

## MODELLING AND SIMULATION OF GENOMIC SEQUENCING PLATFORM OPERATIONS

Jules Le Lay  
Lionel Perrier

Vincent Augusto  
Xavier Boucher  
Xiaolan Xie

Direction of Clinical Research and Innovation  
Centre Leon Berard  
Lyon, FRANCE

Mines Saint-Etienne  
Univ Clermont Auvergne  
INP Clermont Auvergne  
CNRS, UMR 6158 LIMOS  
F - 42023 Saint-Etienne, FRANCE

### ABSTRACT

This paper focuses on the healthcare application field of Genomic Sequencing and addresses the challenge of efficient organization and ramp-up of sequencing platforms. High-throughput sequencing platforms are currently in an industrial prototyping phase in France for large national deployment afterwards. In the current state of our knowledge, there is no scientifically established generic model nor decision-making support at the operational level which could guide the medical authorities in designing organizational rules, then managing the deployment of such platforms at the national level. After analyzing the state of the art, a simulation model of a genome sequencing platform is presented, then used as a decision-making support to manage a ramp-up situation for an application case of a French sequencing platform. These first results are discussed, together with the perspective to develop a generic model and decision-aid approach.

## 1 INTRODUCTION

### 1.1 Sequencing Context

The rapid development of whole genome sequencing since the first sequencing (International Human Genome Sequencing Consortium 2001; Venter et al. 2001) has opened up new opportunities for medical sciences. International programs, such as the 100,000 genomes project supported by the UK, are exploring the application of genomic medicine (The 100,000 Genomes Project Pilot Investigators 2021). In France, the use of genomic sequencing in medicine is coordinated by the French national genomics medicine plan or “Plan France Médecine Génomique” (PFMG) (PFMG 2025 2017). Those two programs are particularly investigating the diagnosis of Rare Diseases (RD) and cancer using sequencing data. For rare genetic diseases, application of Whole Genome Sequencing (WGS), Whole Exome Sequencing (WES) and Transcriptome analysis (RNAseq) techniques could shorten the diagnostic wandering and thus help treat patients (Stranneheim et al. 2021), while providing treatment recommendations for oncology patients, using e.g. RNAseq as in (Pleasant et al. 2022). International cooperation has developed to gather evidence and advancements in the field (Boycott et al. 2017).

The PFMG, aims to change the way patients are diagnosed, prevented and treated. The plan is organized around 3 major objectives: (i) prepare the integration of genomic medicine into the patient care pathway, (ii) develop the national sector of genomic medicine and (iii) install France as a leader amongst the countries involved in the sector (Lethimonnier and Levy 2018; Sanlaville et al. 2021).

To do so, this sector must overcome several challenges, like establish genomic medicine as a tool for personalized medicine, exploring the genomics specificity of pathologies for the therapeutic benefit of the patient, and developing the infrastructure to acquire, treat and interpret genomics data. Additionally, the

sector needs to control the costs of genomics sequencing to ensure the viability of its integration within the overall healthcare system.

In this context, the two French genomics laboratories or “Laboratoires France Médecine Génomique” (LBMMS) dedicated to genomic sequencing have opened to serve develop high-throughput genomic sequencing and answer the challenges posed by this new tool. The prescriptions on French territory are dispatched toward one of the laboratory depending on the location of the prescriber. These LBMMS are each composed on a sequencing facility, calculation and storage and archiving infrastructures.

A list of pre-indications for which patients will be able to benefit from genomic sequencing have been established. Currently, 61 pre-indications have been selected: 51 in the area of RD, 8 in oncology, and 2 in oncogenetic pre-indications.

## **1.2 Problematic**

The organizational design of genomic sequencing platforms is an emerging field, both technically and from a scientific point of view. The literature directly dedicated to this topic remains scarce (see section 2) and no generic model nor decision-making support to provide medical authorities with consistent guidelines have been found. In France, two regional platforms, AURAGEN and SeqOIA, serve as experimental demonstrators to assess the organizational efficiency of the sequencing process. The present research intend to propose a first step to close this research gap. The objective is to initiate the development of a generic tool that medical authorities could utilize for determining consistent organizational rules and processes, adapt them to local and contextual configuration of each potential deployment. To initiate this generic approach, the paper presents:

- A model of genome sequencing platform using Anylogic® software;
- A first decision-making utilization of this model, to address an operational situation of ramp-up for genome sequencing processes.

The next sections describe the organization of the sequencing facility and the operations of the platform. We focus on modelling physical and informational flows in the laboratory. We present a Discrete-event simulation (DES) model of the activities performed at a macroscopic level, with a special attention to resource management. To illustrate the functioning of this model, we simulated the necessary ramp-up of the platform to achieve the objectives of the plan, and analyzed the consequences on the resources needed and stocks of consumables to maintain.

## **2 LITERATURE REVIEW**

### **2.1 Operation Simulation and Optimization**

DES is a simulation technique that models the operations of the system as a discrete sequence of events in time. It models systems as networks of queues and activities, through which the products pass. DES is particularly suited to study systems at a tactical or operational level, as it allows detailed representations of systems (Brailsford and Hilton 2001).

The use of DES in healthcare is abundant in the literature. An umbrella review on simulation modelling in healthcare identified 37 reviews on the subject (Salleh et al. 2017). Among the 14 reviews reporting the simulation method used in the included articles, DES was the most in 11 of them. In total, 58% of articles used this method (without accounting for potential duplicates between the different reviews).

