# Rapid DNA forensics Process

UNIVERSITY OF SURREY

THE FACULTY OF ARTS AND SOCIAL SCIENCES

Surrey Business School

Academic Year 2023-2024

Submission deadline: 1600 BST, Monday 13th May 2024

| Student URN | MSc Programme |
|---|---|
| 6809743 | BUSINESS ANALYTICS |

# Table of Contents

# Introduction

The new rapid DNA forensic processes samples of DNA match faster than before. The police force aims to optimise the new rapid DNA process such that the processing time for a sample is less than 24 hours. In order to find bottlenecks and spaces for improvement we must first look into the individual stages of the DNA forensic process.

The new rapid DNA forensic process begins with a call to report a crime. Once a crime has been reported, a Crime Scene Investigator (CSI) team is sent to the scene of the crime. The CSI is trained to examine the crime scene, collect DNA samples and to input the data in a tablet to upload into the police force database. As soon as the samples are physically collected, a courier service that is hired by the police force comes to collect the samples from the CSI team. The courier team immediately delivers the samples to the DNA forensics lab. The lab technicians are responsible for preparing the samples to go into the new rapid DNA forensic machine. However, since the lab technicians have a range of tasks to preform, they do not begin the sample preparation as soon as they arrive to the lab.

As and when the lab technicians get to completing the sample preparation process the samples are stored in a refrigerator until there is a full batch of 8 samples ready to go into the DNA sequencing machine. The machine will run for exactly 2 hours and once the samples exit the machine, they will be ready for validation. Since the senior lab researcher, who is responsible for the validation process, take some time to get to the sample validation there will be further delays in the validation stage. When the senior lab researcher gets to the sample, they validate them one after the other. As the samples are validated, they are uploaded on the database to find a match which is immediate. The final stage is to match the sequenced DNAs against known profiles in the database, to obtain potential matches.

Conditions:

1. The CSI will attend the crime scene not before the crime is reported, and they will not move onto next crime scene until the uploading of the current sample collection is finished.
2. Lab technicians must handle several tasks prior to the sequencing process, and they would not start working on the samples until about five other tasks had been completed, each taking about 20 minutes. This restricts him to prepare only one sample at a time.
3. The Rapid DNA sequencer operates at full capacity only when all the eight slots are fully occupied.
4. Lab researcher will start the validation after finishing the sequencing run and can validate only one sample at a time.

