



UFOP

Universidade Federal
de Ouro Preto

UNIVERSIDADE FEDERAL DE OURO PRETO
ESCOLA DE MINAS
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO



MATHEUS CORREIA FREITAS SOUZA

**SISTEMA DE GESTÃO DE FILAS VIRTUAIS EM PROCESSOS PÚBLICOS
HOSPITALARES**

OURO PRETO, MG
2026

MATHEUS CORREIA FREITAS SOUZA

**SISTEMA DE GESTÃO DE FILAS VIRTUAIS EM PROCESSOS PÚBLICOS
HOSPITALARES**

Monografia apresentada ao Curso de Engenharia de Produção da Universidade Federal de Ouro Preto como parte dos requisitos necessários para a obtenção de Grau de Engenheiro de Produção.

Orientadora: Irce Fernandes
Gomes Guimarães

**OURO PRETO, MG
2026**

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

S729s Souza, Matheus Correia Freitas.
Sistema de gestão de filas virtuais em processos públicos
hospitalares. [manuscrito] / Matheus Correia Freitas Souza. - 2026.
61 f.

Orientadora: Profa. Dra. Irce Guimarães.
Monografia (Bacharelado), Universidade Federal de Ouro Preto.
Escola de Minas. Graduação em Engenharia de Produção .

1. Engenharia de produção - Arena (Software). 2. Simulação
(Computadores). 3. Administração dos serviços de saúde. 4. Unidade
Pronto Atendimento (UPA). I. Guimarães, Irce. II. Universidade Federal de
Ouro Preto. III. Título.

CDU 658.5

Bibliotecário(a) Responsável: Maristela Sanches Lima Mesquita - CRB-1716



FOLHA DE APROVAÇÃO

Matheus Correia Freitas Souza

Sistema de Gestão de Filas Virtuais em Processos Públicos Hospitalares

Monografia apresentada ao Curso de Engenharia de Produção da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Engenheiro de Produção

Aprovada em 10 de março de 2026

Membros da banca

Dra- Irce Fernandes Gomes Guimarães- Orientadora (Universidade Federal de Ouro Preto- UFOP)
Dra - Clarisse da Silva Vieira Camelo de Souza - (Universidade Federal de Ouro Preto- UFOP)
Mestranda e Engenheira Laura Eliza Ferreira Silva - (Universidade Federal de Ouro Preto- UFOP)

Dra- Irce Fernandes Gomes Guimarães, orientadora do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 25/03/2026



Documento assinado eletronicamente por **Irce Fernandes Gomes Guimaraes, PROFESSOR DE MAGISTERIO SUPERIOR**, em 26/03/2026, às 21:48, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1081807** e o código CRC **DE448BB5**.

Dedico este trabalho à minha família, especialmente à minha mãe, ao meu pai e aos meus irmãos, Felipe e Vinícius, pelo apoio e incentivo constantes. À minha namorada, Clara, pela compreensão e companheirismo ao longo desta trajetória. Dedico também, à República Alforria, que me proporcionou experiências e aprendizados inesquecíveis.

AGRADECIMENTO

Primeiramente, agradeço a Deus por me conceder perseverança, força e sabedoria durante essa caminhada.

À minha mãe, Madalena e ao meu pai, Marcorélio, por todo apoio, dedicação, amor e incentivo constantes, que foram essenciais não só para a realização deste trabalho, mas também para toda a minha vida. Aos meus irmãos, Felipe e Vinícius, pela parceria e amizade.

À Clara, pelo companheirismo, compreensão e apoio durante todo esse período.

À minha orientadora, Irce, pela orientação minuciosa, pela paciência e pelas contribuições fundamentais ao desenvolvimento desta pesquisa.

Por fim, agradeço a toda a minha família, a República Alforria e a todos os meus amigos, que sempre estiveram ao meu lado, oferecendo suporte, carinho e incentivo durante todas as etapas da minha formação.

RESUMO

As filas são parte do cotidiano das pessoas, sendo encontradas em diversos contextos de sua rotina, entretanto, através da teoria das filas é possível entender e modelar a situação para que os usuários aguardem menos tempo para serem atendidos. Neste contexto, o objetivo do estudo é avaliar alguns métodos de gestão de filas virtuais para sistemas públicos de saúde buscando reduzir o tempo de espera físico, otimizar os recursos e apresentar uma melhor experiência ao paciente. Para isso, inicialmente desenvolveu-se um referencial teórico para proporcionar todos os conceitos técnicos necessários, iniciando na gestão de setor hospitalar público abordando a teletriagem e a simulação computacional. Em seguida, foi realizado sobre a teoria das filas, as filas em ambientes hospitalares e as filas virtuais para viabilizar a implementação de um sistema de triagem e atendimento remoto em casos não urgentes. Após a revisão de literatura, foi desenvolvido o estudo de caso na Unidade de Pronto Atendimento de Ouro Preto - MG, em que avaliou o tempo de espera em filas no atendimento presencial e no atendimento remoto. Após a implementação dos modelos simulados, concluiu-se que a simulação computacional possui mais vantagens quando comparado com outras ferramentas por apresentar a visualização de comportamento do sistema de forma fácil, com simulações de diferentes cenários rápida e flexível, proporcionando amplo relatório de dados de saída para validação dos cenários propostos do sistema.

Palavras-chave: Arena. Ferramentas. Simulação. Tempo Atendimento. Unidade Pronto Atendimento.

ABSTRACT

Queues are part of people's daily lives, found in various contexts of their routine. However, through queuing theory, it is possible to understand and model the situation in order to reduce users waiting time. In this context, the objective of this study is to evaluate some methods of managing virtual queues for public health systems, seeking to reduce physical waiting time, optimize resources, and provide a better patient experience. To this end, a theoretical framework was initially developed to provide all the necessary technical concepts, starting with the management of the public hospital sector, addressing teletriage and computer simulation. Subsequently, a literature review on queuing theory, hospital queues, and virtual queues was conducted to support the implementation of a remote triage and care system for non-urgent cases. After the literature review, a case study was developed at the Emergency Care Unit of Ouro Preto - MG, where waiting times were evaluated for both in-person and remote care services. After implementing the simulated models, it was concluded that computer simulation has more advantages compared to other tools because it presents an easy visualization of system behavior, with quick and flexible simulations of different scenarios, providing a comprehensive output data report for validating the proposed system scenarios.

Keywords: Arena. Tools. Simulation. Service Time. Emergency Care Unit.

LISTA DE FIGURAS

Figura 1 – Estrutura básica de um sistema de filas M/M/1	27
Figura 2 – Simbologia das características da notação de Kendall	28
Figura 3 – Protocolo Manchester de atendimento	32
Figura 4 – Etapas de um estudo de simulação.....	37
Figura 5 – Unidade de Pronto Atendimento (UPA) Dom Orione, Ouro Preto/MG	41
Figura 6 – As principais etapas de um fluxograma.....	43
Figura 7 – Fluxograma com duas possibilidades de filas	44
Figura 8 – Modelo ARENA	48
Figura 9 – Segundo modelo ARENA.	50
Figura 10 – Entity	52
Figura 11 – Entity: Parte 2.....	55

LISTA DE TABELAS

Tabela 1 – Distribuição estatística dos processos em que os dados foram fornecidos por relatório de <i>Software</i>	39
Tabela 2 – Distribuição estatística dos processos em que os dados foram fornecidos por relatório de <i>Software</i>	39
Tabela 3 – Porcentagens e atributos correspondentes a cada cor de prioridade	47
Tabela 4 – Taxa de utilização de cada atendente.....	51
Tabela 5 – Tempo médio de espera de pacientes para cada atendente	51
Tabela 6 – Número médio de espera de pacientes para cada atendente.....	52
Tabela 7 – Tempo médio de espera de acordo com cada classificação.....	53
Tabela 8 – Taxa de utilização de cada atendente.....	53
Tabela 9 – Tempo médio de espera de pacientes para cada atendente	54
Tabela 10 – Número médio de espera de pacientes para cada atendente.....	54
Tabela 11 – Tempo médio de espera de acordo com cada classificação.....	56
Tabela 12 – Tempo médio de espera de acordo com a classificação de risco na UPA convencional e com triagem <i>online</i>	56
Tabela 13 – Taxa de utilização dos servidores na UPA convencional e com triagem <i>online</i>	57

LISTA DE SIGLAS

FIFO – *First in First Out*

LIFO – *Last in first Out*

PO – Pesquisa Operacional

PRI - *Priority*

SIRO – *Served in a Random Order*

SUS – Sistema Único de Saúde

UPA – Unidade de Pronto Atendimento

SUMÁRIO

1	INTRODUÇÃO	14
1.1	OBJETIVOS	16
1.1.1	Objetivo geral	16
1.1.2	Objetivos específicos	16
1.2	JUSTIFICATIVA/RELEVÂNCIA	17
1.3	HIPÓTESES	17
1.4	ESTRUTURA DA MONOGRAFIA	18
2	REFERENCIAL TEÓRICO	20
2.1	GESTÃO DO SETOR HOSPITALAR PÚBLICO	20
2.1.1	Contribuições da Pesquisa Operacional na Gestão de Filas Hospitalares	21
2.1.2	Teletriagem e simulação computacional	22
2.2	TEORIA DE FILAS	25
2.2.1	Características das filas	26
2.2.2	Notação de Kendall	28
2.2.3	Notação de Kendall em pronto atendimento	28
2.2.4	Crterios de avaliação das filas	29
2.3	FILAS EM AMBIENTES HOSPITALARES	30
2.3.1	Teoria da filas em ambientes hospitalares	30
2.3.2	Protocolo de <i>Manchester</i>	31
2.3.3	Importância da gestão de filas	33
2.4	FILAS VIRTUAIS	34
2.4.1	Vantagens das filas virtuais	34
3	METODOLOGIA	36
3.1	REVISÃO DE LITERATURA	36
3.2	ESTUDO DE CASO	38
3.3	COLETA DE DADOS	38
4	RESULTADOS E DISCUSSÕES	41
4.1	LOCALIDADE DO ESTUDO	41
4.2	MAPEAMENTO DO PROCESSO	42

4.3	MODELOS SIMULADOS.....	46
4.4	RESULTADOS	51
4.4.1	Análise da simulação da UPA atualmente.....	51
4.4.2	Análise da simulação da UPA com aplicativo de triagem.....	53
4.5	COMPARAÇÃO DOS RESULTADOS OBTIDOS EM CADA MODELO.....	56
5	CONSIDERAÇÕES FINAIS.....	58
	REFERÊNCIAS.....	60

1. INTRODUÇÃO

As filas fazem parte do cotidiano das pessoas e podem ser encontradas em diversas situações e contextos, como por exemplo nos supermercados em que os clientes esperam para serem atendidos pelos caixas e em bancos, no qual é comum ver os clientes esperando em filas para realizar transações no caixa ou em outros serviços.

A teoria das filas, também conhecida como teoria de espera, é um campo de estudo da matemática aplicada e da engenharia de sistemas que busca analisar e aprimorar o comportamento das filas de espera. Essa teoria é utilizada para entender e modelar situações em que clientes aguardam por um serviço em sistemas como bancos, supermercados, aeroportos, entre outros.

De acordo com Oliveira et al. (2017) a teoria das filas envolve o estudo de diversos elementos, como a taxa de chegada dos clientes, o tempo de serviço, o número de servidores disponíveis, a capacidade do sistema, a disciplina de atendimento (por exemplo, atendimento em ordem de chegada ou por prioridade), as características da população de clientes, entre outros fatores. O objetivo é analisar o desempenho do sistema, medindo métricas como o tempo médio de espera, o tempo médio de atendimento, a taxa de ocupação dos servidores, a probabilidade de não ser atendido imediatamente, entre outros indicadores.

Dito isto, uma forma de otimizar o processo das filas seria a utilização das filas virtuais. A gestão de filas virtuais é um método de organizar e gerenciar o fluxo de atendimento em um ambiente de forma digital. Essas filas são utilizadas principalmente em situações em que há uma demanda por serviços ou atendimento ao cliente que não pode ser facilmente acomodada simultaneamente.

Para implementar a gestão de filas virtuais, é necessário utilizar plataformas, redes digitais, aplicativos ou *software específico*, que podem oferecer recursos como agendamento de atendimento, notificações por SMS ou e-mail, integração com sistemas de gestão, entre outros. É importante escolher uma solução adequada às necessidades da empresa e que ofereça suporte técnico e atualizações regulares para garantir o bom funcionamento do sistema.

Dentre os inúmeros benefícios de se implantar uma ferramenta tecnológica adequada aos serviços públicos, tem-se a redução de custos sem queda na qualidade, maior produtividade nas atividades sem muito retrabalho e atendimento às diferentes demandas dos serviços de saúde em tempo ágil.

Um exemplo de empresa que adota o uso das filas virtuais é a *Apple*, que é conhecida por sua abordagem inovadora no atendimento ao cliente, e o uso de filas virtuais é um dos exemplos dessa estratégia. A empresa implementou um sistema de filas virtuais em suas lojas para melhorar a experiência de compra e evitar aglomerações.

A *Apple* introduziu o sistema de filas virtuais como parte de sua iniciativa "*Genius Bar*", que é o serviço de suporte técnico e atendimento ao cliente oferecido nas lojas da *Apple*. Ao visitar uma loja da *Apple*, os clientes podem reservar horários de atendimento por meio do aplicativo móvel *Apple Store* ou do site da empresa. Eles selecionam a loja desejada, escolhem o tipo de suporte necessário e agendam um horário conveniente para receber atendimento personalizado.

Essa abordagem elimina a necessidade de os clientes esperarem em filas físicas para obter suporte técnico ou assistência na loja. Em vez disso, eles podem fazer outras atividades ou explorar a loja enquanto aguardam seu horário de atendimento. O sistema de filas virtuais da *Apple* permite que os clientes otimizem seu tempo e recebam atendimento individualizado de acordo com suas necessidades.