A recent review on DES applied to healthcare systems identified 231 papers, with 40% of them published between 2018 and 2021, which highlights the increasing interest in this technique (Vázquez-Serrano et al. 2021). Identified papers tend to focus on operational problematic, with 36 papers studying primarily resource allocation and resource schedules, and 84 on process time and efficiency.

Production ramp-up is a critical step in a product's life-cycle, and thus for industrial companies, which results in a lot of investigation on the subject. The main characteristics of ramp-up phase as Interruptions / Uncertainty / Defects / Learning / Demand growth / Price reduction / Finite horizon (Glock and Grosse 2015) . In this context, simulation is a well identified decision support tool. For instance, it can help to anticipate the modification of capacity requirements induced by ramp-up, or to evaluate the impact of lot-sizing on the production.

Medical laboratory analysis using simulation is a covered area in the literature. For instance, DES to optimize resource allocation is used in (Lote et al. 2009). Optimizing routes to deliver samples to the lab, and reallocating resources allowed them to standardize resource load. DES has been used by (Yang et al. 2014) to assess the performance of several lean management techniques to improve laboratory organization. The most commonly used performance indicators in laboratory performance analysis are detailed in (Tsai et al. 2019). Turn-around time is the most used, however, its definition differs per laboratory and discipline.

The use of DES at operational levels can also be seen in Health Economics analysis. For Health Economics evaluation in the field of Personalised Medicine (PM), (Degeling et al. 2017) reports that Markov modelling and decision-tree analysis are the most commonly used simulation techniques in the field. However, the authors note that there is a significant increase in the use of alternative simulation and modelling methods, like DES, since 2011. Following this review, (Marshall et al. 2020) presents DES and Agent-Based Modelling (ABM) methods and their potential uses for Health Economics evaluation in the field of PM, especially for the customization of treatment pathways. This article also recommends to integrate simulation and modelling techniques to the Health Economics toolkit.

## **2.2 Sequencing and Simulation**

The rapid development of high-throughput sequencing opened new opportunities for personalized medicine, which incited institutions and researchers to investigate its use at large scale.

As a consequence, multiple studies have investigated the economical viability of genomic sequencing. Vu et al. (2021) implemented DES with infinite resources to study the costs and resource implications of different strategies of gene panels sequencing. Similarly, Wu et al. (2022) uses DES for the economic evaluation of using genomic sequencing of children mitochondrial disorder. Incerti et al. (2022) and Li et al. (2021) conducted cost-effectiveness analyses of genome-wide sequencing, in the case of undiagnosed RD for the former, and on developmental disabilities and multiple congenital anomalies for the latter. If Incerti et al. (2022) uses an online deterministic tool to calculate the cost-effectiveness of WGS, Li et al. (2021) chose to use DES modelling.

In France, the PFMG prepares the integration of genomic medicine into the patient pathway. Such integration requires to organize the healthcare system around the genomic facilities. Such organization of healthcare networks around high throughput sequencing in the Netherlands is investigated in van de Ven et al. (2022) using a hybrid ABM and DES model to calculate the cost of WGS. This article uses ABM to simulate the network of hospitals in the Netherlands and the referral of patients to other hospitals or sequencing facilities. DES is used for the modelling of prescription and hospitals internal processes. Based on the simulation model developed in van de Ven et al. (2022), Soltanolkottabi et al. (2022) implemented System Dynamics (SD) and Game Theory elements to model WGS implementation for evaluating bio-marker testing strategies.

In addition to the presented studies, several available online study protocol are reporting to use simulation in their methodology for economic evaluation. Berglund et al. (2022) and Hayeems et al. (2022), highlighting once again the possibilities of DES to study process operations.

## **3 MODELLING AND SIMULATION**

The objective of this study is to model and simulate the operations of a genomic sequencing facility. This section aims at describing the methodology used to build this model.

### **3.1 Case Study**

The model presented hereafter is focusing on the AURAGEN LBMMS. The organization is composed of a grouping of health cooperation or “Groupement de Coopération Sanitaire” (GCS) formed by the four university hospital or “Centre Hospitalier Universitaire” (CHU) and two cancer centres or “Centre de Lutte Contre le Cancer” (CLCC) from Auvergne Rhône-Alpes region (AURA) and multiple academic and research partners, including the École des Mines de saint-Étienne and Claude Bernard University of Lyon. AURAGEN is in charge of sequencing samples from the southern and eastern parts of mainland France, and from French overseas department and territories.

A national global organization of care was installed to facilitate the process. Prescriptions are submitted to a national centralised tool and validated by dedicated Pluridisciplinary Consultation Meeting or “Réunion de Concertation Pluridisciplinaire (RCP) for oncology and RD care. Once the prescription is validated, samples are collected and transported to the LBMMS. For AURAGEN, samples are sequenced in the Hospices Civils de Lyon (HCL) site. The High Performance Infrastructures where sequences are treated using bio-informatics algorithms are located in the city of Grenoble. The development of bio-informatics algorithms are performed in the Cancer Center Léon Bérard for cancer-related sequencing and in the CHU Grenoble Alpes for RD. These two entities also supervise the interpretation of the results by experts of the domain affiliated to the AURAGEN LBMMS.

In the present paper, we intend to model the overall patient’s pathway in genomic medicine, while modelling in detail the operations taking place on the platform, from the reception of samples until the handling of the sequence analysis to the prescribing health professional.

### **3.2 Simulation Model**

The first step toward building the simulation model consisted in mapping the different steps of the process, identifying the actors and resources involved at each steps and linking the activities to each other. We detail here-after the methodology used to understand this process and build a representation as a base for the simulation.

#### **3.2.1 Methodology**

This representation was built in co-construction with key actors of the process. The model was built in a series of interviews with each of them. The number and agenda of interviews were adapted to the contacts.

