

## **FILAS NO SISTEMA DE SAÚDE: GESTÃO DA CAPACIDADE DE ATENDIMENTO EM UMA UNIDADE DE PRONTO ATENDIMENTO**

**Patrícia de Aquino Lannes Brites**

Universidade Federal do Espírito Santo  
Av. Fernando Ferrari, 514, Goiabeiras, Vitória - ES  
patricia\_alannes@hotmail.com

**Fabício Broseghini Barcelos**

Instituto Federal do Espírito Santo  
Av. Governador José Sette, s/n, Itacibá, Cariacica - ES  
Fabricio.barcelos@ifes.edu.br

**Fernanda Amorim Ludgero**

Universidade Federal do Espírito Santo  
Av. Fernando Ferrari, 514, Goiabeiras, Vitória - ES  
fernanda\_ludgero@hotmail.com

### **RESUMO**

Organizações da área de saúde possuem operações complexas e sua qualidade e *lead time* estão ligados à vida de seus clientes. A escassez de recursos, altos custos e aumento da demanda são pontos relevantes para a gestão desse tipo de instituição. Assim, cresce a necessidade de ferramentas que auxiliem a gestão da capacidade de atendimento. Este estudo analisou um Pronto Atendimento utilizando a Teoria das Restrições e a Simulação Computacional, para encontrar operações limitantes no sistema e, por meio dos resultados, sugerir e simular alternativas. Foram propostos 3 cenários. Os resultados do primeiro cenário indicam que o gargalo mudou da fila de triagem para a de exame de sangue. O segundo cenário apresentou uma redução de 94% na espera da fila de exame de sangue, mas necessita de alto investimento. O terceiro cenário, como alternativa ao segundo, gerou redução de 45% do tempo, porém sem a necessidade de realizar investimentos.

**PALAVRAS-CHAVE: Simulação, Teoria das Restrições, Saúde.**

**TÓPICO: SIM**

### **ABSTRACT**

Healthcare organizations have complex operations and their quality and lead time are linked to the life of its customers. The scarcity of resources, high costs and increasing demand are difficult points to the management of this type of institution. Thus, grows the need for tools that help manage the service capacity. This study examined an Emergency Unit using the Theory of Constraints and Computer Simulation, in order to find the limiting operations in the system and, through the results, suggest and simulate alternatives. It is likely, that further restrictions arise and it is needed to implement a process of continuous improvement. Were proposed three scenarios. The first transferred the bottleneck from the medical screening queue to the blood test. The second scenario showed a 94% reduction in waiting time for blood test, but it has the need for high investment. The third scenario, as an alternative to the second, presented a reduction of 45% of the time on the same queue, but without the need for investments.

**KEYWORDS: Simulation, Theory of Constraints, Health.**

## 1. Introdução

Organizações da área de saúde podem ser consideradas verdadeiros desafios gerenciais. Algumas dificuldades que os gestores enfrentam são a simultaneidade entre produção e consumo, a escassez de recursos, os altos custos e o aumento crescente na demanda.

Além disso, pesquisas apontam que os processos nesses sistemas são dinâmicos, as operações são de elevada complexidade e sua qualidade e *lead time* estão ligados não somente à satisfação do cliente, mas à sua saúde e dignidade. Falhas de gestão podem custar vidas.

No setor de atendimentos de urgência e emergência, o fator tempo de resposta é crítico. Pacientes graves necessitam de atendimento rápido, pois, quanto maior seu tempo de espera, mais prolongado é seu sofrimento e, menores suas chances de restabelecimento. No entanto, apesar da meta no setor estar direcionada à velocidade no atendimento, é observado receio e revolta de usuários devido às longas esperas.