Diante desse cenário, observa-se que a organização do fluxo de atendimento em unidades públicas de saúde representa um desafio relevante, principalmente em função da elevada demanda e da limitação de recursos disponíveis. Em ambientes hospitalares, o tempo de espera pode variar conforme fatores como a gravidade dos casos, a disponibilidade de profissionais e a forma como o sistema de atendimento é estruturado, sendo comum a utilização de protocolos de classificação de risco, como o Protocolo de Manchester, para priorização dos pacientes.

Nesse contexto, o uso de tecnologias digitais aplicadas à gestão de filas surge como uma alternativa capaz de contribuir para a melhoria da organização do atendimento, permitindo antecipar etapas do processo e reduzir a permanência desnecessária dos pacientes na unidade de saúde. A adoção de sistemas de triagem online e filas virtuais pode favorecer uma melhor distribuição da demanda, além de auxiliar na utilização mais eficiente dos recursos disponíveis.

Dessa forma, este trabalho busca analisar a aplicação de métodos de gestão de filas virtuais em sistemas públicos hospitalares, utilizando ferramentas de simulação computacional para avaliar diferentes cenários de atendimento. Assim, a pesquisa procura responder à seguinte pergunta: Quais tecnologias e metodologias podem auxiliar as instituições públicas de saúde a minimizar a espera em filas em ambientes de atendimento de pacientes em postos de saúde e hospitais?

1.1. OBJETIVOS

Este estudo estabelece, a seguir, o objetivo geral e os objetivos específicos que orientam o desenvolvimento da pesquisa e a análise dos resultados obtidos.

1.1.1. Objetivo Geral

O objetivo geral deste estudo é apresentar diretrizes para a implementação de alguns métodos de gestão de filas virtuais para sistemas públicos de saúde, visando a redução do tempo de espera físico, otimização de recursos e uma melhor experiência para o paciente.

1.1.2. Objetivos específicos

- Diagnosticar os modelos de gestão de filas hospitalares físicas e virtuais;
- Identificar a importância da minimização das filas nos hospitais públicos para os casos graves e não graves;
- Apontar quais métodos aplicados de gestão de filas virtuais podem ser aplicados nos hospitais públicos;
- Simular um sistema de triagem *online* para casos não graves;
- Analisar e simular os resultados por meio de casos reais e compará-los às situações do estudo.

1.2. JUSTIFICATIVA/RELEVÂNCIA

A oportunidade de se fazer um estudo explorando as tecnologias e metodologias de atendimento à saúde adotadas pelas instituições durante o período pandêmico, bem como a utilização destes recursos para atender as solicitações de serviços de saúde por meio de aplicativos e *delivery* formalizou o interesse para a elaboração deste estudo na área de Engenharia de Produção. Notou-se, também, a viabilidade de elaborar um projeto de pesquisa com ênfase na Teoria de Filas e Simulação para minimizar o tempo de espera de pacientes em locais de atendimento e cuidados da saúde.

Visto o tempo médio de espera nos hospitais públicos e postos de saúde e a relativa demora para atendimento dos pacientes, o objetivo desta pesquisa é buscar soluções para a minimização das filas presenciais para os casos não graves (principalmente na classificação azul e verde, de acordo com o protocolo de *Manchester*) utilizando o banco de dados do hospital para proporcionar melhor conforto para as pessoas por terem prioridade baixa em relação aos outros pacientes de caso clínico mais preocupante.

O objetivo é antecipar o sistema de triagem para esses casos de forma *online*, para que a pessoa saiba quando está chegando a sua vez de ser atendida sem precisar estar presente no local. Além disso, espera-se que, com o uso da capacidade e agilidade do sistema, os agentes de saúde poderão dar uma atenção ainda maior para os casos de maior prioridade.

1.3. HIPÓTESES

Considerando o problema identificado quanto ao tempo de espera em unidades de pronto atendimento e com base na literatura sobre teoria das filas, gestão de processos hospitalares e teletriagem, este estudo parte das seguintes hipóteses:

Hipótese 1: A implementação de um sistema de triagem *online* para pacientes classificados como casos não graves (cores verde e azul, segundo Protocolo de *Manchester*) pode contribuir para a redução do tempo de espera presencial na UPA Dom Orione.

Hipótese 2: A adoção de uma fila virtual pode contribuir para a melhoria da organização do fluxo de pacientes, proporcionando maior previsibilidade do atendimento e melhor experiência ao paciente, com redução do tempo de permanência física na unidade.

Hipótese 3: A introdução do sistema de triagem *online* pode otimizar a utilização dos servidores do sistema, reduzindo a sobrecarga operacional dos profissionais de saúde.

Hipótese 4: A aplicação da simulação computacional possibilita analisar cenários alternativos e apoiar a tomada de decisão sem interferir no funcionamento real da unidade.

Hipótese 5: A utilização da simulação computacional e de métodos de gestão de filas virtuais pode contribuir para o aprimoramento da organização sistêmica do atendimento em unidades públicas de saúde, favorecendo decisões mais eficientes na alocação da capacidade de atendimento.

1.4. ESTRUTURA DA MONOGRAFIA

Este trabalho está estruturado em cinco capítulos, organizados de forma a apresentar os fundamentos teóricos, a metodologia adotada, o desenvolvimento da pesquisa e a análise dos resultados obtidos.

O Capítulo 2 contempla o referencial teórico, no qual são discutidos os principais conceitos que fundamentam o trabalho, incluindo a gestão do setor hospitalar público, a teletriagem e a simulação computacional, os fundamentos da teoria das filas, suas características e critérios de avaliação, além da aplicação desses conceitos em ambientes hospitalares.

O Capítulo 3 descreve a metodologia científica adotada, detalhando o tipo de pesquisa, os procedimentos de revisão de literatura, o estudo de caso realizado, bem como os métodos de coleta e tratamento dos dados utilizados na modelagem e simulação do sistema estudado.

O Capítulo 4 apresenta os resultados e discussões da pesquisa. Inicialmente, é realizada a caracterização da Unidade de Pronto Atendimento (UPA) Dom Orione, situada no município de Ouro Preto – MG, destacando sua estrutura e funcionamento. Em seguida, é apresentado o mapeamento do processo atual da

unidade, com a descrição dos fluxogramas representativos do sistema convencional e do modelo proposto com a implementação da triagem online. Posteriormente, são detalhados os modelos simulados desenvolvidos no software Arena, incluindo a construção dos cenários analisados. Por fim, são expostos e discutidos os resultados obtidos por meio das simulações, abrangendo a análise das taxas de utilização dos recursos, tempos médios de espera, comparação entre os cenários simulados e avaliação do impacto da implementação do sistema de triagem online para casos não urgentes.

Por fim, o Capítulo 5 apresenta as considerações finais, destacando as principais conclusões do estudo, as contribuições da pesquisa e as possíveis implicações para a gestão de filas em unidades públicas de saúde.

2. REFERENCIAL TEÓRICO

O referencial teórico contempla temas relacionados à gestão do setor hospitalar público, com foco em ferramentas e estratégias voltadas à otimização dos serviços de saúde, bem como os fundamentos de teoria das filas, seus modelos matemáticos e critérios de avaliação. Desse modo, a organização deste capítulo está estruturada inicialmente com os aspectos relacionados à gestão hospitalar (2.1), seguida pela abordagem dos fundamentos da teoria das filas (2.2) e suas aplicações no contexto hospitalar (2.3), finalizando com a abordagem das filas virtuais (2.4).

2.1. GESTÃO DO SETOR HOSPITALAR PÚBLICO

A gestão do setor hospitalar público é marcada por desafios estruturais que impactam diretamente a eficiência operacional e a qualidade da assistência prestada aos pacientes. No contexto brasileiro, a elevada demanda por serviços de urgência e emergência, associada à limitação de recursos financeiros, humanos e estruturais, impõe constantes desafios à organização do fluxo de atendimento.

Estudos, como Manning e Islam (2023), apontam que o fluxo de pacientes em hospitais públicos é influenciado por fatores como variabilidade da demanda, restrições de capacidade, dificuldades na alocação de recursos e falhas na coordenação entre setores. Esses elementos contribuem para o congestionamento dos serviços e para o aumento do tempo de espera, especialmente em unidades de pronto atendimento e emergência.

Além disso, pesquisas indicam que modelos de gestão hospitalar pública frequentemente enfrentam problemas relacionados à centralização administrativa, fragmentação de processos e limitações na autonomia gerencial, fatores que dificultam a implementação de melhorias operacionais contínuas, destacam Campos e Rates (2024). A superlotação, nesse cenário, não se configura apenas como um problema de excesso de demanda, mas como reflexo de dificuldades sistêmicas na articulação entre oferta e necessidade de atendimento.

Diante dos desafios estruturais enfrentados pela gestão hospitalar pública, a Pesquisa Operacional (PO) destaca-se como abordagem metodológica capaz de oferecer suporte quantitativo à tomada de decisão. Ao incorporar modelos

matemáticos e estatísticos na análise dos processos assistenciais, a PO possibilita compreender o comportamento das filas e avaliar alternativas para melhoria do desempenho operacional.

2.1.1. Contribuições da Pesquisa Operacional na Gestão de Filas Hospitalares

Uma das contribuições centrais da Pesquisa Operacional na gestão de filas hospitalares refere-se ao dimensionamento da capacidade de atendimento. Por meio da modelagem matemática, é possível estimar níveis adequados de servidores, leitos e infraestrutura, considerando taxas de chegada e tempos médios de serviço. Conforme destaca Green (2006), pequenas variações na taxa de utilização podem provocar aumentos significativos no tempo de espera, especialmente quando o sistema opera próximo ao limite de capacidade. Assim, a PO permite identificar faixas operacionais mais estáveis e compatíveis com níveis aceitáveis de desempenho.

Outra contribuição relevante diz respeito à análise de políticas de atendimento e priorização. Sistemas hospitalares frequentemente operam com diferentes classes de pacientes, exigindo critérios de prioridade clínica. A PO possibilita avaliar o impacto dessas regras sobre indicadores como tempo médio de espera e nível de ocupação, permitindo equilibrar eficiência operacional e equidade no atendimento. De acordo com Hall (2013), a incorporação de critérios de priorização em modelos quantitativos é essencial para compreender os *trade-offs* entre desempenho global do sistema e atendimento a casos de maior gravidade.

A PO também contribui para o planejamento estratégico de redes hospitalares, ao integrar decisões de curto prazo (como escala de profissionais) com decisões estruturais (como ampliação de capacidade ou reorganização de fluxos). Revisões estruturadas da literatura, como Hulshof et al. (2012) indicam que métodos de PO vêm sendo amplamente aplicados em decisões de planejamento em saúde, abrangendo alocação de recursos, organização de fluxos e análise de desempenho sistêmico.

Além disso, a abordagem quantitativa proposta pela PO fornece suporte para avaliação de cenários sob incerteza, elemento inerente aos serviços de saúde. A consideração explícita da variabilidade nas chegadas e nos tempos de atendimento

permite antecipar situações de congestionamento e subsidiar decisões baseadas em evidências, fortalecendo a gestão hospitalar pública.

Dessa forma, a Pesquisa Operacional consolida-se como base conceitual e metodológica para a análise de sistemas de filas no setor hospitalar, estruturando a aplicação da teoria das filas e demais ferramentas quantitativas que serão aprofundadas nas seções subsequentes.

Entre essas ferramentas da PO, destacam-se a teletriagem e a simulação computacional, que têm sido apontadas na literatura como alternativas capazes de apoiar a tomada de decisão e contribuir para a melhoria do desempenho operacional em unidades de saúde.

2.1.2. Teletriagem e Simulação Computacional

De acordo com Sousa (2017), a partir dos anos 80, a teletriagem começou a ser utilizada por profissionais da área da saúde com o intuito de evitar o uso desnecessário do espaço físico hospitalar, permitindo que o ambiente hospitalar mantenha maior foco nos casos de urgência, além da redução de custos. Esse processo de triagem remota é geralmente realizado por meio de tecnologias de comunicação, como telefone, videoconferência ou aplicativos de mensagens.

A ideia principal dos serviços de teletriagem consiste em otimizar o uso dos recursos assistenciais, evitando que pacientes sem necessidade de atendimento imediato ocupem estruturas hospitalares destinadas a casos de maior complexidade, possibilitando que esses recebam orientação remota por meio de dispositivos móveis ou telefone. Em situações de emergência médica, por meio da interação telefônica, o paciente pode obter orientação quanto à gravidade de seus sintomas e receber recomendações sobre o próximo passo a ser tomado, seja orientação de autocuidado, agendamento de consulta, atendimento imediato ou acionamento de serviços de urgência. Segundo O’Cathain et al. (2004), serviços de triagem telefônica podem influenciar diretamente o comportamento de busca por atendimento, direcionando pacientes ao nível de cuidado mais adequado e contribuindo para a racionalização da demanda. Esse tipo de serviço mostra-se especialmente relevante em regiões com acesso limitado aos serviços de saúde ou em situações em que o deslocamento do paciente é dificultado.

De maneira geral, a teletriagem configura-se como ferramenta capaz de ampliar o acesso aos cuidados de saúde, agilizar o processo de classificação inicial e contribuir para que os recursos assistenciais sejam direcionados prioritariamente aos casos de maior gravidade.

No entanto, para que a teletriagem produza resultados efetivos, sua implementação deve seguir uma sequência estruturada de etapas. Inicialmente, torna-se necessária a definição de protocolos clínicos padronizados, estabelecendo critérios objetivos para classificação de risco e tomada de decisão. Conforme apontam Bunn, Byrne e Kendall (2005), sistemas de triagem baseados em protocolos estruturados contribuem para maior consistência clínica e redução de encaminhamentos inadequados aos serviços de urgência.