The first session’s objective was for us to present our work and methodology, and for our contacts to present the sub-process they are in charge of. Based on their understanding of the process, the authors built a first version of the model based on the explanations of experts. Subsequent interviews all followed the same agenda: the authors presented the model to the experts, that reviewed and corrected it. Identified changes were incorporated and validated in the following interview. The process was repeated until experts deemed the model representative.

The simulation model is a hybrid DES and ABM simulation that emulates the flow of samples in a french high-throughput genomic sequencing facility. It contains the main operations performed on the samples at a macroscopic level. Indeed, the operations are modelled as aggregated steps, without detailing the machine or samples preparation or other performed by the technicians. The following paragraphs detail the model and the different mechanisms at stake.

#### **3.2.2 Non-technical documentation**

The presented simulation model is a hybrid simulation combining DES and ABM elements. It models the operations of a high-throughput genomic sequencing facility, the patients’ pathway, the physical flow of sequenced samples and the bio-informatics process that allow sequence analysis.

This model is dedicated to the operational study of the platform. It will be used to identify potential bottlenecks, additional resource requirements or impact of production planning modifications on performance indicators.

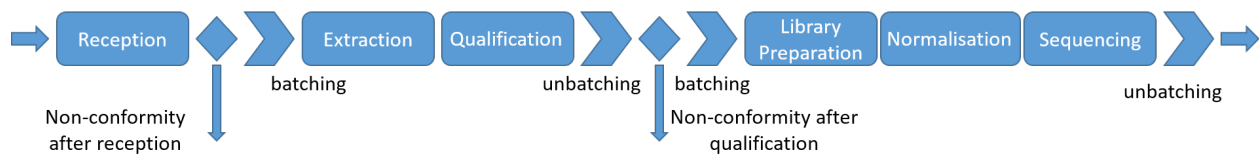
The model is built around two agents, the *Patient* and the *Sample*. State-charts are used to model the agents evolution.

Patient agents evolve through the genomic medicine pathway. They are generated at predefined rate and distributed between RD and cancer patients. The initial and consent consultation, the pluridisciplinary meeting which validates the prescription and the blood sampling are the first states of the patients statechart. Patients are then blocked until the completion of the sample flowchart. The final pluridisciplinary meeting validating the sequencing results and the final consultations are the two final flowchart's states.

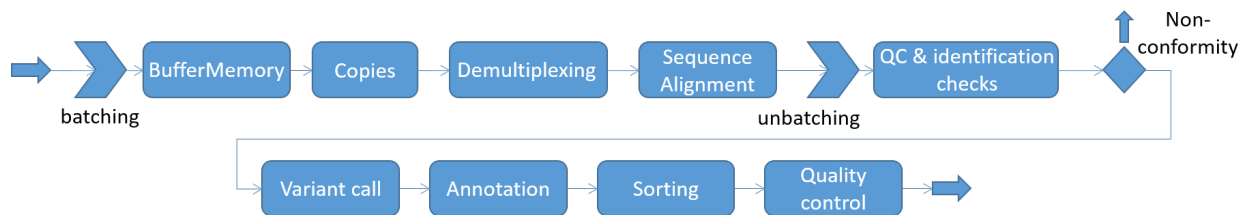
Sample agents' flowchart model the LBMMS process with 3 main steps: sequencing and bio-informatics analysis, interpretation of sequencing's results. An additional step is here to handle the various errors that could occur in the sequencing. Each state of the chart triggers DES micro-processes detailing the operations performed on patients and samples.

Patient's states trigger simple processes composed of a delay block modelling the time spent with the health professional. These states will not be discussed further but are present for futur development of the model. Patients' sampling process triggers the creation of the corresponding samples. Sequencing and bio-informatics states trigger more complex DES processes detailing the operations performed on the sample in the sequencing lab. These DES processes are presented in Figures 1a and 1b.

For those two processes, we developed a set of rules to model the flow management of samples and files in the LBMMS. Priority rules are coded using AnyLogic 'queue' blocks. The management of batches is performed in 'batch' blocks, with additional coding allowing the weekly constitution of undersized batches to launch cancer process steps. Cancer and RD samples are batched separately. Resource seizing and delays duration are done in 'seize' or 'service' blocks using functions that allow to dynamically seize the appropriate resource and enter the delay duration corresponding to the step and nature of samples.



(a) Schematic DES diagram of the sequencing process.



(b) Schematic DES diagram of the bio-informatics analysis process.

Figure 1: Schematic DES diagrams of the main sub-processes of the laboratory.

The inputs used for the simulation are the laboratory usage data. We used the number of received prescriptions and received samples to calculate the admission rate of patients in the France Genomics pathway. Operation and waiting times have been evaluated on site during the observations or estimated by the operators and laboratory engineers.

The modeled resources are divided in two groups: human resources (namely receptionists and technicians), and machines. There are two receptionists that operate on opening hours. In the sequencing process, the five technicians perform the various process steps. We included several resource pools of technicians in the model, each dedicated to a precise step. A standard weekly planning, see Table 1, was established in consultation with the laboratory actors based on the platform’s activity, to replicate their organization.

Three machines are included in the simulation for RD extraction, cancer extraction, library preparation. Their capacities are defined using schedules on opening hours to facilitate calculation and interpretation of occupation rates. There are four sequencers in the laboratory having 2 independent slots each. A schedule is used to plan their weekly maintenance (4 hours every monday morning). Appendix A details the process steps’ duration and resources used, as well as batch management throughout the process.

Table 1: Details of technicians pool implementation and assignments to the different sequencing process tasks.