Muitos pesquisadores chamam a atenção para a necessidade de ferramentas que auxiliem a gestão da capacidade de atendimento em hospitais, levando em consideração o fluxo de pacientes e os recursos disponíveis. É fundamental, constantemente, repensar as práticas profissionais ou intervir sobre a forma de organização dos serviços, visando seu aperfeiçoamento. [Ramos e Lima, 2003]

Dessa maneira, esse estudo se justifica pela necessidade de aplicação de ferramentas na busca das causas de longas esperas no atendimento de pacientes, e da sugestão de alternativas que aumentem a eficiência do serviço de modo a não prejudicar a população, não sobrecarregar os servidores e ser viável financeiramente à instituição.

O objetivo do presente trabalho é a análise do fluxo de pacientes em uma Unidade de Pronto Atendimento utilizando a Teoria das Restrições e a Simulação Computacional.

Em termos específicos visa-se mapear o sistema de atendimento, compreender cada processo que o mesmo envolve, colher dados para construção do modelo computacional, analisar os resultados da simulação e, aplicar os cinco passos para melhoria contínua contidos no conceito da Teoria das Restrições de modo a gerar alternativas que extingam ou atenuem o impacto de fatores limitantes sobre o escoamento da demanda.

## 2. Materiais e Métodos

O *software* escolhido para a execução da simulação computacional foi o Arena®. O método de construção do modelo simulado segue simplificada, os seguintes passos de concepção: coleta de dados *in loco*, formulação do modelo conceitual e geração do modelo computacional. A literatura possui vários exemplos de modelos computacionais construídos para simular ambientes hospitalares [Best, 2014; Pegoraro, 2012; Sabbadini, 2005], no entanto uma característica marcante desse tipo de organização está na sua variabilidade de etapas, processos, horários etc. Assim, geralmente, cada autor constrói o modelo adequado à realidade do ambiente a ser estudado.

Para execução da coleta de dados foi necessária permissão da autoridade relacionada à instituição pesquisada, porém passos burocráticos, posteriores a autorização, não foram concluídos dentro do cronograma desta pesquisa. Por este motivo, foram realizadas entrevistas junto aos funcionários para obter os dados.

Devido a este método de coleta e à falta de acompanhamento dos dados pela instituição, foi somente possível determinar o valor mínimo, a moda e o valor máximo para cada variável. Portanto, usou-se a distribuição triangular para determinar o comportamento estatístico das variáveis, como sugerido por vários autores.

Os dados foram inseridos ao modelo e, passou-se para a etapa de verificação e validação. No primeiro foi realizada a comparação do modelo computacional com o conceitual, para analisar se um representa o outro. E na validação foi estabelecido o intervalo de valores que contem a média da população, ou a confiança estatística do modelo.

Para estabelecer os parâmetros de validação o modelo foi executado com 30 replicações e a cada replicação a média das variáveis de interesse foi exportada para uma planilha eletrônica na qual foram realizados os cálculos do desvio padrão e do semi-intervalo de confiança  $h$  (*half width*). Esse último, ao ser somado e subtraído da média, fornece o intervalo de valores que contém o verdadeiro valor do parâmetro. [Neves 2009]

O cálculo de  $h$  foi realizado considerando um nível de confiança padrão de 95%, pela equação 1.

$$h = t_{(n-1, \alpha/2)} \cdot (s/\sqrt{n}) \quad (1)$$

Onde:

$t_{(n-1, \alpha/2)}$  = Valor tabelado da distribuição t-student para distribuição com (n-1) graus de liberdade e nível de confiança (1- $\alpha$ );

S = Desvio Padrão Amostral;

n = Quantidade de replicações.

O  $h$  desejado, ou  $h^*$ , deve ser menor ou igual a 10% da média [Chwif e Medina 2010]. Assim, se  $h > h^*$ , deve-se calcular o número de replicações ideal, por meio da equação 2, para que os dados de saída do modelo sejam confiáveis.

$$n^* = [ n \times ( h \div h^* )^2 ] \quad (2)$$

Onde:

$n^*$  = número de replicações ideal;

n = número de replicações inicial;

$h^*$  = semi-intervalo desejado;

h = semi-intervalo inicial.