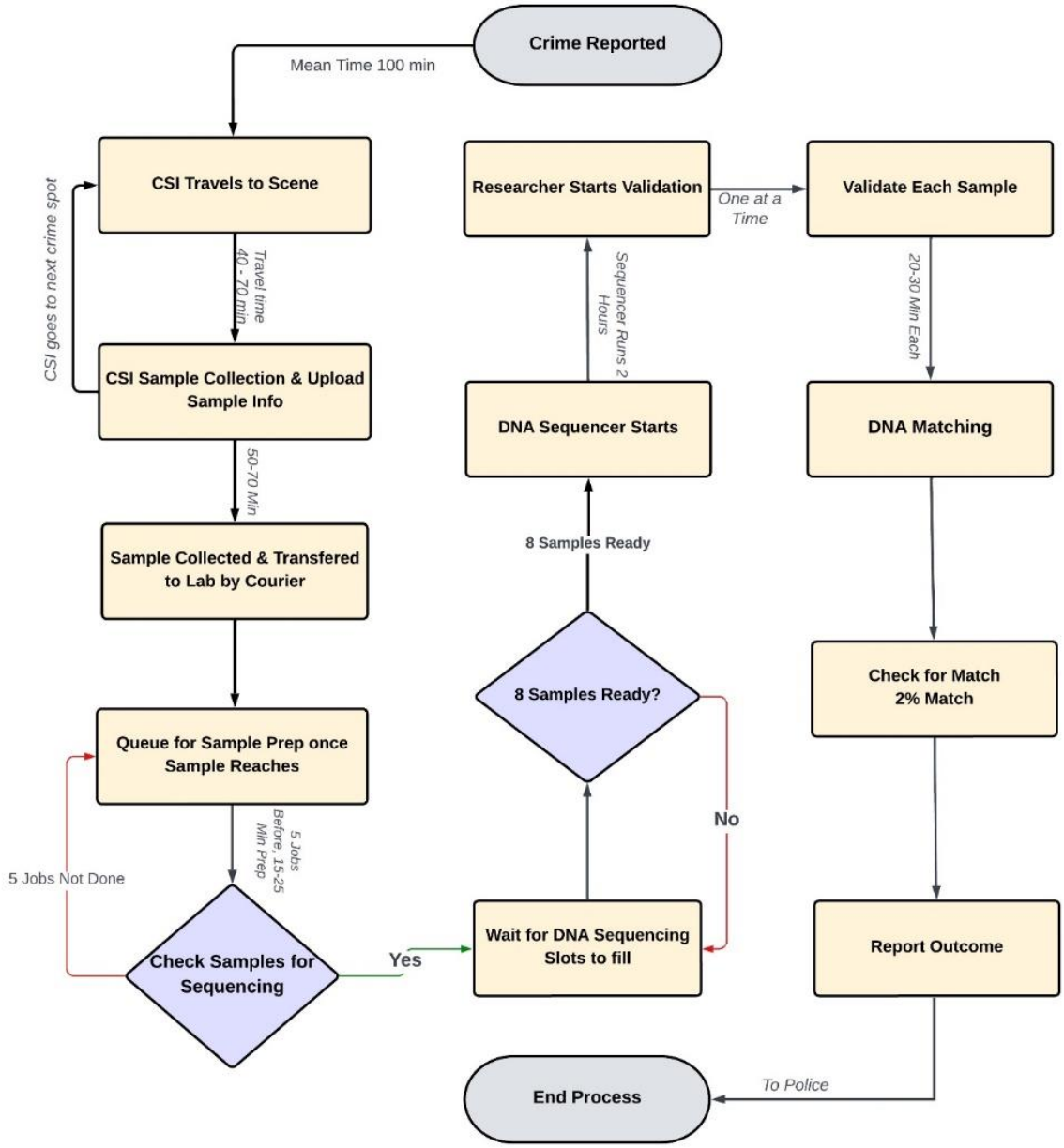


*Figure 1. Flow Diagram*

# Modelling

## Overview

Risk simulation model was carried out in 10,000 iterations with a fixed number seed of 9743 for proper convergence. Microsoft excel faciliated this Analysis supported by the @Risk 5.0.0 add-in provided by Palisade Corp.

The process starts with a crime reporting. Crimes are reported every 100 minutes, setting a consistent inter-arrival reporting time throughout the day irrespective of time of day and week.

$$t_1 \sim \ \text{RiskGamma}(1{,}100)\,\text{min} \rightarrow \ \text{RiskGamma}\left(1, \frac{1}{\lambda}\right)$$

A CSI is sent to the scene of the crime as soon as the crime is reported. The average travel time to the location is 50 minutes but it can take a minimum of 40 and a maximum of 70 minutes. A conditional aspect in this stage of the process is that the CSI cannot leave the present crime scene until they have taken samples and uploaded them to a tablet. The CSI will go straight to the crime scene if they are not at another crime scene. The queueing period here is the time between the reporting time of the crime and the minute that the CSI arrives to the crime scene.

$$t_2 \sim \text{RiskPert}(40{,}47.5{,}70)\text{min} \rightarrow \ \ \text{RiskPert}(\text{Min}, \text{ML}, \text{Max})$$

*Calculation of Most Likely ( Courier Collection )*

*They have given Mean value , so we have to find Most Likely.*

$$mew\ (\mu)\ = \frac{a+4b+c}{6}$$ *, Most Likely = 47.5*

Upon arrival, the CSI immediately starts collecting and uploading samples, which takes between 50 to 70 minutes. This stage is efficient, with no waiting time as the CSI begins processing right after reaching the scene.

$$t_3 \sim RiskUniform(50{,}70)\ \text{min} \rightarrow \ RiskUniform(Min, Max)$$

After sample collection is complete, the CSI hands the samples to a courier who delivers them to the Central Forensic Lab. This transition occurs immediately after uploading is finished, ensuring there is no queue.

$$t_4 \sim RiskWeibull(7{,}100)min \ \rightarrow RiskWeibull(\alpha, \beta)$$

The lab staff must finish procedures that take around 20 minutes per sample before beginning the sample preparation procedure. The five tasks are initiated by the lab technicians as soon as the samples arrive.

$$t_5 \sim RiskGamma(5{,}20)min \rightarrow RiskGamma\left(\alpha, \frac{1}{\lambda}\right)$$

Every sample requires 15 to 25 minutes to prepare. When they are brought to the lab technicians for the prerequisite tasks, the sample preparation time starts, and it finishes when both the prerequisite tasks and the sample preparation are completed. The time it takes the technician to finish the five tasks is the amount of time the samples must wait in queue for the technicians to finish the five tasks.

$$t_6 \sim \text{RiskUniform}(15{,}25)\text{min} \rightarrow \ \text{RiskUniform}(\text{Min}, \text{Max})$$

The rapid DNA machine only runs once 8 samples are ready to run in the sequencing machine. The machine operates on a set time for precisely 2 hours. The waiting of sample until there are 8 ready to go into the machine introduce the wait time in this stage. The queuing therefore is highest for the first sample of the batch and is zero for the 8th sample of the batch.

$$t_7 \sim 120\,\text{min} \rightarrow 2\ \text{Hours Constant}$$

The senior lab researcher spends 75 to 135 minutes and most likely 97 minutes to get to the validation process. This stage goes through samples in batches as they come out of the sequencer machine.

$$t_8 \sim \text{RiskPert}(75{,}97{,}135)\text{min} \rightarrow \ \ \text{RiskPert}(\text{Min}, \text{ML}, \text{Max})$$