Step	Monday	Tuesday	Wednesday	Thursday	Friday
Reception	0	0	0	0	0,5
Extraction RD	1	1	0.5	0	0
Extraction Cancer	1	1	0	0	0
Qualification	0	1	0.5	1	0
Genotyping	1	1	0.5	0.5	0
Library Preparation RD	1	1	0.5	0.5	0
Library Preparation Cancer	0	0	1	2	2
Normalisation RD	0	0	0	0	0
Normalisation Cancer	0	0	0	0	0
Sequencing	1	0	1	0	1

The model was validated using the following methods, as recommended by (Sargent 2013).

1. Animation: The authors validated the model’s behavior through the visualization of the simulation’s execution.
2. Face validity: The model was built in collaboration with experts of the sequencing and bio-informatics process. They validated the final model’s logic in regard of the real system.
3. Internal validity: Several runs were made to determine the amount of internal stochastic variability in the model.
4. Sensitivity analysis: We varied the level of the admissions to assess their impact on the simulation’s outcome.

This model is intended to serve as an operational analysis and exploration tool. It could be adapted to other applications, such as Health Economics analysis, although an additional development effort would be necessary to add several important aspects that would allow evaluation of the sequencing costs.

The main limitation of the model resides in the complexity of the laboratory operations. Indeed, two types of patients follow the sequencing process, RD and cancer patients, and 3 types of sequencing are performed: WGS, WES and RNAseq. RD patients provide one blood sample, on which WGS is performed. Samples from cancer patients give 2 samples to the platform: a blood sample and a tumor sample. From the blood sample, only WGS is performed, although it might change in the futur, and from the tumor sample, all three techniques are performed. The low proportion of cancer patients forces the staff to perform manually most of the process steps. It results in a multiplication of micro-activities performed by technicians that are difficult to model precisely.

The platform is in constant evolution and the studied process have changed several times in the course of the project. As a consequence the model must be regularly updated or it will become obsolete and unusable for its purpose.

## **4 EXPERIMENTAL DESIGN**

Our model is intended to offer an operational analysis of the AURAGEN platform, a genomics laboratory part of PFMG. To this end we designed a set of ramp-up experiments to illustrate the model's use as an organizational analysis tool. The present section details these experiments and their settings.

### **4.1 Experiments**

We implemented several experiments to showcase here the potential of the model to use it as a decision-aid tool. We present in this section three different experiments. First, we implemented a validation scenario to assess the internal stochastic variability of the model. An approximation of the current demand was passed as the sequencing demand and we ran the simulation for 20 replications, to achieve a statistical accuracy without impacting on calculation time. Secondly, we ran a sensibility analysis to assess the robustness of the simulation with different arrival rates. Lastly, a ramp-up of the sequencing demand was implemented, to take it from the actual level to the maximal demand that can be extrapolated from PFMG roadmap, and analyze the potential consequences for the platform.

### **4.2 Validation experiment**

For the validation experiment, we set the arrival rate of patients in the PFMG pathway to match the volume of samples arriving on the platform each week. This number was estimated to about 200 RD samples and 10 cancer files (composed of a blood and a tumor sample). The base arrival rate was fixed to 2 patients per hour between 9am and 5pm each day of the week. The warm-up period is set to 182 days to allow the system to stabilize. This duration was selected after several attempts.

### **4.3 Sensitivity analysis**

In the second experiment, we analyze the influence of the arrival rate on the system. We set the arrival time between 1.43 and 4.05 patients per hours, with a step equals to 0.24, which would correspond to 8.3 files per week. These arrival rates correspond to the reception of between 150 to 425 samples per week (0 - 142 files per week). This resulted in 11 experiments to perform for this sensitivity analysis. The same warm-up period is used.

### **4.4 Ramp-up experiment**

The ramp-up experiment is composed of 9 experiments. Duration and warm-up period are the same as in the validation experiment. Each is set in three different phases:

1. A stable phase of 365 days using the validation arrival rate;
2. A linear arrival rate augmentation for a fixed period of 547 days (one and a half year);
3. A final stable phase at the final arrival rate between 225 and 400 samples per week (between approximately 67 and 142 prescriptions per week), with a step of 8 prescriptions per week, resulting in a total of 9 experiments.

### **4.5 Result treatment**

To evaluate the operational capabilities of the laboratory, we measured several indicators relative to the process' performance and capacities. These indicators are calculated between the end of the warm-up and the end of the simulation and are detailed here under:

- Number of patients prescriptions completed;
- Lead time per sample, measured from reception of samples in the laboratory until results are returned for interpretation: mean and std deviation;

- Mean utilization per resource (by resource pool).

## 5 RESULTS

We now present the results of the experiment described in Section 4.1. The validation experiment, using an arrival rate of 210 samples per week results in the completion of 3,260.90 files per year ( $\sigma = 22.88$ ) on average during the period of interest of 2.5 years. The mean file lead time is 45.24 days ( $\sigma = 0.4$ ).

The sensitivity analysis' results - performance indicators and laboratory's human resources occupation - are displayed in Figure 2. Ramp-up's performance indicators are presented in Figure 3. Technicians' occupation is not displayed, the process is saturated and the graph shows no evolution for steps between genotyping and sequencing.