Nesse caso o número de replicações ideal encontrado foi de 109. Para a validação do modelo primeiramente foi realizado o teste de animação que representou adequadamente o funcionamento e fluxos do sistema. A seguir, foi analisada a sensibilidade do modelo através de pequenas mudanças em alguns dados para se observar se o sistema se comportaria como esperado. E por fim, como utilizado na literatura [Sabbadini et al., 2006; Pegoraro, 2012] o sistema foi submetido ao teste de Turing, onde os dados foram apresentados a um profissional experiente do Pronto Atendimento que os aprovou como representativos da realidade. Assim, seguiu-se para a análise crítica dos processos e resultados por meio da Teoria das Restrições para encontrar fatores limitantes do sistema, descobrir suas causas de ser e criar medidas que melhorem a eficiência do processo.

### 3. Revisão de Literatura

#### 3.1 Sistema Único de Saúde

O Sistema Único de Saúde é o instrumento pelo qual o governo deve prestar saúde pública e gratuita a todo cidadão e deve ser administrado de acordo com as necessidades dos usuários locais. [Pacheco e Cardoso, 2005]

No entanto, apesar de suas premissas, o SUS não possui credibilidade junto à sociedade, a crença generalizada é a de que o setor é ineficiente. Segundo Lima 1999, uma queixa frequente está ligada aos altos tempos de espera para atendimento. Uma pesquisa realizada pelo Instituto Brasileiro de Opinião Pública e Estatística (IBOPE) em 2002 mostrou que 36% dos usuários do SUS se queixam de filas, 17% reclamam que emergências não são atendidas, 9% afirmam que os médicos são mal preparados e 9% se queixam de demora.

Rocha et al (2005) relatam que, apesar de apresentar certa produtividade na análise de eficiência, os serviços médicos não conseguem absorver a demanda, criando longas filas de espera. Essas filas podem ter consequências dramáticas, pois, os autores afirmam que na maioria

das doenças o sucesso do tratamento está vinculado à agilidade do atendimento.

Além disso, o acesso a informações que possibilitem estudos na área é restrito e de permissão lenta devido à burocracia envolvida.

### 3.2 Simulação

Segundo Banks (1998), simulação é uma técnica de solução de problemas pela análise do modelo que descreve o comportamento de um sistema. No presente estudo, o sistema simulado foi o atendimento de urgência e emergência de um Pronto Atendimento de Vitória.

A construção do modelo computacional se deu através do *software* Arena®. O programa foi escolhido por sua interface gráfica amigável ao usuário, não sendo necessário conhecimentos em programação, por sua variedade de recursos de análise de dados e sua acessibilidade.

O modelo apresentado é classificado como matemático, pois utiliza relações matemáticas para determinar o comportamento do sistema; dinâmico, pois os resultados variam com a passagem de tempo e; discreto, pois seu estado varia discretamente no tempo.

Gonçalves (2004) menciona algumas características dos modelos de simulação que tornam essa tecnologia um instrumento adequado na solução de problemas da área de saúde como: baixo custo e pouco risco em vista da aplicação real; suportam flutuações estatísticas; oferecem resultados, conforme os critérios almejados.

Na área de saúde, os recursos são escassos, sendo assim Yang, Koziel e Leifsson (2013) citam que a maioria dos modelos envolve incerteza nas propriedades do material e parâmetros. Portanto, frequentemente, é necessário contentar-se com opções sub-ótimas, pois afinal, pretende-se que os modelos ajam de forma viável e com suficiente precisão. Dessa maneira, a simulação torna-se uma ferramenta eficaz no diagnóstico de processos e uma aliada ao estudo de novas alternativas. Autores como, Pegoraro (2012) e Sabbadini, et. al (2006), aliam à essa ferramenta o uso da Teoria das Restrições (TOC) e aprovam sua eficácia, na constatação e gestão de gargalos.