Once the senior lab researcher gets to the validation process, they validate the samples one by one. The researcher does not waste any time between samples to validate them. This stage of the process takes between 20 and 30 minutes for each sample. The wait time for each sample in this stage is the time that sample is waiting to be validated after it comes out of the sequencing machine.

$$t_9 \sim RiskUniform(20{,}30)\,\text{min} \rightarrow \ RiskUniform(Min, Max)$$

After Validation , The Sequencer outputs are forwarded to the Integrated DNA Databse System, where each sample is compared against criminal profiles. Only about 2% of the samples match profiles within the database. This match shows the importance of maintaining high standards to ensure that each validated sample is as accurate as possible.

$$t_{10} \sim RiskBernoulli(0.02) \ \ \rightarrow \ RiskBernoulli(Probability)$$

## Method

| Parameter | Description | PDF | Mean | 5th %ile | 95th %ile |
|---|---|---|---|---|---|
| $t_1$ | Crime reported | RiskGamma(1,100) | 100 | 5.13 | 299.57 |
| $t_2$ | CSI to arrive | RiskPert(40,47.5,70) | 50 | 42.293 | 59.722 |
| $t_3$ | CSI collect & upload | RiskUniform(50,70) | 60 | 51 | 69 |
| $t_4$ | Courier company to deliver sample | RiskWeibull(7,100) | 93.544 | 65.422 | 116.969 |
| $t_5$ | Total Minutes for lab technician to complete 5 jobs before sample prep | RiskGamma(5,20) | 100.00 | 39.40 | 183.07 |
| $t_6$ | Total Minutes for sample preparation | RiskUniform(15,25) | 20 | 15.500 | 24.500 |
| $t_7$ | Total sequencing time | 120 | - | - | - |
| $t_8$ | Total Minutes for Researcher to start validation | RiskPert(75,97,135) | 99.667 | 82.422 | 119.055 |
| $t_9$ | Total Minutes for Sample validation | RiskUniform(20,30) | 25 | 20.500 | 29.500 |

*Table 1. Model Parameter Estimation. Units in minutes.*

$$\text{Overall Time in System} = t_{csi\ wait} + t_{sample\ collect\ service} + t_{courier\ service} + t_{sample\ prep\ service\ 2+wait} + t_{dna\ service+wait} + t_{validation\ service+wait}$$

# Results

## Average Model

1. Number of Samples processed by the end of seven days : 80 Samples (nearest ten)
   Number of Samples result in DNA match : 2 Samples
2. Mean time to process per sample: 1435.71
3. Percentage of samples that meet the police force ambition of processing samples in under a day : 51.25%
4. Yes, there are about four stage of process which looks under-resourced:
   a) **CSI Collection & Data Entry**: Due to the fact that it takes longer, CSIs now spend more time gathering evidence and entering data onto a tablet. It has the impact of increasing their time spent at a crime scene by about 20 minutes.
   b) **Lab Processing and Sample Prep**: For instance, they are reported to start preparing their samples only after they are done with the other tasks. In addition, technicians work on five other operations that take roughly twenty minutes each, so there will be a significant wait before even a sample is ready for sequencing.
   c) **New Rapid DNA Machine**: This machine has less holding capacity of 8 samples as compared to 100 in the older machines. On the other hand, the machine is only run when it obtains 8 samples at a time.

d) **CSI Validation Results**: The availability of only one senior lab researcher responsible for this task. Since each validation takes 20 to 30 minutes per sample, all the samples must wait in a queue for this one person to process them one at a time.

## Monte Carlo Simulation

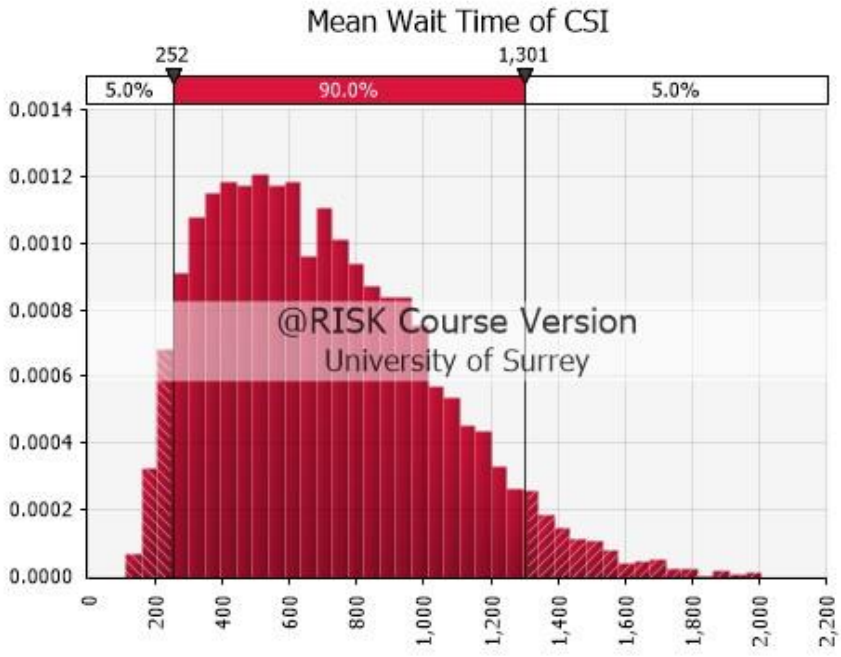


*Figure 2. The mean wait time for CSI to arrive.*

From simulating the mean wait time for CSI to arrive to the crime scene in figure 2, we observed that the distribution was positively skewed meaning that wait times were increasing over time. However, this stage between 252 and 1295 minutes. We also observed that 90% of the samples had a wait time in the median was roughly 653 minutes.