Após a definição dos protocolos, procede-se à organização do fluxo operacional do serviço, incluindo a capacitação dos profissionais responsáveis, a definição das ferramentas tecnológicas utilizadas e a integração com os demais níveis da rede de atenção à saúde. Durante o atendimento remoto, ocorre a coleta sistematizada de informações clínicas, análise dos sintomas relatados e identificação de possíveis sinais de gravidade. A partir dessa avaliação, realiza-se a classificação do nível de urgência e o direcionamento adequado do paciente, caracterizando a teletriagem como processo sequencial de avaliação, classificação e encaminhamento da demanda assistencial. De acordo com o Ministério da Saúde do Brasil (2012), a organização de serviços de Telessaúde no âmbito do SUS exige definição de fluxos assistenciais, qualificação das equipes e integração com a Rede de Atenção à Saúde, garantindo padronização do atendimento e continuidade do cuidado.

Para compreender os impactos dessa estratégia sobre o fluxo hospitalar, torna-se necessário analisar o sistema como um todo, considerando variáveis como demanda, capacidade e tempo de atendimento. Nesse contexto, a simulação computacional apresenta-se como ferramenta capaz de representar o funcionamento do sistema hospitalar e avaliar o comportamento operacional diante da implementação da teletriagem.

A simulação computacional é amplamente reconhecida como instrumento estratégico para modelagem e análise de sistemas complexos. Por meio da construção de modelos que representam o funcionamento de sistemas reais, é possível testar diferentes cenários, avaliar alternativas e antecipar resultados sem intervenções diretas no ambiente estudado, reduzindo custos e riscos operacionais,

conforme Freitas, Moser e Santos (2015). Além disso, destaca Gerônimo et al. (2018) que a simulação permite identificar gargalos, otimizar processos e apoiar decisões baseadas em evidências quantitativas.

Nos últimos anos, a aplicação da simulação tem se mostrado particularmente relevante na área da saúde, permitindo analisar operações hospitalares e validar estratégias de melhoria no fluxo assistencial, de acordo com Almeida et al. (2021). Segundo Banks et al. (2010), o desenvolvimento de um estudo de simulação inicia-se com a definição clara do problema e dos objetivos do modelo. Em seguida, realiza-se a coleta e análise dos dados do sistema real, incluindo taxa de chegada de pacientes, tempos médios de atendimento e capacidade instalada.

Posteriormente, elabora-se o modelo conceitual, que representa de forma simplificada os principais elementos e interações do sistema. Após essa etapa, o modelo é implementado em *software* específico e submetido aos processos de verificação e validação, fundamentais para garantir que o modelo represente adequadamente o sistema real. Conforme Law (2015), a validação é etapa crítica para assegurar credibilidade aos resultados obtidos por meio da simulação.

Além disso, a simulação de eventos discretos destaca-se como uma das abordagens mais utilizadas na modelagem de sistemas hospitalares, por permitir representar explicitamente a dinâmica de filas, alocação de recursos e variabilidade nos tempos de atendimento. Conforme ressaltam Kelton, Sadowski e Zupick (2015), essa técnica possibilita reproduzir o comportamento estocástico dos sistemas reais, incorporando distribuições probabilísticas para chegadas e atendimentos. Em ambientes hospitalares, essa característica é especialmente relevante, uma vez que a demanda apresenta elevada variabilidade e imprevisibilidade.

Adicionalmente, a simulação computacional contribui para a avaliação de políticas de priorização e dimensionamento de equipes, permitindo comparar cenários antes da implementação prática das mudanças. Robinson (2014) destaca que a principal vantagem da simulação está na possibilidade de experimentação virtual, na qual gestores podem testar intervenções sem comprometer o funcionamento do sistema real. No setor de saúde, essa capacidade é estratégica, pois decisões inadequadas podem impactar diretamente a qualidade do atendimento e a segurança do paciente.

No contexto do sistema de saúde, a construção de um modelo de simulação exige a identificação detalhada de variáveis operacionais e assistenciais. Entre os

principais elementos necessários estão a taxa de chegada de pacientes, a classificação de risco, os tempos médios de atendimento por categoria, a disponibilidade de profissionais, a capacidade física instalada e as regras de priorização entre setores. Além disso, deve-se considerar a interdependência entre etapas do atendimento, como recepção, triagem, consulta médica e eventual internação. Conforme destacam Jun, Jacobson e Swisher (1999), a modelagem de serviços de saúde demanda compreensão integrada dos processos clínicos e administrativos, dada a complexidade desses ambientes.

Dessa forma, a integração entre teletriagem e simulação computacional permite não apenas estruturar o atendimento remoto, mas também avaliar quantitativamente seus impactos sobre indicadores como tempo de espera, taxa de ocupação e utilização de recursos hospitalares. Essa abordagem fortalece a tomada de decisão baseada em evidências e contribui para maior eficiência operacional no setor hospitalar público.

2.2. TEORIA DE FILAS

Teoria da fila é considerado uma técnica de planejamento organizacional, permitindo aprimorar e melhorar os serviços, visto que as filas são inevitáveis. Atualmente, há sistemas rápidos que facilitam o manuseio e otimizam o tempo em que o cliente fica na fila. É importante ressaltar que as filas não são recentes, iniciando sobre o tema no início do século XX devido ao crescimento populacional acelerado em que o tempo é caro e custa dinheiro, tornando necessário que tudo seja rápido e eficiente, conforme Silva (2020).

A teoria das filas é baseada em modelos matemáticos e estatísticos que permitem prever o comportamento do sistema de filas em diferentes cenários. Esses modelos podem ser aplicados para encontrar soluções que melhorem a eficiência e o desempenho do atendimento, como dimensionar o número de servidores necessários, analisar o impacto de diferentes políticas de atendimento ou identificar gargalos e pontos de estrangulamento no sistema.

Segundo Macedo et al. (2017), a teoria das filas consiste na coleta e análise de dados, testando hipóteses sobre um sistema. Este modelo é utilizado como uma

das principais ferramentas da pesquisa operacional, para solucionar problemas de congestionamento e para minimizar os custos pela otimização do sistema estudado.

É importante ressaltar que as filas são voltadas diretamente ao atendimento e a qualidade de serviço fornecido, embora o tempo de espera possa variar muito conforme a qualidade do atendimento do servidor, a pesquisa operacional busca utilizar os recursos de forma mais adequada para resolver questões de gargalos e dimensionamentos, de acordo com Praia e Gomes (2013).

2.2.1. Características das filas

De acordo com Bruns et al. (2001), uma fila é caracterizada por um processo de chegadas (podendo este ser de diferentes entidades como pessoas, veículos etc.) até chegar ao sistema de atendimento, que pode ser formado por uma ou mais unidades de serviço.

Como cita Figueiredo e Rocha (2010), para modelar uma fila pode-se destacar 5 principais características:

1) Chegada de usuários: tempo entre chegadas de usuários ao local, o tempo pode ser determinístico ou uma variável aleatória, sendo que este último é o mais frequente. Através de estudos amostrais, pode-se descobrir se o processo de chegadas dos usuários funciona de acordo com alguma distribuição de probabilidades.

2) Serviço: É importante observar e estudar, em um sistema de filas, o número de clientes atendidos em um determinado tempo em cada posto de serviço, com o intuito de informar a distribuição de probabilidade de duração de cada atendimento.

3) Quantidade de atendentes disponíveis: Esse componente remete ao número de atendentes que realizam ao mesmo tempo o atendimento aos usuários.

4) Capacidade para atendimento dos usuários: essa capacidade refere-se ao número máximo de usuários permitidos dentro do estabelecimento estudado, ao mesmo tempo, levando em consideração também os que se encontram na fila de espera. O sistema pode ter capacidade finita ou infinita, na qual ele não irá possuir um limite de número permitido de usuários no estabelecimento.

5) Disciplina da Fila: Segundo Guimarães (2005) é necessário estabelecer uma lógica ordenada que determina como as pessoas do sistema irão utilizar os

serviços quando um servidor estiver livre. Logo, essa característica determina a ordem em que os usuários serão atendidos, podendo ter diferentes critérios, sendo os mais comuns:

- FIFO (*first in first out*): Esse método segue a ordem cronológica de chegadas, na qual o primeiro a chegar é o primeiro a ser atendido.

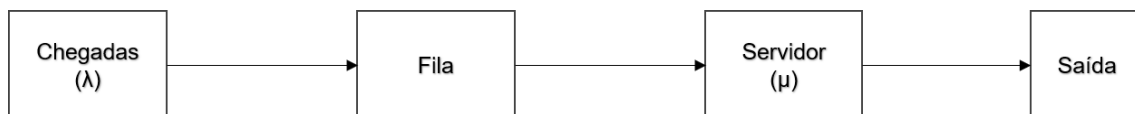
- LIFO (*last in first out*): Ao contrário do FIFO, neste caso o último a chegar é atendido primeiro.

- SIRO (*served in a random order*): é feito de forma aleatório a ordem de atendimento

- PRI (*priority*): é definida alguma prioridade para a ordem de atendimentos, com base em algum critério definido.

A estrutura básica de um sistema de filas pode ser representada de forma esquemática, evidenciando os principais elementos que compõem o modelo, conforme ilustrado na figura 1.

Figura 1 – Estrutura básica de um sistema de filas M/M/1



Fonte: Adaptado de Gross e Harris (1998).

Conforme apresentado na Figura 1, o sistema de filas é composto pelo processo de chegada dos usuários ao sistema, representado pela taxa λ , pela formação da fila de espera quando o servidor se encontra ocupado, pelo atendimento, representado pela taxa de serviço μ , e pela saída dos clientes após conclusão do atendimento. Esse modelo corresponde à estrutura clássica de um sistema M/M/1, amplamente utilizado na teoria das filas para análise de desempenho de sistemas de serviço, de acordo com Gross e Harris (1998) e Hillier e Lieberman (2013).

A representação esquemática do sistema permite compreender a dinâmica entre as taxas de chegada e de atendimento, possibilitando a análise de indicadores como tempo médio de espera, tamanho médio da fila e nível de utilização do servidor. Segundo Hillier e Lieberman (2013), o equilíbrio entre λ e μ é fundamental para garantir a estabilidade do sistema, evitando crescimento indefinido da fila e sobrecarga operacional. Dessa forma, a representação gráfica apresentada constitui

base conceitual para modelagens matemáticas e aplicações de simulação computacional em sistemas reais.

2.2.2. Notação de Kendall

A notação de Kendall fornece uma visão concisa da natureza física do sistema, seus processos e a disciplina de atendimento adotada. Isso permite a definição de modelos de sistemas de filas com base em características observadas. Cada modelo possui uma modelagem matemática específica para cada critério de desempenho, sendo crucial identificar inicialmente qual modelo de sistema está sendo estudado.

A notação de Kendall é descrita da seguinte forma: $A/B/c/K/m/Z$, na qual (A) indica a distribuição de tempo entre chegadas, (B) o tipo de distribuição de probabilidade do tempo de serviço, (c) é a capacidade de atendimento ou quantidade de atendentes, (K) é a capacidade máxima do sistema, (m) é o tamanho da população da qual vem cliente e (Z) é a disciplina de atendimento da fila.

Na Figura 2 são apresentados alguns símbolos padrões para estas características.

Figura 2 - Simbologia das características da notação de Kendall.

Característica	Símbolo	Descrição
(A) Distribuição de tempo entre chegadas	M	Lei de Poisson (exponencial)
	D	Determinístico
	E_k	Erlang com parâmetro k
	H_k	Hiperexponencial com parâmetro k
(B) Distribuição de tempo de serviço	M	Lei de Poisson (exponencial)
	D	Determinístico
	E_k	Erlang com parâmetro k
	H_k	Hiperexponencial com parâmetro k
(c) Número de canais de serviço	1,2,..., infinito	
(K) Capacidade do sistema	1,2,..., infinito	
(m) Tamanho da população	1,2,..., infinito	
(Z) Disciplina da fila	FIFO	Primeiro que chega, primeiro a ser atendido
	LIFO	Último que chega, primeiro a ser atendido
	PRI	Atendimento por prioridade
	SIRO	Atendimento de forma aleatória

Fonte: Marujo et al., 2010.

2.2.3. Notação de Kendall em pronto atendimento

As características típicas das filas de um pronto atendimento podem ser descritas usando a notação de *Kendall* da seguinte maneira:

Distribuição de Chegada (A): No pronto atendimento, a distribuição de chegada de pacientes geralmente segue um padrão irregular e imprevisível. Em alguns momentos, pode haver picos de demanda devido a emergências, enquanto em outros momentos o fluxo pode ser mais estável.

Distribuição de Atendimento (B): A distribuição de atendimento pode variar dependendo dos serviços oferecidos pelo pronto atendimento. Por exemplo, para consultas médicas simples, o tempo de atendimento pode seguir uma distribuição exponencial. No entanto, para procedimentos mais complexos ou exames, o tempo de atendimento pode ser mais variável e seguir uma distribuição diferente, como a distribuição de *Weibull*.

Número de Servidores (C): O número de servidores em um pronto atendimento pode variar de acordo com a capacidade e demanda do local. Pode haver várias áreas de atendimento com diferentes especialidades médicas, cada uma com seu próprio conjunto de servidores. Portanto, o número de servidores (ou canais de atendimento) pode ser variável e depende da infraestrutura e recursos disponíveis.

Assim, a notação de *Kendall* para descrever as filas em um pronto atendimento pode ser algo como "A/B/C", onde "A" descreve a distribuição de chegada dos pacientes, "B" descreve a distribuição de atendimento ou serviço, e "C" representa o número de servidores disponíveis para atender os pacientes.