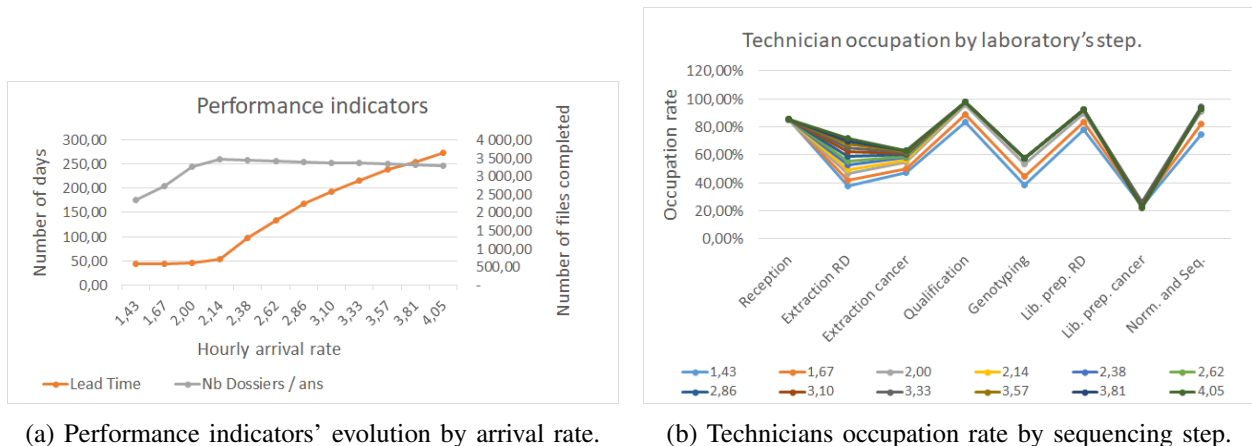


Figure 2: Sensitivity analysis results.

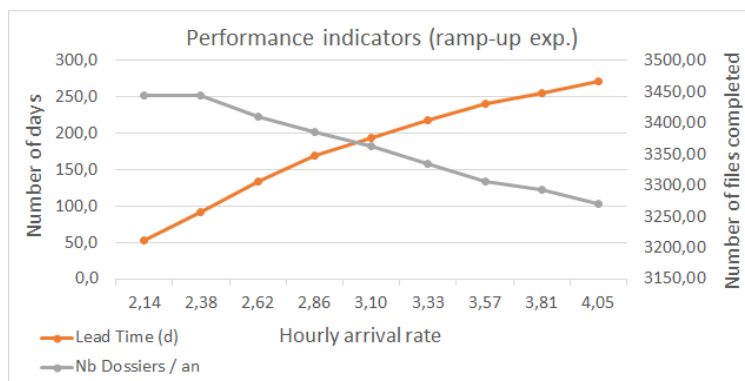


Figure 3: Ramp-up experiment results: performance indicators as a function of final arrival rate.

## 6 DISCUSSION

When comparing simulation's output to the laboratory's performance indicator, we have observed that the simulated lead time is quite close to what is observed in the lab. However, this result does not account for the high variability observed in the laboratory. Lead time highly varies from one file to another. Issues with samples or consentment papers at reception, or insufficient DNA quality may a new sampling, and thus delay the other samples of the file. In some rare cases, those issues might cause inconsistencies in



the data. A detailed statistical study should help modelling these aspects of the sequencing process. The number of simulated completed files is about 16% lower than the actual laboratory results.

According to the sensitivity analysis, the AURAGEN laboratory is at near maximum capacity. Indeed the experiment's results show that the number of files reaches a peak at the fourth experiment (75 prescriptions generated per week) at 3,454.16 files ( $\sigma = 52.87$ ) and then slowly decreases while the average lead time increase. This indicates that arrivals exceed the laboratory's processing capacity and that the work-in-process is increasing. Discussions with the laboratory technicians and management, and a study of the laboratory planning indicates that the laboratory is not this close to saturation, and genotyping is not a process bottleneck. Modifications of the implemented planning might be necessary to reflect more the flexibility of the laboratory.

A possible improvement is the addition of details to each process steps. For now we considered each step as a homogeneous block, detailing the sub-processes (machine preparation, cleaning, and so on) would allow more precise resource utilization.

The present model is intended to be generalized to other platform organizations and to be made available to the French laboratories to evaluate the impact of strategic decisions on the organization of production. This model should be used to explore the impact of the perspectives for the genomic industry on the platforms. For instance, validating whole-genome and transcriptome approaches (Wrzeszczynski et al. 2018), or proving the clinical interest of whole genome and transcriptome sequencing, as suggested by several publication (Cuppen et al. 2022; Pleasance et al. 2022), will lead to the generalization of this technique and an increase sequencing demand. Adding human resources' involved in the bio-informatics part of the process is also necessary to improve the operational analysis of this process.

## **7 CONCLUSION**

In this paper, we presented a simulation model for the operational analysis of a high-throughput sequencing facility in France. We described the process mapping in collaboration with key collaborators, and the simulation model co-construction. We also provide a thorough description of the model and simulation experiments.

This model still needs several improvements to be fully operational. Additional details should be added to the steps of the laboratory process and account for machines preparation times, loading times, etc. Additional consolidated data is needed to feed this improved model. Further experiments run are still needed to adjust the model to the laboratory.

In addition, adapting this model to other high-throughput sequencing facilities would allow a cross-analysis of the different organizations and help the development of new sequencing laboratories in France.

## **ACKNOWLEDGMENTS**

The authors thank all the collaborators for the time invested in this project. The help of all members of the AURAGEN steering committee and the support of the Direction Générale de l'Offre de Soins (DGOS) was invaluable in the implementation of this project. We would like to thank in particular Mrs. Christine Vinciguerra and Anne Thomas, for the organization of our laboratory visits and carefully reviewing the "wet" part of the model. We also thank the technicians for letting us follow them in their activities on the platform. We are also very grateful for the participation of Virginie Bernard, Anthony Ferrari and Julien Thèvenon in the modelling of the "dry" part of the platform.