| **Stage of process** | **Mean** | **Median** | **5th %ile** | **95th %ile** |
|---|---|---|---|---|
| Mean time to wait for CSI to arrive | - | 653.55 | 252.93 | 1295.92 |
| Mean time to wait for sample preparation | 99.983 | | 91.821 | 108.467 |
| Mean time in queue for DNA sequencer | - | 397.07 | 368.26 | 449.84 |
| Mean time to wait from end of sequencing until senior lab researcher is ready to validate samples | 186.928 | - | 180.586 | 193.228 |
| Mean number of samples processed | 78.182 | - | 72.00 | 80.00 |
| Mean number of samples matched to the DNA database, in one week. | 1.5600 | - | 0.00 | 4.00 |
| Mean time to process a sample | - | 1658.61 | 1287.64 | 2295.78 |
| Percentage of samples that are processed in under 24 hours | - | 33.33 | 11.111 | 71.250 |

*Table 2. Summary of the times in each queue of the Rapid DNA forensics Process. Units in minutes.*

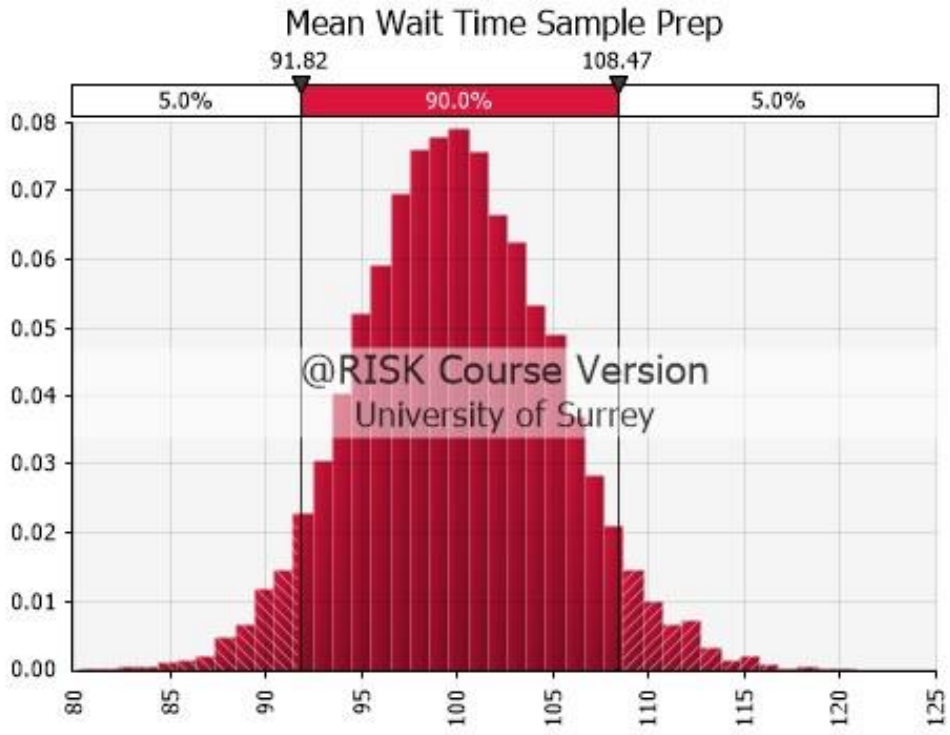


*Figure 3. The mean wait time for sample preparation.*

The second instance of queueing occurred in the sample preparation stage as shown in figure 3. This happened because the la technicians had to prioritize 5 tasks before they could get to the sample preparation. The mean wait time I this stage was around 100 minutes with the 5$^{th}$ percentile being 91 minutes and the 95% percentile being 108 minutes. The narrow distribution tells us that the range for the wait time is fairly small which leads us to believe that the wait time is predictable.