### 3.3 Teoria das Restrições (TOC)

A teoria das restrições ou TOC (*Theory of Constraints*) foi desenvolvida e divulgada pelo físico israelense Eliyahu M. Goldratt. Goldratt (2002) determina quatro conceitos que fundamentam a TOC, são eles: Toda organização é um sistema tangível, e seu desempenho depende da relação de seus processos entre si; toda organização possui uma meta e seus recursos devem estar direcionados a alcançá-la; todo sistema tangível possui, pelo menos, uma limitação; a gestão deve direcionar seus esforços aos recursos que mais impactarão sobre a meta.

A TOC demonstra que melhorias aplicadas ao fator limitante resultam numa melhora total do sistema. Se um hospital ganhar uma hora em um recurso gargalo, por exemplo, em uma triagem, ele estará ganhando uma hora em todo o sistema de atendimento do hospital [Pegoraro, 2012]. Com o objetivo de estabelecer mudanças nos lugares certos, Goldratt (2002) estabeleceu cinco passos para o gerenciamento das restrições: identificar a restrição do sistema (ou gargalo); explorar a restrição com ações que aumentem sua eficiência; subordinar os outros processos à restrição; elevar a eficiência do gargalo ao nível mais alto; e se uma restrição for elevada, voltar ao 1º passo, ou seja, melhoria contínua.

A aplicação da TOC se faz possível em diversas organizações, nas quais Sabbadini et. al (2006) incluem as da área de saúde, devido à amplitude de seus conceitos.

Os autores Rocha et al (2005), explicam que o fluxo de pacientes são entidades em um sistema composto por uma sequência de atividades que configuram uma cadeia de processos, com diferentes taxas médias de atendimento. Eles ressaltam que, a cadeia é tão eficiente quanto o processo com a menor taxa média de pacientes atendidos.

Sabbadini et. al (2006) consideram fundamental que a capacidade dos serviços de saúde corresponda à demanda. Isso se faz possível pela identificação dos recursos limitantes, onde a alteração da capacidade vai influenciar a capacidade total do sistema. Porém, segundo Rocha et al

(2005), a elevação da capacidade de um determinado setor considerado gargalo não pode ser obtida em detrimento da qualidade dos serviços prestados. Por isso o setor é um desafio à TOC.

#### **4. O Estudo de Caso**

Este estudo analisa uma Unidade de Pronto Atendimento da Grande Vitória, a instituição atende aproximadamente 1600 adultos e crianças semanalmente. Os pacientes são originários, principalmente, dos municípios de Vitória, Vila Velha, Cariacica e Serra.

A proposta da organização é, como o nome diz, o pronto atendimento de casos de urgência e emergência sem necessidade de prévio agendamento. Sua meta é atender os indivíduos em um intervalo de tempo em que seu estado não seja agravado, nem seu sofrimento prolongado.

No entanto a Unidade enfrenta diárias manifestações de repúdio de alguns de seus pacientes devido ao longo tempo de espera a que são submetidos quando buscam o atendimento. Apesar de possuir um sistema informatizado que acelera o processo de cadastro e a análise de exames, essa medida ainda não é suficiente para extinguir o problema das filas.

O período de funcionamento do local é de 24 horas por dia, divididos em dois turnos de 12 horas. O diurno se dá entre 7h e 19h e o noturno de 19h às 7h. Os turnos influenciam a capacidade da organização, que é baseada em escala.

O turno diurno conta com 4 clínicos gerais, 2 pediatras, 2 auxiliares de laboratório, 1 técnico radiologista, 3 enfermeiros, 8 técnicos de enfermagem e 2 recepcionistas. O noturno, no entanto, conta com apenas 2 clínicos gerais, 1 pediatra, 1 enfermeiro e 5 técnicos de enfermagem. A capacidade dos demais recursos é fixa. Essa diferença visa diminuir gastos com pessoal, visto que, à noite o fluxo de pacientes é muito menor, chegando a ser quase inexistente entre o período de meia-noite e 6h da manhã.