## **A SIMULATION MODEL DOCUMENTATION**

We provide here-under the documentation on the simulation model developed for this study, following recommendations for the reporting of simulation models like Eddy et al. (2012) and Sargent (2013).

Different batch sizes are used depending on the prescriptions' type (cancer or rare disease) and on the process step. RD are grouped by plates or racks of 96 samples for extraction and qualification. Controls are

performed on individual samples and grouped again for library preparation and normalization. Sequencing is performed on pools of 24 samples. For cancer, as the volume of prescriptions is lower, the laboratory starts the production of a new batch every week regardless of the number of samples. The maximum number of samples processed at once is limited to 12, due to the constraints on the current process.

Table 2 details the process times and resources involved in each activities. The delay times are extrapolated from estimations made by experts. The sequencing for RD samples is performed twice, to normalize the pool of samples, and then to sequence it.

Table 2: Summary of each process step duration and necessary resources.

Service block	Description	Delay time	Resource
Reception	Administrative sample verifications and registration.	15 min	Receptionist
Extraction	DNA extraction from biological samples.	Cancer: 2 * 7h Rare disease: 7h	Technician + QiaSymphony or Maxwell (automata)
Qualification	Evaluation of the DNA samples quality and quantity.	Cancer: 6h Rare disease: 11h	Technician
Library Preparation	Fractionning and preparation of DNA strands.	Cancer: 16h Rare disease: 11h	Cancer: technician Rare disease: technician + 1 MicroLabStar
Sequencing Preparation	Preparation of the sequencers.	Cancer: 4h Rare disease: 4h + 4h	Technician + 1 sequencer
Sequencing	Sequencing of samples.	Cancer: 44h Rare disease: 14 + 44h	Technician + 1 sequencer
Buffer	Save sequencing results in a memory	1s	Informatics resources
Copies	Copy the sequencing twice on the server (archive and analysis).	triangular(180, 240, 300) (min)	Informatics resources (based on usage data)
Demultiplexing	Recreates the sample's sequence.	triangular(120, 180, 240) (min)	Informatics resources
Sequence Alignment	Sequence alignment using custom algorithms.	triangular(1440, 1620, 1800) (min)	Informatics resources
Run QC couv	Quality control step 1.	30 min	Informatics resources + bio-informatician
Scan QC couv	Quality control step 2.	30 min	Informatics resources + bio-informatician
ID checks	Identityvigilance.	30 min	Informatics resources + bio-informatician
Variants Call	Automated identification of gene variants.	triangular(300, 390, 480) (min)	Informatics resources
Annotation	Automated addition of comments on identified variants of interests.	triangular(120, 180, 240) (min)	Informatics resources
Sorting	Sorting variants of interests.	triangular(5,10,15) (min)	Informatics resources
QC 2	2nd quality control.	30 min	Informatics resources

## REFERENCES

- Berglund, E., G. Barbany, C. Orsmark-Pietras, L. Fogelstrand, J. Abrahamsson, I. Golovleva, H. Hallböök, M. Höglund, V. Lazarevic, L.-A. Levin, J. Nordlund, U. Norèn-Nyström, J. Palle, T. Thangavelu, L. Palmqvist, V. Wirta, L. Cavalier, T. Fioretos, and R. Rosenquist. 2022. "A Study Protocol For Validation And Implementation Of Whole-Genome And -Transcriptome Sequencing As A Comprehensive Precision Diagnostic Test In Acute Leukemias". *Frontiers in Medicine* 9:12.
- Boycott, K. M., A. Rath, J. X. Chong, and et al. 2017, May. "International Cooperation To Enable The Diagnosis Of All Rare Genetic Diseases". *The American Journal of Human Genetics* 100(5):695–705. Publisher: Elsevier.
- Brailsford, S., and N. Hilton. 2001. "A Comparison Of Discrete Event Simulation And System Dynamics For Modelling Healthcare Systems". In *Planning for the Future: Health Service Quality and Emergency Accessibility.*, edited by J. Riley, pp 18–39. Operational Research Applied to Health Services (ORAHs).
- Cuppen, E., O. Elemento, R. Rosenquist, S. Nikic, M. IJerman, I. D. Zaleski, G. Frederix, L.-r. Levin, C. G. Mullighan, R. Buettner, T. J. Pugh, S. Grimmond, C. Caldas, F. Andre, I. Custers, E. Campo, H. van Snellenberg, A. Schuh,

