. This has now been clarified on Page 10.

•        I understand the interaction of exponential/Poisson as regards arrival times, but at the same time there is only one queue for the patient arrivals. Since a booked service doesn't really lend itself to exponential distribution modelling, could the authors provide more detail about how the entry to the model operates. In particular, since the patients who turn up at the start of the day hoping to be seen are allocated slots that are evenly distributed throughout the day, then would they not find it more convenient to come back nearer the time? Furthermore, some patients presumably do not get slots? How are these accounted for, since under the current model it seems to be assumed that they do not attend again?

Response:

In SIMUL8, arrival data can be sampled through the "Day Planner" option wherein it is possible to use a distribution option such as "Random (Poisson number/Exponential inter-arrival)" together with an option to select "Equally spaced within interval". We have added the following text on Page 8: "We did not select the "Equally spaced within interval" option as it is not recommended for life-like simulations. Given that patients arrive to the Unity clinic via one queue, some patients with booked appointments might be late in reporting to the reception. As such, we did not select the "Equally spaced within interval" option since this might not always be adhered to in a life-like situation via a common entry to the clinic. To mitigate this, the inter-arrival times were set for the hourly segments of each day rather than a longer time segment. We applied the same approach for the analysis of all scenarios to produce valid comparisons."  We ran a scenario with selecting the "Equally spaced within interval" option and compared this with the same scenario without selecting the "Equally spaced within interval" option. As expected, the average patient waiting times were found to be slightly lower when the "Equally spaced within interval" option was selected.

We also added the following text on Page 8: "It is possible that a walk-in patient can turn up in the morning, get a slot in the afternoon, and come back nearer the time. We did not model this delay because this patient will essentially be one of the hourly 11 patients arriving later in the day. If we were to model this delay, it would have had an unrealistic impact on the average walk-in patient

waiting times in the clinic. A patient who does not get a slot can always come back another day and be included in that day's arrival." Incorporating these for the purposes of this model was felt to be unnecessarily complex. We feel this was a reasonable compromise and is unlikely to alter the overall conclusions which were drawn from the model.

• The lab seems to be a particular bottleneck, but if online capacity continues to increase, how do the authors envisage the lab service coping? What next for the service when this capacity is full?

Response:

Page 11: "We identified some bottlenecks under the current system, particularly in the consultation and treatment queues for male and female walk-in patients." The introduction of an online service for STI self-sampling is only intended for asymptomatic patients. The costs associated with processing online based self-taken samples and testing are estimated separately (Page 10; Appendix 1). This has no impact on patient waiting times in the clinic.

Page 10: "Based on findings from the literature,3 6 we assumed that 50% of asymptomatic patients would opt for an STI or HIV self-sample." Page 13: "Our results suggest that the provision of online-based STI and HIV self-sampling alongside existing clinic-based services would decrease the average patient waiting times and reduce the cost of staff time for managing the patients. In addition, it would save money and release resources when provided in-house, which could be used to provide funding for the increased new demand that such a service will generate.4"

• I understand that the authors had very little data on service times and how the provision of the average distribution in simul8 helped the difficulty, but that still leaves the problem that the service times are unknown, and I found little in Maull et al that persuaded me that this use of a normal distribution is a way round the problem, rather, only a demonstration that in Maull et al the normal did not fit their data! To allow Simul8 to fix the variance doesn't seem to entirely solve the problem in the sense that a decision is still made as to what the variance is, and in the normal case the variance is entirely independent of the mean. I agree that a one parameter distribution such as exponential would likely not be a good choice. A chi square could be tried, as they have some skew at lower df, and could be programmed in terms of its gamma formulation if a chi2 is not available explicitly. Also, Simul8 licences are expensive, and health trusts as a rule do not have them. Open source methods such as Jaamsim or through SimPy are available, but require coding knowledge. So could the authors at least suggest what reproducers of this model might do in such a situation? Or comment on the desirability of this data being collected? Good models need good data.

Response:

To avoid confusion, we have deleted Maull's reference from the text. The reference of Maull was not intended to suggest that the use of a normal distribution is a way round the problem. The fact that the use of a normal distribution did not fit the data well in Maull's case does not necessarily mean it is always wrong to use a normal distribution for activity times in other situations. Due to data unavailability, it was not possible to determine how well any distribution fits the data set. A usual supposition is that the distribution for most activities in a project network is right skewed, but it does

not guarantee the use of a specific distribution. For example, a three-parameter beta distribution may perform better than a normal or log normal distribution for activity times. The use of a distribution entirely depends on the data set at hand; one can only speculate what might have been more appropriate in the absence of real data. There are several distributions available for use for service times, including beta, normal, log normal, triangular, etc. However, in the absence of data, the provision of Average distribution in SIMUL8 provides a workable way of dealing with such a problem – three clinical experts agreed that activity times would vary by ±25% of the estimated mean. A more appropriate choice might have been to use a log normal distribution for service time, but we could not obtain credible standard deviation for using this distribution. We have acknowledged this as a limitation on Page 14: "However, service time data were estimated by 3 clinical consultants (MC, PH and HW) since these data were not readily available from the EPR. A usual supposition is that the distribution for most activities in a project network is right skewed, but it does not guarantee the use of a specific distribution. It might have been more appropriate to use a Log Normal distribution to sample the service time data, but we used an Average distribution in simulations as we only needed to input our estimate of the average (mean) of the time and a large variability was desired. An Average distribution in SIMUL8 is a Normal distribution with the standard deviation set to ±25% of the mean. This was mitigated by using the same distributions for the analysis of all scenarios. The use of a Log Normal distribution would require one to collect two parameter values (mean and standard deviation), to potentially incorporate uncertainty more appropriately."

A recent systematic review of the use of computer simulation modelling of patient flow within UK emergency departments found that about 50% of the studies used SIMUL8 software for running their model (https://www.ncbi.nlm.nih.gov/pubmed/28487459). SIMUL8 can provide an insight into the workings of a system through visual animation. Advising on which computer software would be of best use for replicating our model is outside the scope of this paper.

## VERSION 3 – REVIEW

| REVIEWER | Andrew Salmon<br>PenCHORD<br>University of Exeter UK |
|---|---|
| REVIEW RETURNED | 21-May-2020 |

| GENERAL COMMENTS | Thank you for addressing the additional comments. |
|---|---|