O processo de atendimento consiste nas seguintes etapas. O paciente chega e realiza o cadastro junto à recepção, em seguida é chamado para triagem, onde um profissional da área de enfermagem faz uma breve avaliação de seu quadro clínico avaliando seus sinais vitais, glicemia, sintomas, etc.

Após essa avaliação, o funcionário o classifica em um dos níveis de gravidade de acordo com o Protocolo de Manchester. Essa técnica trata-se de outra medida adotada pela empresa na tentativa de tornar o processo de atendimento mais eficiente. A metodologia do Protocolo de Manchester consiste na classificação dos pacientes pelo nível de gravidade e no estabelecimento de metas de tempo de espera pelo atendimento através de cores.

Pacientes classificados com a cor vermelha (emergentes) recebem atendimento imediato. Os classificados com a cor laranja (muito urgentes) devem esperar no máximo 10 minutos para serem atendidos. Quem recebe a cor amarela (urgentes) pode esperar até 60 minutos. Para os definidos com a cor verde (pouco urgentes) a espera não deve passar de 120 minutos. E por fim, a cor azul (não urgentes) determina uma espera máxima de 240 minutos.

Apesar de o sistema conter 5 cores para designar os graus de risco, a Unidade em questão adota apenas as 4 primeiras e os pacientes de classificação azul são incluídos ao grupo verde. Outrossim, pacientes emergenciais, pulam o processo de cadastro e triagem, pois sua estabilização deve ser feita imediatamente. Sua identificação de gravidade é, geralmente, feita na ambulância durante o trajeto, ou por observação de suas condições no momento em que adentra o local.

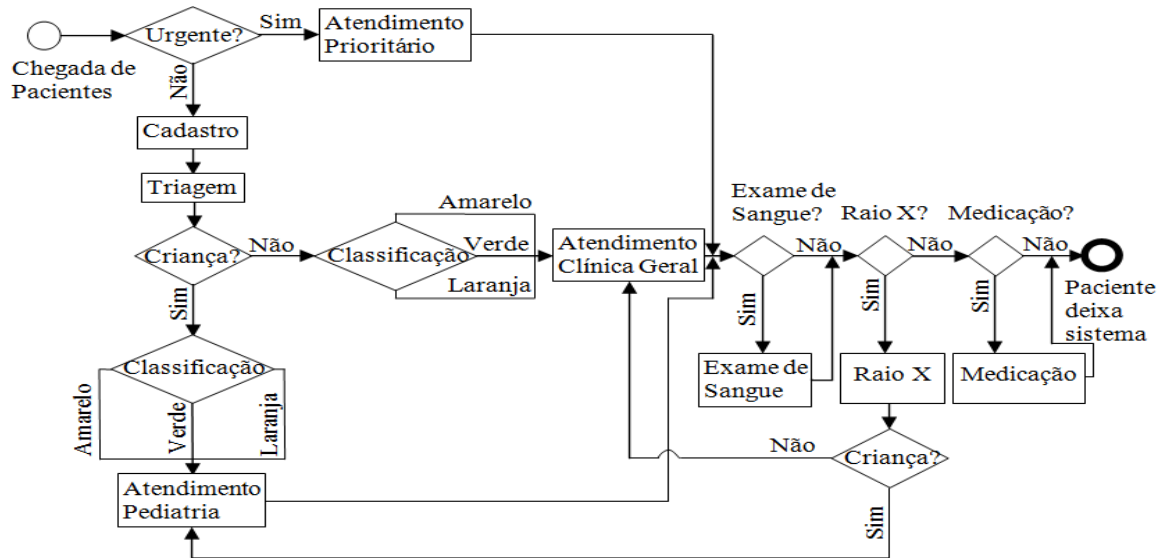
O método garante um atendimento mais rápido para pessoas com problemas mais graves, ou seja, a ordem da fila se dá por prioridade. Ainda assim, com exceção de casos de prioridade máxima, os alvos de espera máxima não são atingidos.