- H. Nakagawa, C. von Kalle, T. Haferlach, S. Fröhling, and V. Jobanputra. 2022. "Implementation Of Whole-Genome And Transcriptome Sequencing Into Clinical Cancer Care". *JCO Precision Oncology* 6:e2200245. PMID: 36480778.
- Degeling, K., H. Koffijberg, and M. J. IJzerman. 2017, January. "A Systematic Review And Checklist Presenting The Main Challenges For Health Economic Modeling In Personalized Medicine: Towards Implementing Patient-Level Models". *Expert Review of Pharmacoeconomics & Outcomes Research* 17(1):17–25.
- Eddy, D., W. Hollingworth, J. Caro, J. Tsevat, K. McDonald, and J. Wong. 2012. "Model Transparency And Validation: A Report Of The ISPOR-SMDM Modeling Good Research Practices Task Force-7". *Value in Health* 15(6):843–850.
- Glock, C. H., and E. H. Grosse. 2015, November. "Decision Support Models For Production Ramp-Up: A Systematic Literature Review". *International Journal of Production Research* 53(21):6637–6651.
- Hayeems, R. Z., F. Bernier, K. M. Boycott, T. Hartley, C. Michaels-Igbokwe, and D. A. Marshall. 2022, October. "Positioning Whole Exome Sequencing In The Diagnostic Pathway For Rare Disease To Optimise Utility: A Protocol For An Observational Cohort Study And An Economic Evaluation". *BMJ Open* 12(10):e061468.
- Incerti, D., X.-M. Xu, J. W. Chou, N. Gonzaludo, J. W. Belmont, and B. E. Schroeder. 2022, January. "Cost-Effectiveness Of Genome Sequencing For Diagnosing Patients With Undiagnosed Rare Genetic Diseases". *Genetics in Medicine* 24(1):109–118.
- International Human Genome Sequencing Consortium 2001, February. "Initial Sequencing And Analysis Of The Human Genome". *Nature* 409(6822):860–921. Number: 6822 Publisher: Nature Publishing Group.
- Lethimonnier, F., and Y. Levy. 2018, April. "Genomic Medicine France 2025". *Annals of Oncology: Official Journal of the European Society for Medical Oncology* 29(4):783–784.
- Li, C., S. Vandersluis, C. Holubowich, W. J. Ungar, E. S. Goh, K. M. Boycott, N. Sikich, I. Dhalla, and V. Ng. 2021, March. "Cost-Effectiveness Of Genome-Wide Sequencing For Unexplained Developmental Disabilities And Multiple Congenital Anomalies". *Genetics in Medicine* 23(3):451–460.
- Lote, R., E. J. Williams, and O. M. Ülgen. 2009. "Simulation Of Medical Laboratory Operations To Achieve Optimal Resource Allocation.". In *ECMS 2009 Proceedings*, edited by J. Otamendi, A. Bargiela, J. L. Montes, and L. M. D. Pedrera, 249–255. European Council for Modeling and Simulation.
- Marshall, D. A., L. R. Grazziotin, D. A. Regier, S. Wordsworth, J. Buchanan, K. Phillips, and M. IJzerman. 2020, May. "Addressing Challenges Of Economic Evaluation In Precision Medicine Using Dynamic Simulation Modeling". *Value in Health* 23(5):566–573.
- PFMG 2025 2017. "France Médecine Génomique 2025". <https://pfm2025.aviesan.fr/>. online, accessed 22-03-2023.
- Pleasant, E., A. Bohm, L. M. Williamson, J. M. T. Nelson, Y. Shen, M. Bonakdar, E. Titmuss, V. Csizmok, K. Wee, S. Hosseinzadeh, and et al. 2022, September. "Whole-Genome And Transcriptome Analysis Enhances Precision Cancer Treatment Options". *Annals of Oncology* 33(9):939–949. Publisher: Elsevier.
- Salleh, S., P. Thokala, A. Brennan, R. Hughes, and A. Booth. 2017, May. "Simulation Modelling In Healthcare: An Umbrella Review Of Systematic Literature Reviews". *Pharmacoeconomics* 35:937–949.
- Sanlaville, D., M. Vidaud, F. Nowak, and F. Lethimonnier. 2021, December. "Plan France Médecine Génomique 2025 : La France Entre Dans L'ère De La Médecine Génomique". *La revue du praticien* 71(10):1061–4.
- Sargent, R. G. 2013, February. "Verification And Validation Of Simulation Models". *Journal of Simulation* 7(1):12–24.
- Soltanolkotabi, M., H. A. Khorshidi, and M. J. IJzerman. 2022, October. "Modeling of Whole Genomic Sequencing Implementation Using System Dynamics And Game Theory". In *Proceedings of the IISE Annual Conference & Expo 2022*, 1061–4. Institute of Industrial and Systems Engineers: arXiv. arXiv:2210.01551 [cs, q-bio].
- Stranneheim, H., K. Lagerstedt-Robinson, M. Magnusson, M. Kvarnung, and et al. 2021, March. "Integration Of Whole Genome Sequencing Into A Healthcare Setting: High Diagnostic Rates Across Multiple Clinical Entities In 3219 Rare Disease Patients". *Genome Medicine* 13(1):40.
- The 100,000 Genomes Project Pilot Investigators 2021. "100,000 Genomes Pilot On Rare-Disease Diagnosis In Health Care — Preliminary Report". *New England Journal of Medicine* 385(20):1868–1880. PMID: 34758253.
- Tsai, E. R., A. N. Tintu, D. Demirtas, R. J. Boucherie, R. De Jonge, and Y. B. De Rijke. 2019, October. "A Critical Review Of Laboratory Performance Indicators". *Critical Reviews in Clinical Laboratory Sciences* 56(7):458–471.
- van de Ven, M., M. IJzerman, V. Retèl, W. van Harten, and H. Koffijberg. 2022, March. "Developing a Dynamic Simulation Model To Support The Nationwide Implementation of Whole Genome Sequencing in Lung Cancer". *BMC Medical Research Methodology* 22(1):83.
- Vázquez-Serrano, J. I., R. E. Peimbert-García, and L. E. Cárdenas-Barrón. 2021, January. "Discrete-Event Simulation Modeling In Healthcare: A Comprehensive Review". *International Journal of Environmental Research and Public Health* 18(22):12262. Number: 22 Publisher: Multidisciplinary Digital Publishing Institute.
- Venter, J. C., M. D. Adams, E. W. Myers, and et al. 2001, February. "The Sequence Of The Human Genome". *Science* 291(5507):1304–1351. Publisher: American Association for the Advancement of Science.