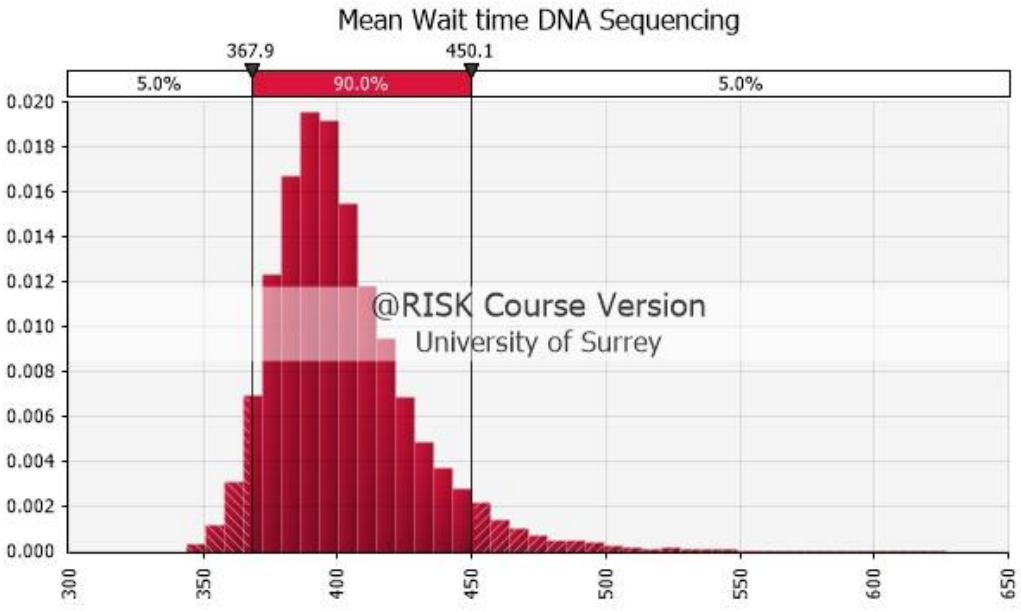


*Figure 4. The mean wait time for the DNA sequencing*

The third wait time in our model was in the DNA Sequencing stage as the samples were waiting to go into the sequencing machine in batches of 8 as shown in figure 4. The distribution had a thin right tail and had a median of 397 minutes and 90% of the of the per sample wait time in this stage was between 368 and 449 minutes.

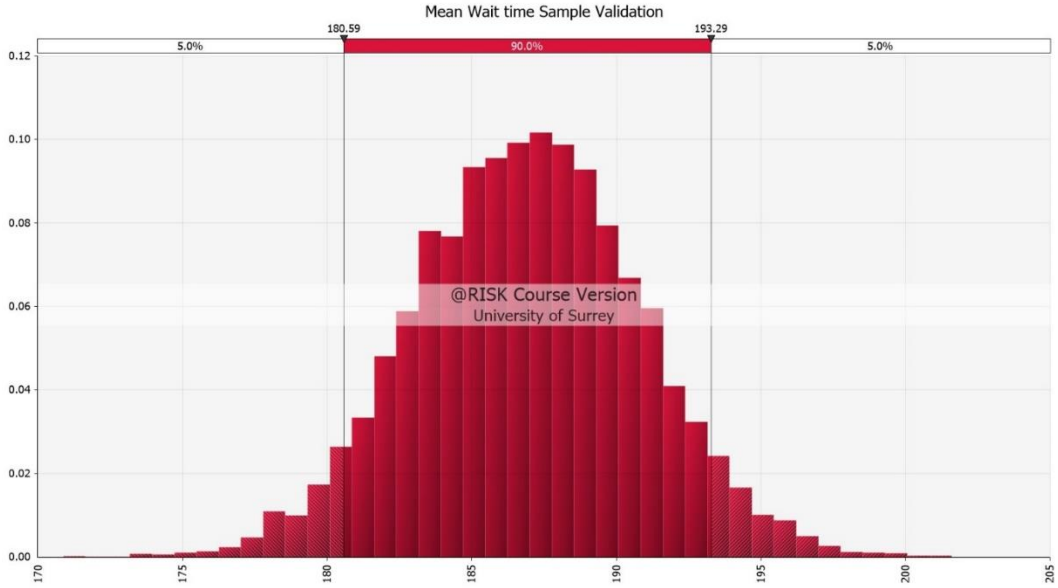


*Figure 5. The mean wait time for sample validation.*

The last wait time in the DNA forensics process was in the sample validation stage as shown in figure 5. This was because the sample were validated one after another by one senior lab researcher as they finished sequencing. The distribution for the wait time resembled a normal distribution with a mean time of around 186 minutes and 90% of the wait times lied within the range of 180 and 193 minutes. This means that the wait times were roughly consistent for each batch.