2.2.4. Critérios de avaliação das filas

Os critérios de avaliação de filas são fundamentais para analisar o desempenho e a eficiência dos sistemas para várias situações, como atendimento ao cliente, processos de produção e operações logísticas. Segundo Torres (1966) e Marujo et al. (2010), para avaliar e administrar as filas, pode-se utilizar alguns critérios:

Tempo médio de espera: Através deste critério, pode-se avaliar o tempo médio que a entidade estudada espera para ser atendida ou processada na fila. Minimizar esse tempo é crucial para melhorar a satisfação do cliente e a eficiência do serviço.

Tamanho médio de fila: Este critério refere-se ao número médio de clientes ou elementos que estão na fila em um determinado momento. Se esse número for grande, indica que a fila está sobrecarregada e pode levar a tempos médios de espera mais longos. Portanto, controlar e manter um tamanho médio de fila adequado é essencial para manter a eficiência do sistema.

Lead time: O *lead time* é o tempo total necessário para que a entidade estudada atravesse todo o sistema de fila. Reduzir o lead time geralmente é um dos objetivos mais importantes para melhorar a eficiência e a agilidade do sistema de filas.

Taxa de chegada: Esse critério representa a taxa na qual a entidade estudada entra no sistema. Uma taxa de chegada alta pode sobrecarregar o sistema de fila, resultando em tempos médios de espera mais longos e tamanho médio de fila maior.

Taxa de atendimento: Essa taxa remete à velocidade que os elementos de uma fila são atendidos ou processados. Quanto maior for a taxa de atendimento, maior será a chance de garantir que a fila seja processada de maneira rápida e eficiente.

Layout: O *layout* representa o arranjo físico dos elementos no ambiente em que ocorre o processo de fila, podendo incluir balcões de atendimento, máquinas, instalações, pessoal da produção, entre outros aspectos espaciais.

2.3. FILAS EM AMBIENTES HOSPITALARES

A aplicação da teoria das filas no contexto hospitalar visa compreender e otimizar o fluxo de pacientes em unidades de saúde. Desta forma, esta seção discute a utilização do Protocolo de *Manchester* como instrumento de classificação de risco e a importância da gestão eficiente de filas para a melhoria da qualidade do atendimento e da eficiência operacional.

2.3.1. Teoria das filas em ambientes hospitalares

Um estudo feito por Orlandin et al. (2019) utilizou a teoria das filas em um ambulatório do Rio Grande do Sul, para dimensionar e analisar o seu sistema de atendimento, do ponto de vista operacional, com o intuito de propor melhorias para os serviços prestados. O estudo concluiu, através da análise feita pelo estudo de caso,

que a taxa de ocupação de 96,87% justifica as filas observadas durante alguns momentos do dia. Através de simulação, foi analisado que a taxa de ocupação cairia para 64,58% se o ambulatório contratasse um atendente a mais, o que aumentaria os custos, porém deixaria o sistema mais eficiente.

Galindo e Santos Junior (2012) apresentaram um estudo de caso sobre o processo de gestão das filas no Hospital de Nova Odessa, com o objetivo de reduzir o tempo de espera na fila para atendimento dos pacientes e melhorar o atendimento oferecido. Através do estudo feito e utilizando a teoria das filas e a teoria das restrições, o autor concluiu que seria importante implementar um departamento de triagem para guiar os pacientes para a especialidade médica necessária. Além disso, a pesquisa indicou que seria crucial a contratação de mais um médico para os plantões diurnos.

Pode-se concluir, através das pesquisas citadas, que a teoria das filas vem sendo estudada e aplicada como uma importante ferramenta para promover uma gestão mais eficiente e eficaz nos ambientes hospitalares, garantindo que os recursos sejam utilizados de forma mais eficaz para atender às necessidades dos pacientes com mais qualidade e rapidez.

2.3.2. Protocolo de *Manchester*

Geralmente para classificar a enfermidade do paciente, os sistemas hospitalares brasileiros, em especial as unidades de pronto atendimento, seguem o protocolo de *Manchester*. O tempo médio de espera nos hospitais públicos pode variar de acordo com alguns fatores, como a localização geográfica, a demanda de pacientes, a disponibilidade de recursos e a gravidade dos casos atendidos, que, de acordo com Teixeira, Oselame e Neves (2014) é uma ferramenta que indica a prioridade de urgência e determina a ordem de atendimento com base na urgência clínica do paciente.

O objetivo principal do protocolo de *Manchester* é garantir que os pacientes sejam atendidos de acordo com a gravidade de suas condições médicas, priorizando aqueles que necessitam de tratamento imediato. Ele baseia-se em um sistema de cores, que são atribuídas às diferentes categorias de pacientes, cada uma representando uma prioridade de atendimento. Essas cores são, do mais grave para o menos grave: Vermelho, laranja, amarelo, verde e azul. Ao chegar ao hospital, os

pacientes passam por uma triagem inicial, na qual são avaliados por um profissional de saúde treinado. Com base em sua condição médica e sintomas apresentados, o paciente é classificado em uma das categorias de cores denominadas pelo protocolo.

De acordo com Amaral (2017), segundo o sistema de classificação do protocolo de *Manchester*, o tempo de espera para a cor azul é de até 240 minutos, significando que se trata de um caso não urgente. A pulseira verde representa os quadros urgentes, no qual o tempo de espera é de até 120 minutos. A pulseira de cor amarela é dada para os pacientes de caso urgente, mas que podem aguardar o atendimento, seu tempo de espera é de até 60 minutos. A cor laranja significa que o paciente possui um quadro muito urgente e não pode demorar para ser atendido, o tempo de espera é de até 10 minutos. Por fim, a pulseira de cor vermelha é dada para o paciente que necessita de atendimento imediato, então não pode ter tempo de espera, como demonstrado na Figura 3.

Figura 3 - Protocolo *Manchester* de atendimento.



Fonte: BRASIL ESCOLA, 2024.

De acordo com o demonstrado na Figura 3, o prazo máximo de espera para atendimentos descreve-se conforme a classificação, no qual pode apresentar variações e atrasos nos casos de menor urgência, em que os atendimentos verde ou azul podem ser realizados atendimentos nas Unidades Básicas de Saúde próximas à residência.

Segundo Roncalli et al. (2017), o uso do protocolo de *Manchester* ajuda na seleção de fluxogramas com base nos sintomas que os pacientes estão sentindo, o que direciona o tempo de atendimento pela ordem de prioridade, através da escolha da cor das pulseiras de cada paciente. Além disso, o sistema de classificação de risco possibilita parâmetros de gestão para a unidade de pronto atendimento. Através da planilha de dados, o gestor terá acesso a informações sobre filas, tempo de espera dos pacientes, tempo de atendimento dos funcionários, entre outras informações, o que facilita o controle e a tomada de decisão.

2.3.3. Importância da gestão de filas

Pode-se ressaltar que a gestão de filas em pronto atendimentos desempenha um papel crucial na eficiência operacional, na qualidade do serviço prestado e na satisfação do paciente. Uma boa gestão de filas pode acarretar os seguintes resultados:

Tempo de Espera Reduzido: Os prontos atendimentos costumam lidar com uma grande variedade de casos, desde emergências graves até problemas de saúde menos urgentes. Uma gestão eficaz de filas pode ajudar a reduzir o tempo de espera dos pacientes, priorizando aqueles com condições mais graves e otimizando o fluxo de pacientes de acordo com a gravidade de suas condições.

Melhoria na Triagem e Priorização: A gestão de filas permite uma triagem mais eficiente dos pacientes, identificando rapidamente aqueles que necessitam de atendimento imediato. Isso ajuda a garantir que os recursos sejam alocados de maneira adequada e que os pacientes mais críticos recebam atenção prioritária.

Utilização Eficiente de Recursos: Os recursos em um pronto atendimento, como médicos, enfermeiros, equipamentos e salas de exames, são limitados. Uma gestão eficaz de filas ajuda a otimizar a utilização desses recursos, garantindo que sejam empregados de maneira eficiente e que não haja subutilização ou sobrecarga.

Redução do Estresse para Pacientes e Equipe: Esperar em filas pode ser estressante para os pacientes, especialmente quando estão lidando com problemas de saúde urgentes ou desconfortáveis. Uma gestão eficiente de filas pode ajudar a reduzir esse estresse, proporcionando uma experiência mais organizada e previsível. Além disso, uma gestão eficaz de filas pode contribuir para reduzir o estresse da

equipe médica, garantindo que eles possam trabalhar de maneira mais eficiente e focar nos pacientes que mais precisam de atenção.

Melhoria na Experiência do Paciente: Uma gestão eficaz de filas contribui para uma experiência geral mais positiva para o paciente. Pacientes que recebem atendimento rápido e eficiente tendem a estar mais satisfeitos com os serviços prestados e são mais propensos a retornar ao pronto atendimento se necessário, além de recomendar o serviço a outras pessoas.

Monitoramento e Melhoria Contínua: Ao implementar sistemas de gestão de filas, os prontos atendimentos podem monitorar o desempenho ao longo do tempo e identificar áreas de melhoria. Isso pode incluir ajustes nos processos, realocação de recursos ou investimentos em tecnologia para otimizar ainda mais o fluxo de pacientes e a qualidade do atendimento.

Em resumo, a gestão eficaz de filas desempenha um papel fundamental na garantia de um atendimento rápido, eficiente e de qualidade nos prontos atendimentos, contribuindo para a satisfação do paciente, a eficiência operacional e a qualidade geral dos serviços prestados.

2.4. FILAS VIRTUAIS

Na literatura existem várias propostas com a finalidade de reduzir os impactos e transtornos causados pelas filas, uma delas é a implementação da fila virtual, na qual os clientes/pacientes não precisam estar no mesmo espaço físico, podendo utilizar esse tempo para realizar outras atividades.

2.4.1. Vantagens das filas virtuais

Existem várias aplicações e benefícios da gestão de filas virtuais. Algumas das principais vantagens são apontadas por Pordeus (2021) e Senna et al. (2020), sendo elas:

- **Organização eficiente:** Com a gestão de filas virtuais, é possível organizar os clientes ou usuários de forma eficiente, garantindo que todos sejam atendidos de maneira justa e ordenada.

- Redução de tempo de espera: As filas virtuais ajudam a reduzir o tempo de espera, pois os clientes não precisam ficar fisicamente presentes em um local específico para aguardar o atendimento. Em vez disso, eles podem realizar outras atividades enquanto aguardam em uma fila virtual.

- Flexibilidade e conveniência: Com a gestão de filas virtuais, os usuários podem acessar a fila de qualquer lugar, por meio de dispositivos como computadores, smartphones ou tablets. Isso proporciona maior flexibilidade e conveniência, permitindo que os usuários aguardem o atendimento sem restrições de localização.

- Distribuição equitativa: As filas virtuais podem ser configuradas para distribuir os clientes ou usuários de forma equitativa, evitando favorecimentos e privilégios indevidos. Isso contribui para um atendimento mais justo e imparcial.

- Monitoramento e análise: As soluções de gestão de filas virtuais frequentemente oferecem recursos de monitoramento e análise, permitindo que as empresas coletem dados sobre o tempo médio de espera, taxa de atendimento, satisfação do cliente, entre outros indicadores. Essas informações podem ser usadas para otimizar os processos e melhorar a experiência do cliente.

3. METODOLOGIA

Para realizar os objetivos específicos dos estudos, foi inicialmente desenvolvido uma revisão bibliográfica utilizando as palavras-chaves “filas virtuais”, “filas” e “implantação em hospitais públicos”, para realizar a fundamentação teórica do estudo. Em seguida, será desenvolvido um estudo analisando o tempo médio de espera de pacientes que se localizam nos hospitais públicos de Ouro Preto – MG, além de simular um sistema de triagem *online* para avaliar os casos não graves.

Após desenvolver o estudo, será realizada a discussão dos dados obtidos avaliando-os com os dados existentes na literatura para validar se a implementação do novo processo de sistema de triagem *online* se demonstrou efetiva ao ser implementada em um hospital com alta lotação.

Diante disso, o método da pesquisa utilizado será o estudo bibliográfico, análise bibliométrica, caracterizada como do tipo exploratória, sendo utilizado como ambiente de estudo, o cenário real de um setor público de atendimento e cuidados para a saúde, podendo ser utilizados postos de saúde e/ou hospitais. Também serão utilizados os métodos de Simulação.

3.1. REVISÃO DE LITERATURA

Serão selecionados estudos anteriores relacionados ao tema da pesquisa, e realizada uma busca e filtragem de artigos científicos com base no alinhamento da pesquisa nas plataformas Google Acadêmico, *Science Academy*, *Microsoft Academic*, *SCIELO*, *Web of Science* e Portal CAPES.