- Vu, M., K. Degeling, M. Martyn, E. Lynch, B. Chong, C. Gaff, and M. J. IJzerman. 2021, April. "Evaluating The Resource Implications Of Different Service Delivery Models For Offering Additional Genomic Findings". *Genetics in Medicine* 23(4):606–613.
- Wrzeszczynski, K. O., V. Felice, A. Abhyankar, L. Kozon, H. Geiger, D. Manaa, F. London, D. Robinson, X. Fang, D. Lin, M. F. Lamendola-Essel, D. Khaira, E. Dikoglu, A.-K. Emde, N. Robine, M. Shah, K. Arora, O. Basturk, U. Bhanot, A. Kentsis, M. M. Mansukhani, G. Bhagat, and V. Jobanputra. 2018. "Analytical Validation Of Clinical Whole-Genome And Transcriptome Sequencing Of Patient-Derived Tumors For Reporting Targetable Variants In Cancer". *The Journal of Molecular Diagnostics* 20(6):822–835.
- Wu, Y., S. Balasubramaniam, R. Rius, D. R. Thorburn, J. Christodoulou, and I. Goranitis. 2022, May. "Genomic Sequencing For The Diagnosis Of Childhood Mitochondrial Disorders: A Health Economic Evaluation". *European Journal of Human Genetics* 30(5):577–586.
- Yang, T., T.-K. Wang, V. C. Li, and C.-L. Su. 2014, December. "The Optimization Of Total Laboratory Automation By Simulation Of A Pull-Strategy". *Journal of Medical Systems* 39(1):162.

## **AUTHOR BIOGRAPHIES**

**JULES LE LAY** is a research engineer at the Centre Léon Bérard. He holds a Ph.D in industrial engineering from the École Nationale Supérieure des Mines de Saint-Étienne, France. He received an MSc in Engineering from Mines Saint-Étienne, Auvergne-Rhône-Alpes in 2017. Currently, he is a research engineer at Centre Léon Bérard. His research interests include modeling, simulation and optimization of health care systems. His e-mail address is [jules.le-lay@emse.fr](mailto:jules.le-lay@emse.fr).

**VINCENT AUGUSTO** is a professor and director of Center for Biomedical and Healthcare Engineering - Mines Saint-Étienne since January 2020 (70 people). He holds a Ph.D in industrial engineering from École Nationale Supérieure des Mines de Saint-Étienne, France (2008) and a "Habilitation 'a Diriger des Recherches'" (2016) on Health Territories Engineering. His areas of research are industrial engineering techniques, such as simulation and optimization, applied to health care systems, health data and medical decision aids. He was a visiting professor at CIRRELT (Interuniversity Research Center on Business Networks, Logistics and Transport), Laval University, Québec, Canada in 2009 and 2015. He is co-responsible for the living lab MedTechLab, where approach design, experiments and innovation are combined with health care engineering. His e-mail address is [augusto@emse.fr](mailto:augusto@emse.fr).

**XAVIER BOUCHER** is a Professor in Industrial Management at Center for Biomedical and Healthcare Engineering – Mines Saint Etienne. His main research orientations are Territorial Healthcare Systems. Active member of several scientific societies in the field of Industrial Engineering (IFAC, IFIP, SAGIP, SOCOLNET), he has participated to several EU projects, with the role of Coordinator (CODEMO 2023-2026), scientific coordinator for Mines Saint-Etienne for FA 4.0 (2020-2023), DIGI-FoF (2019-2021), OMIKA2 (2015-2018). He is Operational Director for the industrial Chair CORENSTOCK (collaboration IMT-elm Leblanc Bosch Group, 2021-2024) and he has coordinated national research projects for Mines Saint Etienne (ANR PSS-EuroNetwork (2016-2018 MRSEI), ANR ServINNOV (2012-2015) and FUI projects). He is as Associated Editor for Journal of Decision Systems and member of Editorial Board for International Journal on Information Technology Management. His e-mail address is [boucher@emse.fr](mailto:boucher@emse.fr).

**LIONEL PERRIER** is responsible for the innovations and strategies unit in the Clinical Research Direction of the Cancer Centre Leon Berard, Lyon, France, and member of the Economic and Public Health Evaluation Committee (CEESP) of the French National Authority for Health (HAS). He holds a PhD in economics (2001) and qualified to supervise doctoral research (2010). As scientific coordinator, work package leader or scientist, he is in charge of the health economist part of numerous projects and solicited for teaching in this field amongst other in EM Lyon Business School and Centrale Lyon. He is also with Univ Lyon, Leon Bérard Cancer Centre, GATE UMR 5824, F69008 Lyon, France, as well as Human and Social Sciences Department, and the Ethics Focus Group of the Cancer Centre Leon Berard. His e-mail address is [lionel.perrier@lyon.unicancer.fr](mailto:lionel.perrier@lyon.unicancer.fr).

**XIAOLAN XIE** is a professor and head of the department of Healthcare Engineering of the Center for Biomedical and Healthcare Engineering - Mines Saint Etienne. His research interests include data analytics, design, planning and scheduling, supply chain optimization, and performance evaluation of health care and manufacturing systems. He is author/coauthor of 300+ publications, including 110+ journal articles and six books. Pr. Xie is a fellow of IEEE, the founding chair of the IEEE RAS-Technical Committee on Automation in Health Care Management and an IEEE RAS 2017 Distinguished Lecturer. He has been an editor/associate editor for IEEE Transactions on Automation Science & Engineering, IEEE Transactions on Automatic Control, IEEE Transactions on Robotics & Automation and the International Journal of Production Research. He has been a Guest Editor of various special issues on health care engineering and manufacturing systems. His e-mail address is [xie@emse.fr](mailto:xie@emse.fr).