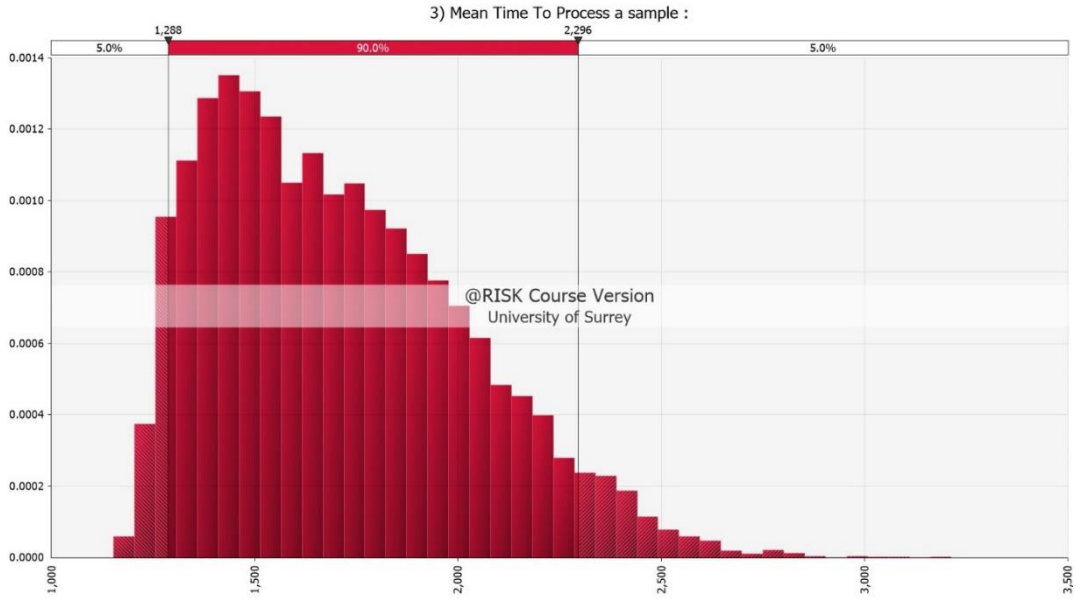


*Figure 6. mean time to process a sample*

The police forces ambition was to process a sample within 24 hours. With the current Rapid DNA process, the mean time to process a sample had a median of 1658 minutes as seen in figure 6. The distribution was observed to have a slight right skew with 95% of the mean time to process samples was between 1287 and 2295 minutes.

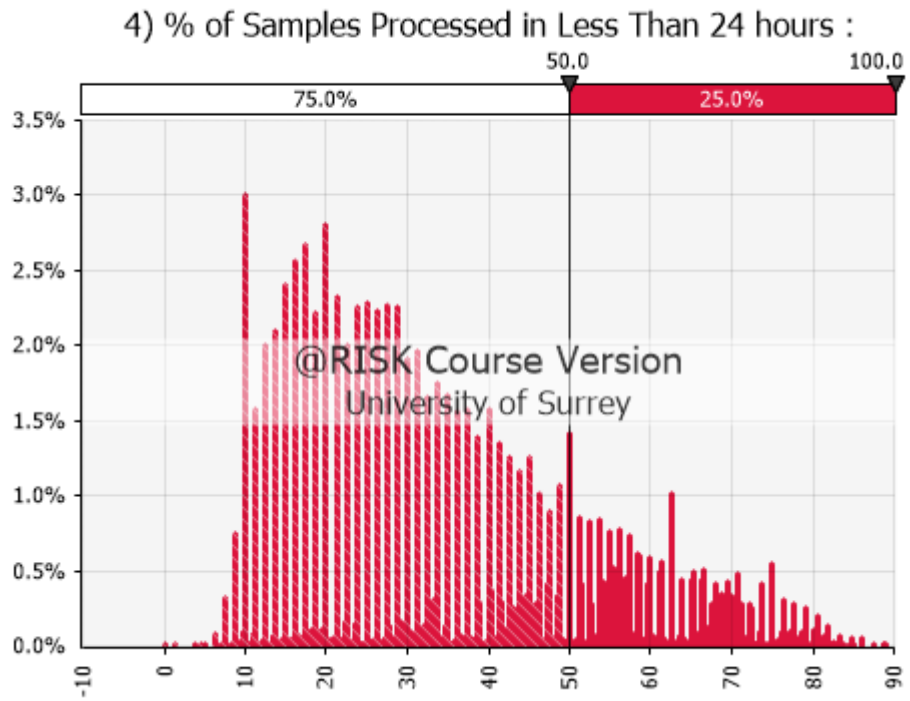


*Figure 7. percentage of samples processed under 24 hours - confidence*

Figure 7 shows the distribution for the samples processed in under 24 hours. The mean percentage of samples processed within 24 hours was 36 samples. Furthermore, we can also observe from figure 7 that we can say with 25% confidence that we can process half of the samples within 2 hours. From the distribution we can also say with 90% confidence that 11% and 71% of our samples within a week with process within 24 hours.