No processo de filtragem/busca dos materiais científicos, os seguintes filtros serão considerados:

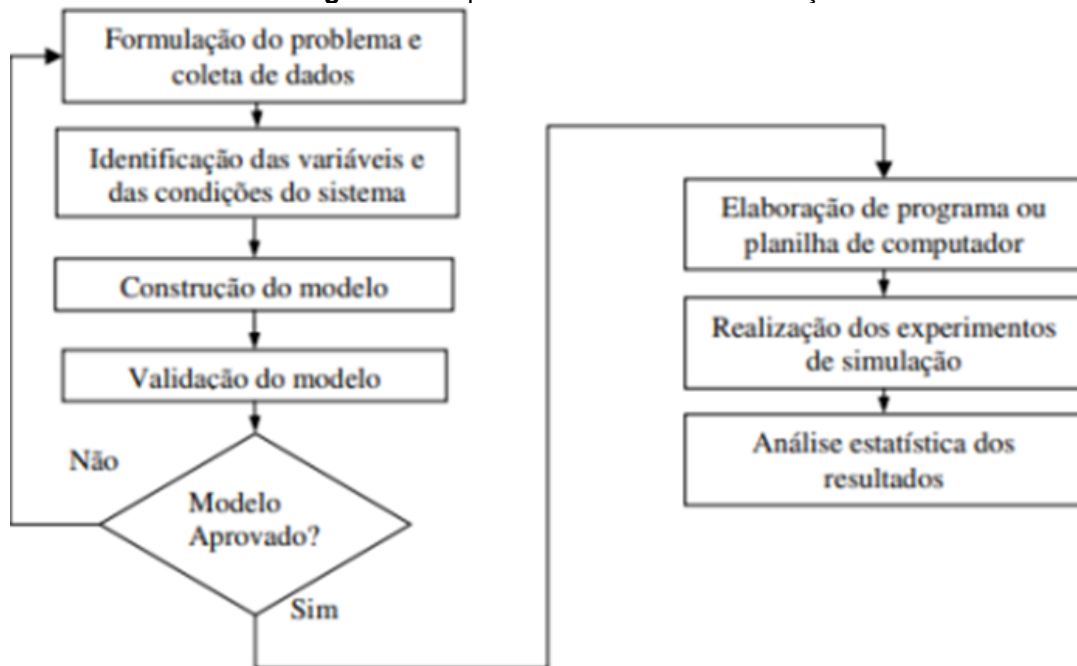
a. a seleção dos artigos, monografias, teses e dissertações nas bases de dados a partir das palavras chaves: Gestão de Filas, Sistema de saúde, Filas Virtuais, aplicativos. Simulação.

b. a filtragem dos materiais selecionados com base no ano de publicação dos estudos a escolha por pesquisa que possuem simulações e desenvolvimento de aplicativos.

c. alinhamento dos títulos e resumo com o tema

Na sequência propõe-se a utilização de um caso real para coletar e analisar dados que serão utilizados na construção de um modelo que será utilizado na Simulação Computacional, por meio de um *Software*, a fim de identificar os processos de filas virtuais e não virtuais, tempo de espera em fila e uso da capacidade e recursos. Na Figura 4 é possível observar as etapas metodológicas propostas para esse estudo. Ao final da simulação e análise bibliométrica, as diretrizes serão elaboradas.

Figura 4 - Etapas de um estudo de Simulação.



Fonte: Elaborado pelo autor, 2024.

3.2. ESTUDO DE CASO

Para coletar os dados, foi feito um estudo de caso realizado na Unidade de Pronto Atendimento (UPA) da cidade. Assim, foi realizada uma transferência de dados referente aos critérios de criticidade de cada atendimento associado ao sistema de triagem para fazer a validação, identificando e desenvolvendo o padrão de informações que ao serem inseridas no sistema gera uma lista de criticidade de acordo com o quadro clínico e os dados do paciente.

Outro ponto importante a se ressaltar é que durante o estudo foi desenvolvido uma quantificação referente ao tempo gasto do paciente desde o início que entra na recepção do hospital até sair atendido. Para isso, foi cronometrado o tempo de cada atendimento separado por setor de recepção, triagem e atendimento médico avaliando os dados de criticidade com o tempo de atendimento para validar se está sendo implementado adequadamente o Protocolo *Manchester*.

As ferramentas utilizadas inicialmente para validar a criticidade optou-se por utilizar o *software* ARENA, enquanto os dados temporizados foram tratados utilizando o Excel para fins analíticos e comparativos, possibilitando a criação de tabelas dinâmicas que auxiliam na análise.

3.3. COLETA DE DADOS

Para a obtenção dos dados utilizados neste estudo, foi necessário realizar solicitações formais à Secretaria Municipal de Saúde, a fim de obter autorização para acesso às informações da UPA. Após a aprovação da solicitação, foi realizada uma entrevista com a responsável pela coordenação da unidade, que autorizou e disponibilizou os dados utilizados na presente pesquisa.

Os dados coletados e fornecidos são referentes aos meses de junho, julho e setembro de 2024. Foi fornecido pela Unidade de Pronto Atendimento, através de um relatório gerado por um sistema que é utilizado pelos funcionários de Tecnologia da Informação do local, os seguintes dados: a quantidade de pacientes que consultaram nos meses, a chegada dos pacientes, a quantidade de pacientes conforme cada classificação do protocolo de *Manchester*, tempo de atendimento da

triagem, horário de chegada dos pacientes na recepção, horário de chegada dos pacientes no consultório e horário de saída dos pacientes do consultório.

Este relatório de triagem foi utilizado como base para calcular a porcentagem de pacientes de acordo com cada uma das cores do protocolo de *Manchester* – com o intuito de assegurar que a ordem de prioridade do protocolo fosse devidamente seguida na simulação. Através dos dados fornecidos de horário de chegada dos pacientes, foi possível calcular o intervalo entre as chegadas. Além disso, através dos dados de horário de chegada no consultório e horário de saída dos pacientes do consultório, foi possível calcular o tempo de atendimento clínico. O modelo também gerou a mediana do tempo de classificação de risco.

Tabela 1 - Distribuição estatística dos processos em que os dados foram fornecidos por relatório de *software*.

Expressões estatísticas	
Atendimento médico	-0.5 + LOGN (17.8, 21.4)
Tempo entre chegadas	-0.5 + LOGN (7.04, 9.2)
Tempo de triagem	Mediana (1.29)

Fonte: Elaborado pelo autor, 2024.

Devido à limitação do *software* em que ele apenas computa o paciente no sistema após o paciente passar pela recepção, o tempo de atendimento da recepção foi necessário coletar os dados pessoalmente. A coleta foi feita durante o mês de setembro, durante períodos aleatórios do dia, tanto noturno quanto diurno.

Tabela 2 - Distribuição estatística dos processos em que os dados foram fornecidos por relatório de *software*

Expressões estatísticas	
Atendimento da recepção	ERLA (0.308, 4)

Fonte: Elaborado pelo autor, 2024.

Após a coleta dos dados, foi realizado o tratamento estatístico das informações com o objetivo de obter distribuições probabilísticas que representassem de forma adequada o comportamento do sistema real. Para isso, utilizou-se o módulo *Input Analyzer*, disponível no *software Arena*, que permite ajustar distribuições estatísticas aos dados observados por meio de testes de aderência.

Foram analisados os tempos entre chegadas dos pacientes, tempos de atendimento nas diferentes etapas do processo e demais variáveis necessárias para

a construção do modelo de simulação. A partir desses dados, foram obtidas distribuições estatísticas que melhor representaram o comportamento do sistema, as quais estão apresentadas nas Tabelas 1 e 2, utilizadas como base para a parametrização do modelo computacional.

Conforme apresentado na Tabela 1, os dados de atendimento médico e tempo entre chegadas foram ajustados a distribuições do tipo lognormal, enquanto o tempo de triagem foi representado por meio da mediana observada. Já na Tabela 2, o tempo de atendimento da recepção foi ajustado a uma distribuição do tipo *Erlang*, obtida a partir dos dados coletados em campo. Essas distribuições foram definidas com auxílio do *Input Analyzer*, que compara diferentes modelos probabilísticos e indica aquele que apresenta melhor aderência aos dados observados.

As distribuições selecionadas foram então inseridas no modelo desenvolvido no Arena, permitindo que a simulação representasse de forma mais fiel o funcionamento real da unidade estudada. Esse procedimento é fundamental para garantir a consistência dos resultados obtidos, uma vez que a qualidade da simulação depende diretamente da adequação das distribuições utilizadas para representar os processos reais observados durante a coleta de dados.

4. RESULTADOS E DISCUSSÕES

4.1 LOCALIDADE DO ESTUDO

O local de estudo foi a UPA Dom Orione, situada na cidade de Ouro Preto, Minas Gerais, foi fundada em 03 de julho de 2020 com a missão de atender às necessidades de saúde da população, é um importante ponto de atendimento de urgência e emergência na região. A estrutura física da unidade pode ser visualizada na Figura 5.

Figura 5 – Unidade de Pronto Atendimento (UPA) Dom Orione, Ouro Preto/MG.



Fonte: Prefeitura Municipal de Ouro Preto, 2025.

Com uma grande equipe de profissionais de saúde, a UPA Dom Orione está estruturada para atender um volume significativo de casos. Conta em sua infraestrutura com 19 leitos de observação, 8 leitos para medicação rápida, consultórios de atendimento, sala de sutura e curativos, sala de coleta de material, sala de inalação, sala de eletrocardiograma, sala de aplicação de medicamentos, sala de exames de radiologia geral, sala de isolamento, além das áreas administrativas, de funcionários e apoio. A unidade opera 24 horas por dia, garantindo suporte contínuo à comunidade, o que a torna um local de estudo relevante para a pesquisa sobre eficiência do atendimento de saúde.

Neste trabalho, a UPA Dom Orione será analisada no contexto da implementação de um modelo de simulação que introduza um sistema de triagem *online* para casos não graves. A proposta visa melhorar a gestão do fluxo de pacientes, permitindo que aqueles com necessidades menos urgentes sejam triados e orientados de forma mais eficaz, reduzindo assim o tempo de espera e otimizando os recursos disponíveis.

A importância de um sistema de triagem *online* reside na possibilidade de direcionar o atendimento de maneira mais eficiente, garantindo que os casos mais críticos recebam a atenção necessária sem comprometer a qualidade do atendimento aos demais pacientes. Essa abordagem não busca apenas melhorar a experiência do paciente, mas também propõe um aprimoramento na organização interna da UPA, potencializando seu papel na saúde pública local.

Com essa perspectiva, a UPA Dom Orione se apresenta como um campo de pesquisa rico, onde é possível explorar soluções que podem impactar positivamente a saúde da comunidade e servir como modelo para outras unidades de saúde em contextos semelhantes.

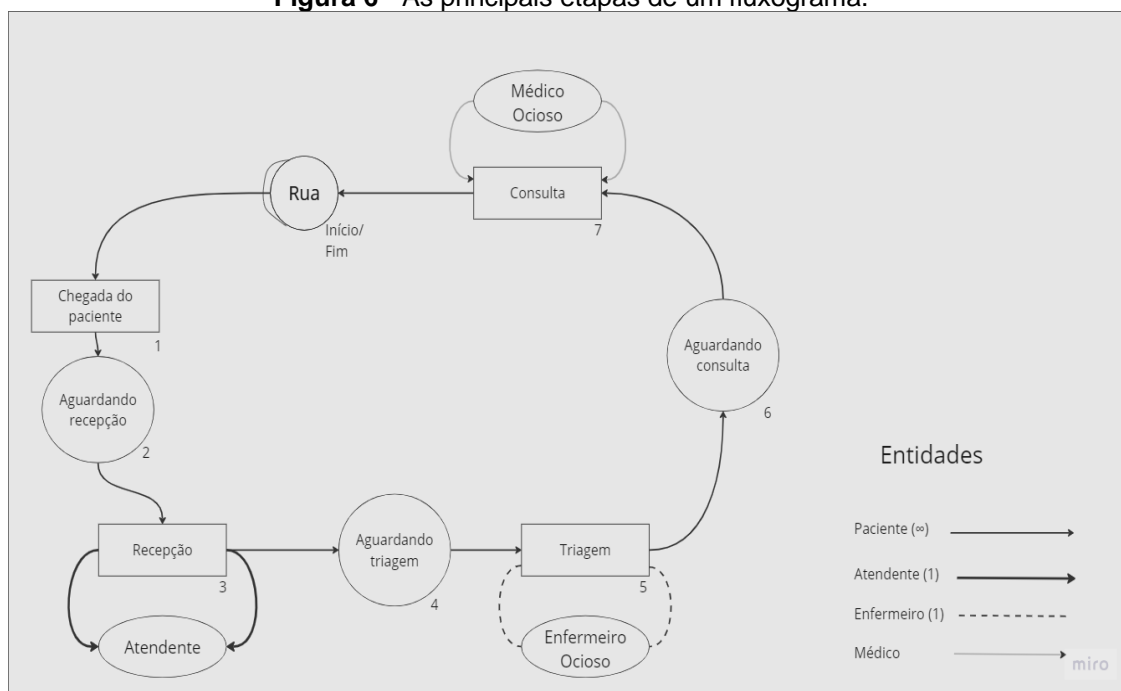
4.2 MAPEAMENTO DO PROCESSO

Para modelar a simulação foi feito um mapeamento dos processos dos principais setores de atendimento da UPA.

Este fluxograma representa como a Unidade de Pronto Atendimento (UPA) funciona atualmente. Os pacientes chegam, aguardam na fila para serem atendidos pela recepção. Depois que são atendidos, aguardam novamente e passam por uma triagem, na qual são avaliados por um profissional de saúde treinado.

Com base nos sintomas apresentados, o paciente é classificado em uma das categorias de cores denominadas pelo protocolo de *Manchester*, o que influenciará no tempo de espera. Após a triagem, o indivíduo espera pela consulta, e, por fim, sai do hospital após ser atendido pelo médico. Essas etapas são apresentadas na Figura 6.

Figura 6 - As principais etapas de um fluxograma.



Fonte: Elaborado pelo autor, 2024.