## Sensitivity Analysis

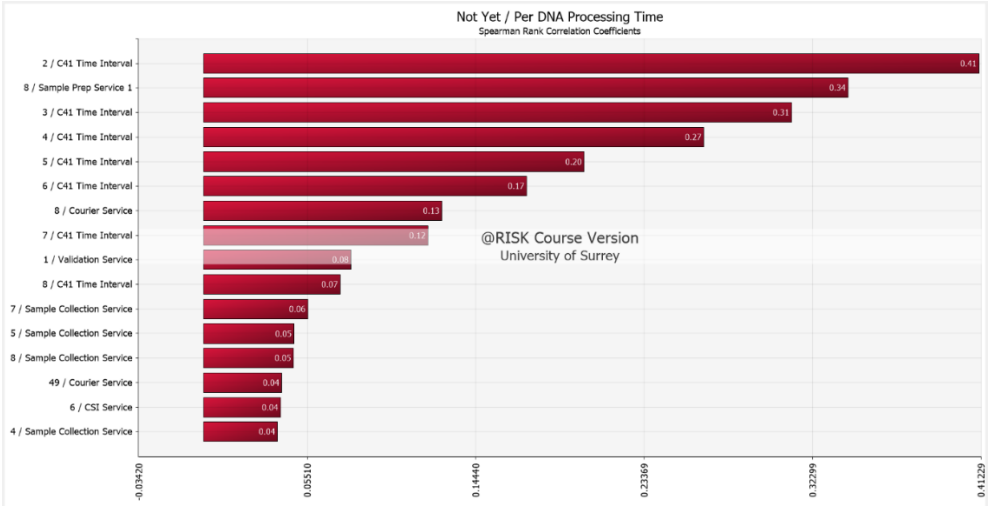


*Figure 8. Sensitivity Analysis*

In our analysis of reducing the processing time per sample to less than 24 hours, it is helpful to explore the strength of correlation between service times of different stages of the DNA forensics process. We ran a Spearman Rank Correlation Coefficient chart as seen in figure 8 which allows us to see which service times have the highest impacts on the total processing time per sample.

The sensitivity analysis shows us that the time with the highest impact on the per sample processing time was inter-arrival time (0.41) which is the time between crimes that is reported. It is essential to look into times that we have a degree of control over. Since inter-arrival time of crimes is not something we can change, we can move forward to the other stages in the figure. The service time with the second highest correlation with total processing time is sample prep stage (0.34). The next highest correlated service time was Courier collection service (0.13).

This analysis highlights parts in the forensics process that can be optimized to strategically reduce the sample processing times. We can intervene in sample preparation stage and the courier collection stage to implement improvements.

### Ad HOC Analysis

In the previous section, it was confirmed that certain stages have a high impact on the overall processing time. These stages were mainly the waiting times with high mean that caused bottlenecks in the process. Improving these queuing stages to lower the waiting time can notably improve the overall processing time. As the police's ambition is to reduce the processing time to be under 24 hours, it is crucial that the unnecessary waiting time is reduced or eliminated. The following what-if scenarios are created to make strategic adjustments to address the bottlenecks in the process.

## What If Scenarios

| Scenario | Median (Old) | Median (New) | 5th ile (Old) | 5th ile 2 (New) | 95th ile 1 (Old) | 95th ile 2 (New) |
|---|---|---|---|---|---|---|
| **Queue for Sample Prep** | 33.75 | 37.5 | 11.25 | 12.5 | 71.23 | 76 |
| **CSI Sample Collection** | 33.75 | 46.98 | 11.11 | 13.63 | 71.25 | 80 |
| **Start Validation** | 33.75 | 35 | 10.84 | 11.25 | 71.42 | 73.68 |

*Table 3. What if Scenarios in percentages*

### What if 1: Queue for Sample Preparation

By adding a trainee technician to the team, the workflow has been reorganised so that the trainee is allotted two jobs while the main technician is assigned the remaining three. This allocation of tasks aims to enhance the percentage of samples processed within 24 hours. Essentially, the update intends to accelerate processing time by incorporating more hands in operations, potentially minimising bottlenecks and increasing total throughput. We can observe this in figure 9 where there has been an improvement in median percentage of samples processed under 24 hours from 33.75% to 37.5%. This improvement tells us that hiring the trainee will reduce our total processing time in our system and bring us closer to our aim of processing more sample within 24 hours.

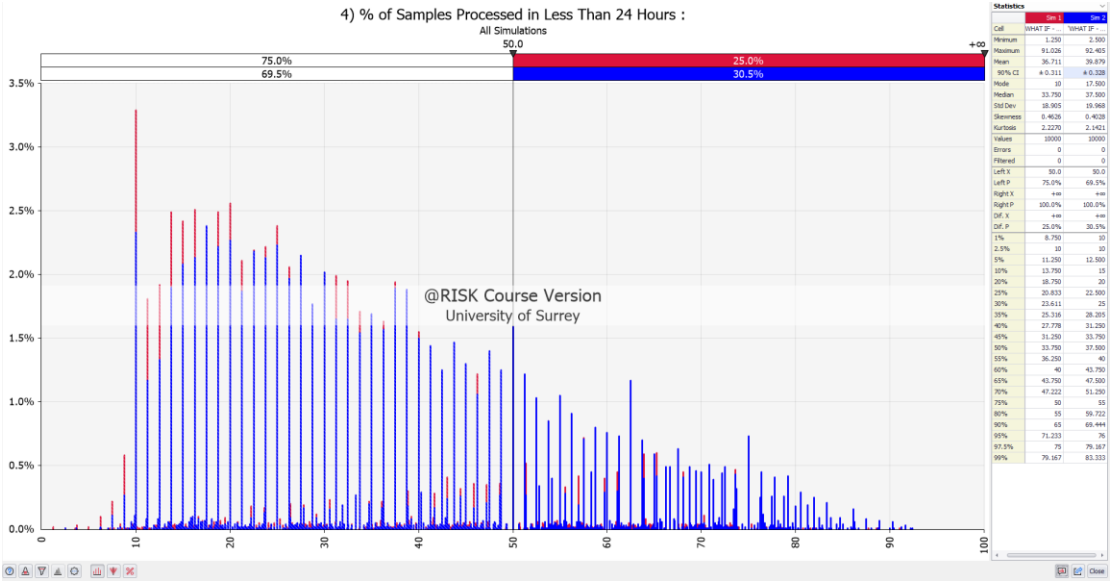


*Figure 9. What If Scenario 1 – Queue for sample preparation*

## What if 2: CSI Sample Collection

For the second scenario, we are aiming to reduce the highly correlated CSI sample collection process time as shown in figure 10. The voice to text technology has been previously implemented in the police report writing process (Wilis, Hidayatulah, Affan, & Dessyanto BP, 2020). Integrating this technology can help the CSI to dictate their observations and complete detailed reports in less time and reduce the physical labour of typing manually. The implementation of this technology has shown to increase the confidence by 20.3%.

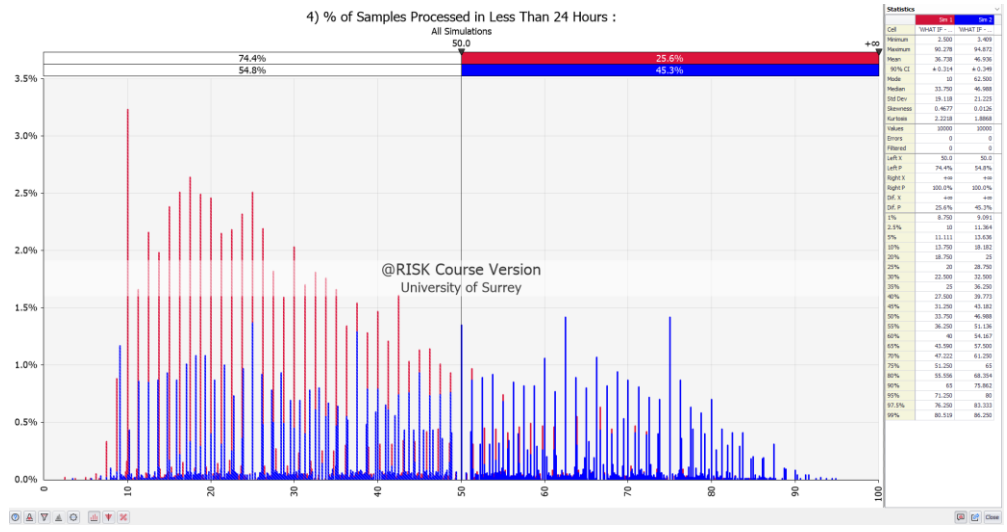


*Figure 10. What If Scenario 2 – CSI Sample collection*

## What if 3: Start Validation

For the final scenario, we are proposing a simple and cost-efficient integration of the sequencing machine's schedule linked to the senior lab researcher's calendar. This system automatically notifies the researcher when a sequencing begins, effectively scheduling the event directly into their agenda. This technology assigns priority labels to different tasks and makes sure the important ones are done on time and it pushes down the tasks with less priority. This is anticipated to reduce the waiting time

before validation, reducing it by 20%. We can say with 28% confidence that at least half of the samples in the week will be processed within 24 hours which is 2.25% (median) improvement from the original process.

## Insights and recommendations

### Enhancing CSI Sample Collection with Voice-to-Text Technology

Integrating voice-to-text technology into CSI sample collection processes represents a significant advancement in reducing manual reporting burdens. By automating the transcription of observations and findings, this technology not only accelerates documentation but also minimizes human errors associated with manual data entry. This improvement ensures a quicker turnaround from sample collection to analysis, enhancing the overall efficiency of the investigative process.

### Workforce Optimization

Incorporating a technician trainee into the sample preparation workflow allows for an effective redistribution of tasks among team members. This strategic role redefinition not only facilitates on-the-job training for new technicians but also optimizes the workflow by increasing the throughput of sample processing. As a result, the number of samples processed daily increases, significantly reducing waiting times and thereby streamlining the entire validation process.

### Streamlining Schedules for Efficient Sample Process:

The sequencing machines operational schedules with the calendar of the senior lab researcher optimizes both time management and resource allocation. This ensures that sample sequencing and subsequent analyses are conducted promptly upon completion, thereby minimizing downtime and expediting the validation process. Such strategic coordination is crucial for maintaining continuous workflow and prioritizing tasks effectively, which helps in reducing the overall cycle time of sample processing.

## References

Wilis, K., Hidayatulah, H., Affan, H. N., & Dessyanto BP, D. B. (2020). Summarization of Speech to Text from Reporter in Police Office with Latent Semantic Analysis (LSA) Method. *International Journal of Control and Automation, 13*(2), 933--943.