Em resumo, as principais etapas do fluxograma estão enumeradas na Figura 4 e podem ser descritas como:

Início - Ponto de partida dos pacientes

1 - Chegada de pacientes no hospital

2 - Após chegar, o paciente aguarda na fila para ser atendido pela recepção

3 - Nessa etapa ocorre o atendimento pelos recepcionistas

4 - Após passar pela recepção, o paciente aguarda a etapa de triagem na fila

5 - No processo de triagem, um profissional de saúde treinado efetua a avaliação da saúde do paciente e o classifica em uma das categorias de cores denominadas pelo protocolo de *Manchester*, que influenciará no tempo de espera

6 - Após ser classificado em uma das categorias de cores denominadas pelo protocolo de *Manchester*, o paciente aguarda a sua consulta na fila de espera

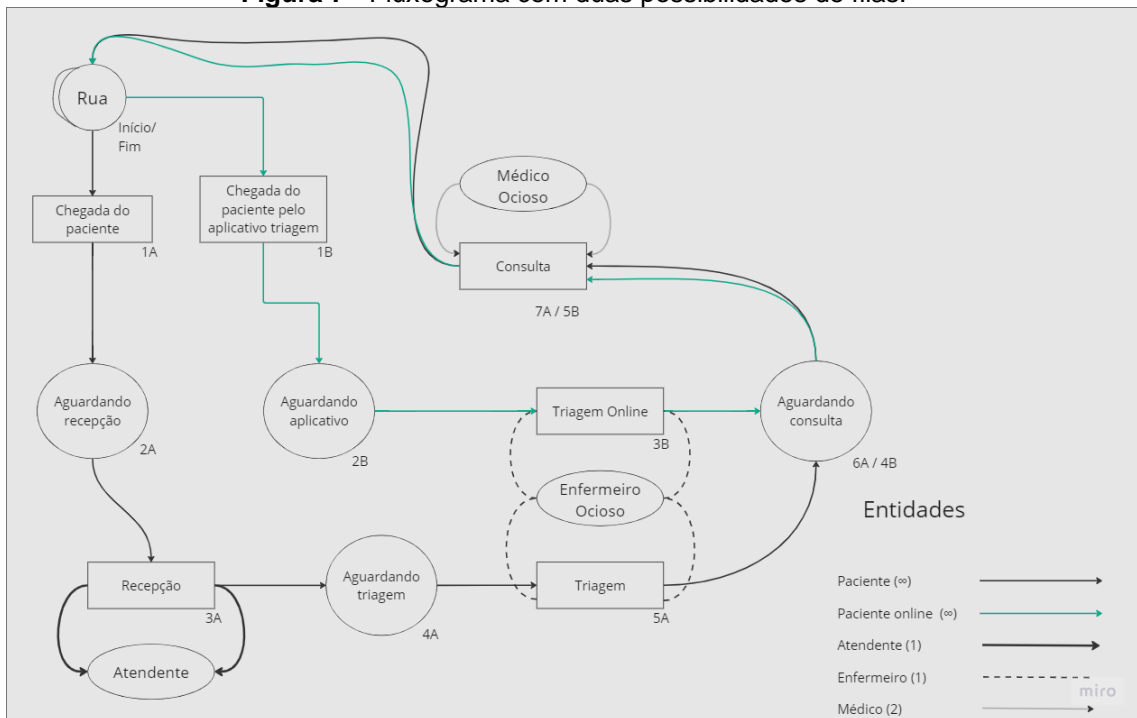
7 - Nessa etapa o paciente é atendido pelo médico

Fim - Ponto de saída dos pacientes

Já na Figura 7, o fluxograma representa como funcionaria a Unidade de Pronto Atendimento após a implementação de um aplicativo de triagem *online*, focado para os casos não graves. A proposta visa permitir que os pacientes utilizem o aplicativo de forma remota, eliminando a necessidade de espera física no hospital. Através dessa plataforma *online*, os pacientes podem acessar a fila de espera e receber notificações sobre o andamento de seu atendimento. Essa abordagem pretende melhorar a eficiência da UPA ao reduzir o tempo de espera percebido pelos pacientes e otimizar o uso dos recursos hospitalares

Dessa forma, o fluxograma se divide em duas possibilidades, sendo a “A” em que o processo ocorrerá como no fluxograma anterior, e “B” em que a triagem ocorrerá de forma *online*.

Figura 7 - Fluxograma com duas possibilidades de filas.



Fonte: Elaborado pelo autor, 2024.

Caminho A:

Início - Ponto de partida dos pacientes

1A - Chegada de pacientes no hospital

2A - Após chegar, o paciente aguarda na fila para ser atendido pela recepção

3A - Nessa etapa ocorre o atendimento pelos recepcionistas

4A - Após passar pela recepção, o paciente aguarda a etapa de triagem na fila

5A - No processo de triagem, um profissional de saúde treinado efetua a avaliação da saúde do paciente e a classificação em uma das categorias de cores denominadas pelo protocolo de *Manchester*, que influenciará no tempo de espera

6A - Após ser classificado em uma das categorias de cores denominadas pelo protocolo de *Manchester*, o paciente aguarda a sua consulta na fila de espera

7A - Nessa etapa o paciente é atendido pelo médico

Fim - Ponto de saída do paciente

Caminho B:

Início - Ponto de partida dos pacientes

1B - Chegada de pacientes pelo aplicativo de triagem *online*

2B - Nessa etapa, ocorre o tempo de processamento do aplicativo

3B - Após iniciar o aplicativo, o paciente preenche todos os dados necessários para fazer a triagem *online*, que fará uma “pré validação” conforme o Protocolo de *Manchester* e será avaliado no final pelo enfermeiro responsável

4B - Após a triagem *online*, o paciente recebe sua classificação do Protocolo de *Manchester* e poderá gerenciar seu tempo e aguardar a consulta enquanto acompanha a fila de espera de forma *online*

5B - Nessa etapa o paciente é atendido pelo médico

Fim - Ponto de saída do paciente

Este fluxograma representa como funcionaria a Unidade de Pronto Atendimento após a implementação de um aplicativo de triagem *online*, focado para os casos não graves. A proposta seria os pacientes utilizarem o aplicativo de qualquer lugar de sua preferência, dessa forma não estariam limitados a esperar seu atendimento no hospital e irão poder acessar a fila de forma *online* para saber quando está chegando a vez de ser atendido.

Já os pacientes que apresentam um grau maior de urgência, assim como no fluxograma anterior, iriam chegar e aguardar na fila (fora os casos de emergência, que necessitam de atendimento imediato) para serem atendidos pela recepção. Depois que são atendidos, aguardam novamente e passam por uma triagem, na qual são avaliados por um profissional de saúde treinado.

Com base nos sintomas apresentados, o paciente é classificado em uma das categorias de cores denominadas pelo protocolo, o que influenciará no tempo de espera. Como o aplicativo de triagem *online* seria focado para os casos não graves, os pacientes que apresentam um quadro de urgência maior serão atendidos mais rápido. Após a triagem, o indivíduo espera pela consulta, e, por fim, sai do hospital após ser atendido pelo médico.

De acordo com a literatura, ao seguir os protocolos de *Manchester*, o tempo médio de espera entre o registro na porta de entrada e a classificação de risco foi de 12 minutos, enquanto o tempo geral de triagem foi de 2 minutos. Somente 32,3% dos pacientes classificados como laranja foram atendidos em até 10 minutos após a classificação, no qual a maioria dos pacientes classificados como amarelo foi atendida pelo médico em até 60 minutos após a classificação. No geral, os pacientes esperaram, em média, 52 minutos entre a chegada ao hospital e o atendimento médico.

4.3 MODELOS SIMULADOS

Foram feitos 2 modelos de simulação, um simulando como a UPA Dom Orione funciona atualmente e o outro simulando a unidade de saúde com o sistema de triagem *online* para os casos não graves. As simulações foram feitas com o intuito de comparar e analisar o impacto de um sistema de triagem *online* na UPA Dom Orione.

No primeiro modelo, ao chegarem na UPA, os pacientes devem responder aos dados na recepção para poder consultar e, se for seu primeiro atendimento no local, devem cadastrar-se no sistema. Depois, devem passar por uma triagem em uma sala próxima para avaliar a prioridade de cada atendimento, conforme o protocolo de *Manchester*.

Após a triagem, os pacientes são organizados de acordo com a urgência de seu atendimento, utilizando porcentagens coletadas e atributos de valor associados à cada uma das cores que representam essa prioridade, conforme é ilustrado na Tabela 3.

Tabela 3 - Percentagens e atributos correspondentes a cada cor de prioridade

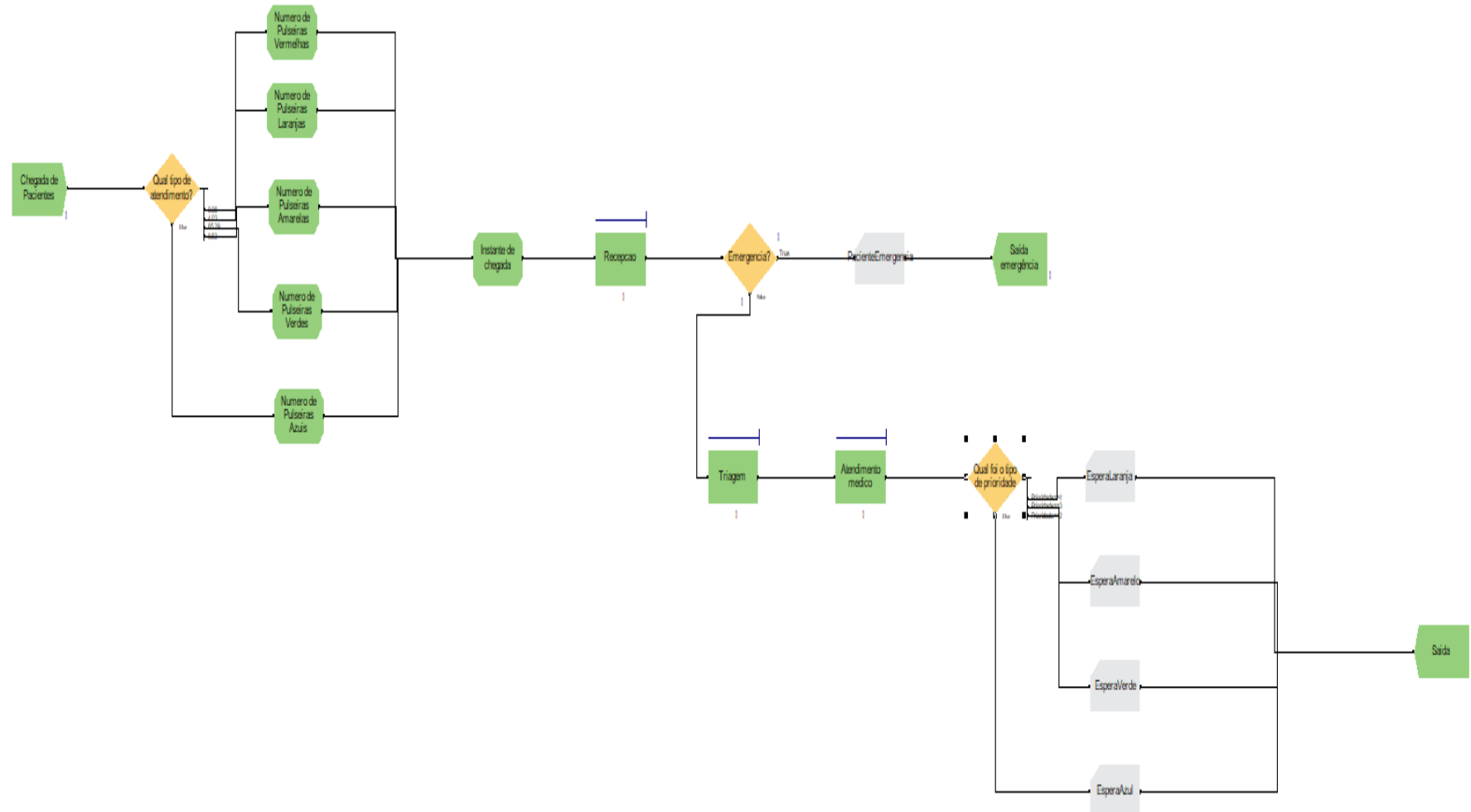
Cor	Porcentagem por cor	Atributo
Vermelho	0,08%	5
Laranja	4,03%	4
Amarelo	9,53%	3
Verde	85,29%	2
Azul	1,07%	1

Fonte: Elaborado pelo autor, 2024.

No modelo, a prioridade é mantida em todos os processos depois da triagem, pois foi criada uma disciplina de filas para atender primeiro os pacientes com o maior valor de atributo.

Em seguida, os pacientes se deslocam para um dos consultórios clínicos e, por fim, saem da UPA. O modelo permite a análise do tempo de espera separado para cada uma das cores de classificação.

Figura 8 - Modelo ARENA



Fonte: Elaborado pelo autor, 2024.

No segundo modelo, o sistema de triagem pode ser feito tanto de forma *online* quanto presencial. Para os casos de forma *online*, o paciente poderá analisar quando está chegando a sua vez de ser atendido sem precisar estar presente no local através de um aplicativo focado para os casos não graves (azuis e verdes). Já para os que optarem ou necessitarem ir presencialmente, o modelo seguirá conforme a simulação anterior.

Os pacientes são organizados de acordo com a urgência de seu atendimento, utilizando porcentagens coletadas e atributos de valor associados à cada uma das cores que representam essa prioridade, conforme foi ilustrado na Tabela 3.

No caso dos pacientes que seguem a triagem *online*, será feita uma validação em que os pacientes devem ter equipamentos necessários para fazer a triagem remota, como o termômetro e o medidor de pressão. Além disso, o aplicativo validará se a classificação realmente foi verde ou azul. Caso contrário o paciente deverá ir de forma presencial.

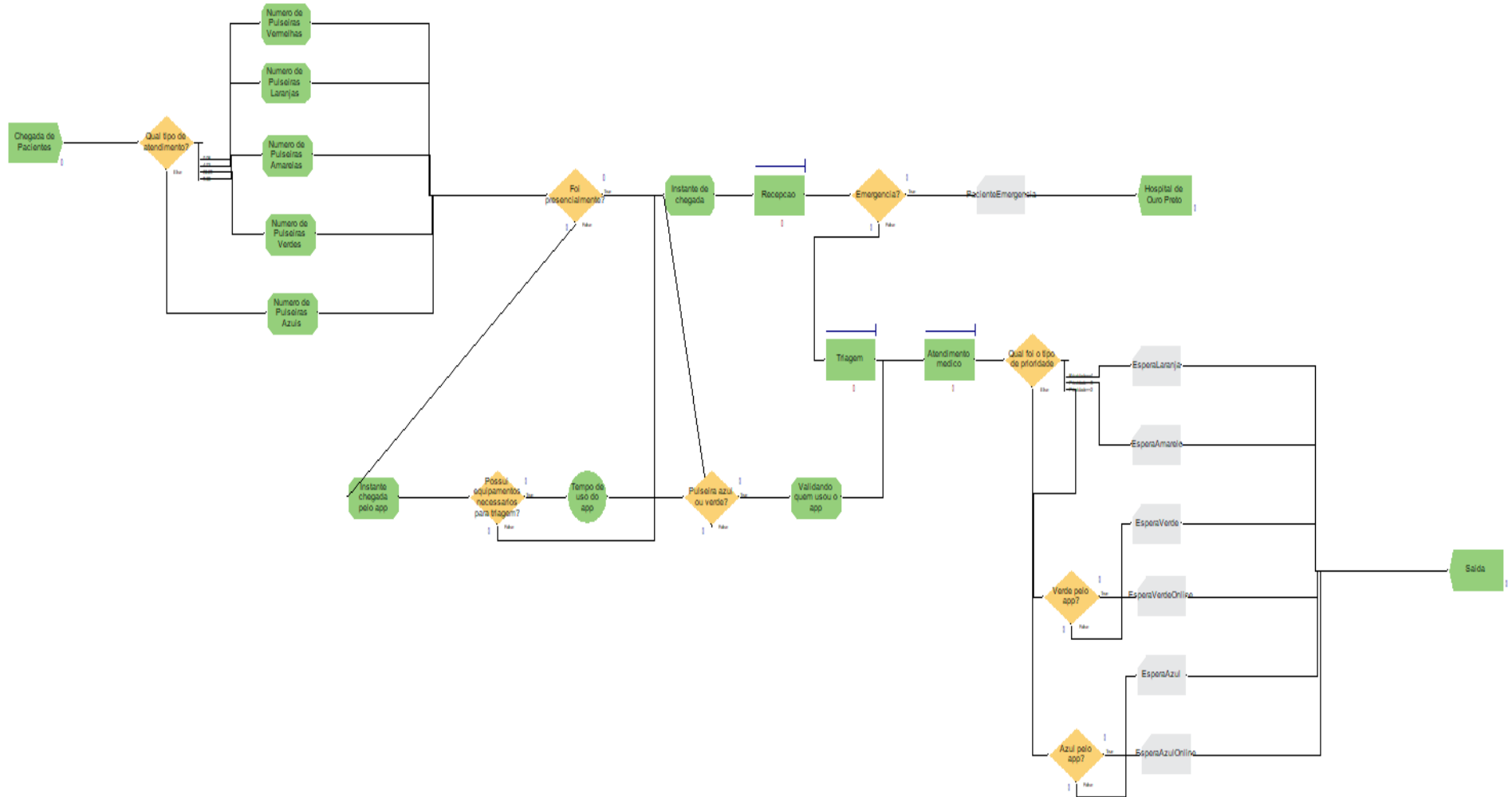
Após a triagem remota, os pacientes que apresentam temperatura dentro do esperado (36,0°C e 37,4°C) e com pressão próxima da normalidade, será realizado um questionário com tópicos para assinalar os seus sintomas, desenvolvendo o anamnese para auxiliar no diagnóstico do médico.

Após o paciente preencher os dados, o sistema integrará com o sistema do hospital para que o médico tenha acesso aos dados internos do paciente e agende a consulta conforme a percepção de importância de retorno e disponibilidade de atendimento presencial.

Durante o atendimento *online*, é informado que em casos de piora nos quadros de sintomas, o paciente deverá se locomover até o pronto atendimento mais próximo ou acionar uma ambulância caso não consiga se locomover.

Após a etapa de triagem *online*, resta ao paciente aguardar o atendimento clínico e, por fim, sair da UPA. Este modelo permite a análise do tempo de espera separado para cada uma das cores de classificação, além do tempo de espera para os casos *online*.

Figura 9 - Segundo Modelo ARENA



Fonte: Elaborado pelo autor, 2024

4.4 RESULTADOS

4.4.1 Análise da simulação da UPA atualmente

Por meio da simulação, o *software* utilizado fornece os dados de taxa de uso dos recursos e os tempos médios de espera para cada cor do protocolo de *Manchester*, assim como os tempos totais médios em que o paciente espera em fila ou se encontra recebendo os serviços. Para obter os dados, foram realizadas três replicações do modelo desenvolvido e foi simulado um período de 7 dias. A unidade de tempo base da simulação é em minutos.

A Tabela 4 apresenta os valores de taxa de utilização dos servidores, que foram gerados pela simulação:

Tabela 4 - Taxa de utilização de cada atendente.

Taxa de utilização		
Recepcionista	Enfermeiro	Médico
0,191	0,235	0,919

Fonte: Elaborado pelo autor, 2024.

De acordo com a Tabela 4, a ordem dos atendentes com a maior taxa de utilização é, respectivamente, médico, enfermeiro e recepcionista. O servidor mais demandado foi o atendimento médico e este apresentou uma ociosidade de 8,1%. Ou seja, em um período de plantão de 12h, os médicos têm aproximadamente 1 hora para realizar outras atividades além do atendimento médico.

O tempo médio de espera em fila para cada um dos serviços da UPA pode ser visto na Tabela 5.

Tabela 5 - Tempo médio de espera de pacientes para cada atendente.

Tempo médio de espera		
Recepcionista	Enfermeiro	Médico
0,065	0,073	49,414

Fonte: Elaborado pelo autor, 2024.

De acordo com a tabela acima, a ordem de maior tempo de espera para receber os serviços é, respectivamente, atendimento médico, triagem e recepcionista.

A Tabela 6 apresenta o número médio de pacientes esperando em cada uma das filas. Analisando o número médio de pessoas esperando nas filas, pode-se concluir que a ordem de maior número médio de pacientes esperando em fila para cada serviço é, respectivamente, atendimento médico, triagem e recepção. Esta informação reforça que o setor mais demandado é o atendimento médico.

Tabela 6 - Número médio de espera de pacientes para cada atendente.

Número médio de espera		
Recepcionista	Enfermeiro	Médico
0,01	0,073	0,01

Fonte: Elaborado pelo autor, 2024.

Como pode ser visto na Figura 10, a simulação fornece o tempo médio que o paciente espera em fila (49,559 minutos), o tempo médio dos atendimentos (20,325 minutos), e o tempo médio total que o paciente ficou dentro do sistema (69,884 minutos).

Figura 10 - Entity.

Entity						
Time						
	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
VA Time						
Pacientes	20.3254	0,62	20.1376	20.6105	0.4472	71.7463
NVA Time						
Pacientes	0.00	0,00	0.00	0.00	0.00	0.00
Wait Time						
Pacientes	49.5593	51,51	33.0108	72.8176	0.00	2920.92
Transfer Time						
Pacientes	0.00	0,00	0.00	0.00	0.00	0.00
Other Time						
Pacientes	0.00	0,00	0.00	0.00	0.00	0.00
Total Time						
Pacientes	69.8847	52,08	53.2390	93.4280	0.4472	2939.20

Fonte: Elaborado pelo autor, 2024.

Outro dado que foi calculado através do relatório do ARENA foram os tempos totais médios de espera em fila de atendimento, ou seja, o tempo total médio

em que cada cor, de acordo com a classificação de risco do paciente, pode permanecer no sistema. Assim, de acordo com a Tabela 7, considerando o tempo em que o paciente chega no hospital procurando um serviço no setor de pronto atendimento, ele deve ficar aproximadamente 26,261 minutos no sistema para os pacientes com a cor laranja, 25,309 minutos no sistema para a cor amarelo, 71,907 minutos no sistema para a cor verde e 485,49 minutos no sistema para a cor azul.

Tabela 7 - Tempo médio de espera de acordo com cada classificação.

Classificação de risco	Tempo média de espera
Laranja	26,261
Amarelo	25,309
Verde	71,907
Azul	485,49

Fonte: Elaborado pelo autor, 2024.

4.4.2 Análise da simulação da UPA com aplicativo de triagem

Por meio da simulação, foi possível projetar como seria a UPA com um aplicativo de triagem *online*, com foco para os casos não graves (pulseira azul e verde). Assim como no modelo anterior, o software utilizado fornece os dados de taxa de uso dos recursos e os tempos médios de espera para cada cor do protocolo de *Manchester*, assim como os tempos totais médios em que o paciente espera em fila ou se encontra recebendo os serviços.

Para obter os dados, foram realizadas três replicações do modelo desenvolvido e foi simulado um período de 7 dias. O tempo de uso estimado para o aplicativo proposto é de aproximadamente 2 minutos e 30 segundos e foi simulado um cenário em que 30% dos pacientes optaram por usar o aplicativo.

A Tabela 8 apresenta os valores de taxa de utilização dos servidores, que foram gerados pela simulação:

Tabela 8 - Taxa de utilização de cada atendente.

Taxa de utilização		
Recepcionista	Enfermeiro	Médico
0,085	0,206	0,848

Fonte: Elaborado pelo autor, 2024.

De acordo com a tabela acima, a ordem dos atendentes com a maior taxa de utilização é, respectivamente, médico, enfermeiro e recepcionista. O servidor mais demandado foi o atendimento médico e este apresentou uma ociosidade de 15,2%. Ou seja, em um período de plantão de 12h, os médicos têm aproximadamente 2 horas para realizar outras atividades além do atendimento médico.

O tempo médio de espera em fila para cada um dos serviços da UPA pode ser visto na Tabela 9.

Tabela 9 - Tempo médio de espera de pacientes para cada atendente.

Tempo médio de espera		
Recepcionista	Enfermeiro	Médico
0,001	0,087	23,15

Fonte: Elaborado pelo autor, 2024.

De acordo com a Tabela 9, a ordem de maior tempo de espera para receber os serviços é, respectivamente, atendimento médico, triagem e recepcionista.

A Tabela 10 apresenta o número médio de pacientes esperando em cada uma das filas. Analisando o número médio de pessoas esperando nas filas, pode-se concluir que a ordem de maior número médio de pacientes esperando em fila para cada serviço é, respectivamente, atendimento médico, triagem e recepção. Esta informação reforça que o setor mais demandado é o atendimento médico.

Tabela 10 - Número médio de espera de pacientes para cada atendente.

Número médio de espera		
Recepcionista	Enfermeiro	Médico
0,001	0,012	3,71

Fonte: Elaborado pelo autor, 2024.

Como pode ser visto na Figura 11, a simulação fornece o tempo médio que o paciente espera em fila (23,244 minutos), o tempo médio dos atendimentos (18,318 minutos), e o tempo médio total que o paciente ficou dentro do sistema (41,954 minutos).

Figura 11 - Entity: Parte 2

Entity						
Time						
VA Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Pacientes	18.3180	0,90	18.0675	18.7318	0.8896	178.07
NVA Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Pacientes	0.00	0,00	0.00	0.00	0.00	0.00
Wait Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Pacientes	23.2444	27,63	16.4575	36.0812	0.00	898.44
Transfer Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Pacientes	0.00	0,00	0.00	0.00	0.00	0.00
Other Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Pacientes	0.3829	0,03	0.3711	0.3963	0.00	2.5000
Total Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Pacientes	41.9452	27,11	35.5605	54.5451	0.8896	911.74

Fonte: Elaborado pelo autor, 2024.

Outro dado que foi calculado através do relatório do ARENA foram os tempos totais médios de espera em fila de atendimento, ou seja, o tempo total médio em que cada cor, de acordo com a classificação de risco do paciente, pode permanecer no sistema. Também foi considerado os tempos dos pacientes que passaram pelo sistema de triagem *online*.

Assim, de acordo com a Tabela 11, considerando o tempo em que o paciente chega no hospital procurando um serviço no setor de pronto atendimento, ele deve ficar aproximadamente 22,680 minutos no sistema para os pacientes com a cor laranja, 22,390 minutos no sistema para a cor amarelo, 43,789 minutos no sistema para a cor verde, 44,075 minutos no sistema para a cor verde pelo aplicativo de classificação de risco *online*, 122,852 minutos no sistema para a cor azul e 96,119 minutos no sistema para a cor azul pelo aplicativo.

Tabela 11 - Tempo médio de espera de acordo com cada classificação

Classificação de risco	Tempo média de espera
Laranja	22,680
Amarelo	22,390
Verde	43,789
Verde <i>Online</i>	44,075
Azul	122,852
Azul <i>online</i>	96,119

Fonte: Elaborado pelo autor, 2024.

4.5 COMPARAÇÃO DOS RESULTADOS OBTIDOS EM CADA MODELO

A partir dos resultados obtidos nas simulações, foi realizada a comparação entre o modelo representando o funcionamento atual da unidade e o modelo proposto com a utilização da triagem online para pacientes classificados como não urgentes. A análise considerou principalmente o tempo médio de espera na etapa de triagem e a taxa de utilização dos servidores, por serem indicadores diretamente relacionados à eficiência do atendimento e à organização do fluxo de pacientes.

Tabela 12 - Tempo médio de espera de acordo com a classificação de risco na UPA convencional e com triagem *online*.

Classificação de risco	Tempos médio de espera (minutos)	
	UPA atualmente	UPA com triagem <i>online</i>
Laranja	26,261	22,680
Amarelo	25,309	22,390
Verde	71,907	43,789
Azul	485,49	44,075
Verde <i>Online</i>	x	122,852
Azul <i>online</i>	x	96,119

Fonte: Elaborado pelo autor, 2024.

Tabela 13 – Taxa de utilização dos servidores na UPA convencional e com triagem *online*.

Servidores	Taxa de utilização dos servidores (%)	
	UPA atualmente	UPA com triagem <i>online</i>
Recepcionista	19,1	8,5
Amarelo	23,5	20,6
Verde	91,9	84,8

Fonte: Elaborado pelo autor, 2024.

Conforme observado na Tabela 12, no modelo atual a triagem presencial apresenta maior tempo médio de espera, principalmente para pacientes classificados nas cores verde e azul, que possuem menor prioridade no atendimento segundo o Protocolo de Manchester. No modelo proposto, parte desses pacientes realiza a triagem previamente por meio do sistema online, reduzindo a quantidade de pessoas aguardando presencialmente e, conseqüentemente, diminuindo o tempo médio de espera nessa etapa. Essa redistribuição da demanda contribui para um fluxo mais equilibrado dentro da unidade, reduzindo a formação de filas e melhorando a organização do atendimento.

Em relação à taxa de utilização dos servidores (Tabela 13), verificou-se que no cenário atual alguns recursos operam próximos ao limite de capacidade, o que aumenta a probabilidade de formação de filas e sobrecarga dos profissionais. No modelo com triagem online, a carga de trabalho é distribuída de forma mais uniforme, resultando em menores taxas de utilização em determinados servidores e maior estabilidade do sistema como um todo. Esse comportamento indica que a introdução de um mecanismo de organização prévia da demanda pode contribuir para melhorar o aproveitamento dos recursos disponíveis sem a necessidade de aumento da estrutura física ou do número de profissionais.

De forma geral, a comparação entre os cenários demonstra que a utilização de triagem online associada a um sistema de fila virtual apresenta potencial para reduzir o tempo de espera presencial, melhorar a utilização dos servidores e aumentar a eficiência operacional da unidade. Esses resultados reforçam a viabilidade da proposta estudada e evidenciam a importância do uso de ferramentas de simulação computacional para avaliar mudanças no sistema antes de sua implementação real.

5. CONSIDERAÇÕES FINAIS

O presente estudo teve como objetivo analisar a aplicação da teoria das filas e da simulação computacional no contexto da UPA Dom Orione, em Ouro Preto – MG, propondo a implementação de um sistema de triagem *online* para pacientes classificados como casos não graves, conforme o Protocolo de *Manchester*. A pesquisa partiu da problemática relacionada ao tempo de espera e à organização do fluxo de atendimento em unidades públicas de urgência e emergência.

A hipótese 1 estabelecia que a implementação de um sistema de triagem *online* poderia contribuir para a redução do tempo médio de espera presencial dos pacientes classificados nas cores verde e azul. A partir da comparação entre os cenários simulados, verificou-se redução no tempo médio de permanência física na unidade para esses pacientes, indicando que a redistribuição prévia da demanda tende a impactar positivamente o fluxo interno de atendimento. Dessa forma, os resultados obtidos corroboram a hipótese formulada.

No que se refere à hipótese 2, que propunha que a adoção de uma fila virtual poderia melhorar a organização do fluxo de pacientes, os dados evidenciaram maior equilíbrio no sistema quando parte da demanda foi direcionada para o atendimento *online*. A redução de gargalos e a melhor previsibilidade do atendimento reforçam que a organização prévia da chegada dos pacientes contribui para maior eficiência operacional.

A hipótese 3 sugeria que a introdução da triagem *online* poderia otimizar a utilização dos servidores do sistema. Os resultados apontaram redução na taxa de utilização de determinados servidores no cenário com triagem *online*, indicando melhor distribuição da carga de trabalho e menor sobrecarga operacional. Isto demonstra que a proposta pode contribuir não apenas para a melhoria do atendimento ao usuário, mas também para a gestão interna da unidade.

A hipótese 4 afirmava que simulação computacional seria capaz de representar adequadamente o sistema real e permitir a análise comparativa de cenários sem interferir na rotina da unidade. A modelagem desenvolvida no *software* Arena possibilitou a visualização do comportamento do sistema, identificação de gargalos e avaliação de diferentes configurações operacionais, confirmando a relevância da simulação como ferramenta de apoio à tomada de decisão.

Por fim, no que se refere à hipótese 5, verificou-se que a simulação computacional e a modelagem do sistema de filas demonstraram potencial como instrumentos de apoio à gestão, permitindo a análise estruturada do fluxo de atendimento e subsidiando decisões relacionadas à organização do sistema. Tal resultado evidencia que a aplicação dessas ferramentas pode contribuir para o aprimoramento da gestão pública em saúde, ao favorecer a utilização mais eficiente da capacidade instalada.

De maneira geral, os resultados indicam que a implementação de um sistema de triagem *online* para casos não graves apresenta viabilidade operacional e potencial para melhorar a eficiência do atendimento na UPA Dom Orione. Ainda que a proposta não elimine integralmente os desafios inerentes à variabilidade da demanda em serviços de urgência e emergência, os resultados indicam que a utilização de ferramentas de simulação pode representar um avanço consistente na organização e eficiência do atendimento.

O estudo reforça a aplicabilidade prática da teoria das filas e da simulação computacional no contexto da saúde pública, contribuindo para o desenvolvimento de soluções baseadas em dados e passíveis de replicação em outras unidades com características semelhantes. Como sugestões para pesquisas futuras, recomenda-se a ampliação da análise considerando diferentes níveis de adesão ao sistema *online*, avaliação econômica da implementação e investigação da percepção dos usuários quanto à experiência com a fila virtual.

REFERÊNCIAS

- ALMEIDA, G. F.; SOBRAL, L. S.; OLIVEIRA, J. A.; BALTAZAR, B. T.; SILVEIRA NETA, A.; TAVAREZ, J. J. DE NOVAIS, B. F.; RODRIGUES, R. M. G.; CAVALCANTI, J. M. M.; DUARTE, L. A.; MARÇAL, G. F.; MACHADO, R. S. S. Uso da simulação computacional para ensino e aprendizagem nos cursos de saúde. *Research, Society and Development*, v. 10, n. 15, p. 1-9, 2021.
- AMARAL, S. F. **O uso do protocolo de Manchester pode auxiliar no atendimento humanizado em uma emergência.** Brasília-DF, 2017.
- BANKS, J.; CARSON, J. S.; NELSON, B. L.; NICOL, D. M. *Discrete-Event System Simulation*. 5. ed. Upper Saddle River: Prentice Hall, 2010.
- BRASIL ESCOLA. **Acolhimento com classificação de risco, um olhar na perspectiva do assistente social.** 2024. Disponível em: <https://meuartigo.brasilecola.uol.com.br/saude/acolhimento-com-classificacao-risco-um-olhar-na-perspectiva.htm>. Acesso em: 10 ago. 2024.
- BRUNS, R. et al. **Pesquisa operacional: uma aplicação da teoria das filas a um sistema de atendimento.** Rio de Janeiro, 2001.
- BUNN, F.; BYRNE, G.; KENDALL, S. **The effects of telephone consultation and triage on healthcare use and patient satisfaction: a systematic review.** *British Journal of General Practice*, v. 55, n. 521, p. 956–961, 2005.
- CAMPOS, Gastão Wagner de Sousa; RATES, Susana Maria Moreira. Segredos e impasses na gestão de um hospital público. **Revista Médica de Minas Gerais**, 2024.
- FIGUEIREDO, D. D.; ROCHA, S. H. Aplicação da teoria das filas na otimização do número de caixas: um estudo de caso. **Revista Iniciação Científica**, 2010.
- FREITAS, L. F.; MOSER, A.; DOS SANTOS, N. A simulação computacional como técnica de pesquisa na administração. **Revista Intersaberes**, Curitiba, v. 9, n. Especial, p. 463–485, 2015.
- GALINDO, A. G.; JUNIOR, R. F. D. O. S. S. Estudo de Caso da Demanda por Atendimento no Hospital Municipal de Nova Odessa Visando Redução no Tempo de Espera na Fila. **Revista dos Alunos de Administração**, v. 1, n. 1, p. 80-92, 2012.
- GERÔNIMO, M. da S. et al. O uso da simulação computacional para melhoria nos processos produtivos: uma aplicação da teoria de filas com o uso de simuladores. **Exacta**, São Paulo, v. 16, n. 3, p. 167–180, 2018.
- GREEN, L. V. Queueing analysis in healthcare. In: HALL, R. W. (ed.). **Patient Flow: Reducing Delay in Healthcare Delivery**. New York: Springer, 2006. p. 281–307.
- GROSS, Donald; HARRIS, Carl M. **Fundamentals of queueing theory**. 3. ed. New York: Wiley, 1998.

GUIMARÃES, I. F. G. **Modelo de Rede de Filas para Avaliação de Desempenho em Trechos Singelos de Malhas Ferroviárias**. 2005. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Minas Gerais. Belo Horizonte, 2005.

HALL, R. W. (ed.). **Patient Flow: Reducing Delay in Healthcare Delivery**. 2. ed. New York: Springer, 2013.

HILLIER, Frederick S.; LIEBERMAN, Gerald J. **Introduction to operations research**. 10.ed. New York: McGraw-Hill, 2013.

JUN, J. B.; JACOBSON, S. H.; SWISHER, J. R. Application of discrete-event simulation in health care clinics: A survey. **Journal of the Operational Research Society**, v. 50, n. 2, p. 109–123, 1999.

HULSHOF, P. J. H.; KORTBEEK, N.; BOUCHERIE, R. J.; HANSS, E. W.; BAKKER, P. J. M. **Taxonomic classification of planning decisions in health care: a structured review of the state of the art in OR/MS**. *Health Systems*, v. 1, n. 2, p. 129–175, 2012.

KELTON, W. D.; SADOWSKI, R. P.; ZUPICK, N. B. **Simulation with Arena**. 6. ed. New York: McGraw-Hill Education, 2015.

LAW, A. M. **Simulation modeling and analysis**. 5. ed. New York: McGraw-Hill Education, 2015.

MACEDO, E. et al. Gestão da Produção e o Estudo de Filas em uma Indústria Moveleira. **Anais da Engenharia de Produção**, v. 1, n. 1, p. 86-97, 2017.

MANNING, Larissa; ISLAM, Md Shahidul. **A systematic review to identify the challenges to achieving effective patient flow in public hospitals**. *International Journal of Health Planning and Management*, 2023.

MARUJO, L. G.; CARVALHO, D.; LEITÃO, M. N. Otimização de layout utilizando-se o SLP combinado com teoria das filas: Um estudo de caso em uma oficina de rodas e freios de aeronaves. **Revista Gestão Industrial**, v. 6, n. 4, 2010.

O’CATHAIN, A.; MUNRO, J.; NICHOLL, J.; KNOWLES, E. **How helpful is NHS Direct? Postal survey of callers**. *BMJ*, v. 328, n. 7431, p. 560, 2004

OLIVEIRA, F. F. et al. Análise de Teoria das Filas: Sistema de Filas de um Serviço de Pronto Atendimento. **Anais da Engenharia de Produção**, v. 1, n. 1, p. 37-46, 2017.

ORLANDIN, B. C. et al. Aplicação de teoria das filas em um ambulatório do estado do Rio Grande do Sul. *In: CONGRESSO BRASILEIRO DE ENGENHARIA DE PRODUÇÃO*. 9. **Anais** [...]. Ponta Grossa, 2019.

PORDEUS, A. C. B. **Tecnologias no serviço público: a possibilidade de melhoria na gestão de filas da UFERSA Angicos**. 2021. Trabalho de Conclusão de Curso. Universidade Federal do Rio Grande do Norte, 2021.

PRAIA, C. R.; GOMES, C. F. S. **Simulação computacional aplicada à modelagem do processo de recebimento de uniformes na Marinha do Brasil**. Natal, 2013.

PREFEITURA MUNICIPAL DE OURO PRETO. **UPA Dom Orione amplia atendimento de urgência e emergência**. Ouro Preto, 2025. Disponível em: <https://www.ouropreto.mg.gov.br/noticia/4515>. Acesso em: 23 fev. 2026.

ROBINSON, S. ***Simulation: the practice of model development and use***. 2. ed. London: Palgrave Macmillan, 2014.

RONCALLI, A. A. et al. Protocolo de Manchester e população usuária na classificação de risco: visão do enfermeiro. **Revista Baiana de Enfermagem**, v. 31, n. 2, 2017.

SENNA, B. et al. Gestão da Fila de Espera para Cirurgias Eletivas em Hospitais do Sistema Único de Saúde. **Brazilian Journal of Surgery & Clinical Research**, v. 30, n. 2, 2020.

SILVA, E. P. S. **Aplicação da teoria das filas em unidade básica de saúde no atendimento odontológico no município de Muricilândia (TO)**. Artigo (Tecnólogo em Logística) – Universidade Federal de Tocantins, Araguaína, 2020.

SOUSA, A. J. A. **Teletriagem: há benefícios para os sistemas de saúde e seus usuários?** 2017. Dissertação (Mestrado em Administração Pública e de Empresas) – Fundação Getúlio Vargas, Rio de Janeiro, 2017.

TEIXEIRA, V. A.; OSELAME, G. B.; NEVES, E. B. O Protocolo de Manchester no sistema único de saúde e a atuação do enfermeiro. **Revista da Universidade Vale do Rio Verde**, v. 12, n. 2, p. 905-920, 2014.

TORRES, O. F. Elementos da teoria das filas. **Revista de Administração de Empresas**, v. 6, p. 111-127, 1966.