#### **5. Construção do modelo**

Foram realizadas visitas à Unidade estudada e observadas e mapeadas as atividades desenvolvidas, as etapas do processo e os tempos de cada atendimento. Foram, também, realizadas entrevistas para a coleta dessas informações.

Após reconhecimento do local, as etapas do processo foram organizadas em fluxograma, como na Figura 1, de maneira a facilitar sua visualização e a geração do modelo.

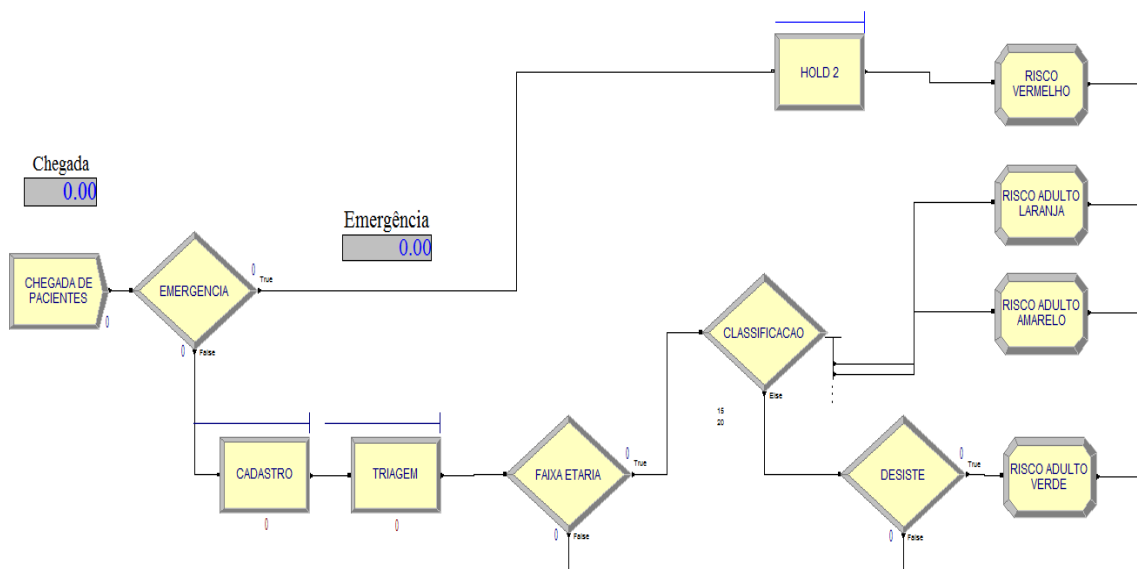
Figura 1: Modelo Conceitual.



O modelo construído considera os seguintes processos: chegada do paciente, cadastro na recepção, triagem, atendimento médico, exame e/ou medicação (caso necessário), retorno à consulta (caso necessário) e saída do sistema.

A Figura 2 apresenta a parte do modelo computacional referente à chegada de pacientes adultos, seu cadastro e triagem, onde recebem o atributo correspondente à classificação de risco.

Figura 2: Modelo Computacional - Chegada.



Devido à sua sazonalidade, a taxa de chegada de pacientes foi determinada por intervalos de tempo e dia da semana. Por exemplo, entre 6h e 8h da manhã de um domingo chegam, em média, 6 pacientes por hora, enquanto no mesmo horário de uma segunda-feira a taxa média é de 20 pacientes por hora.

O sistema, apesar de funcionar 24 horas por dia, não apresenta filas durante todo o período. As filas começam a se formar e se encerram entre 6 horas da manhã e meia-noite. Após esse horário, as chegadas são pontuais e não significativas para se analisar o tempo de espera nas filas.

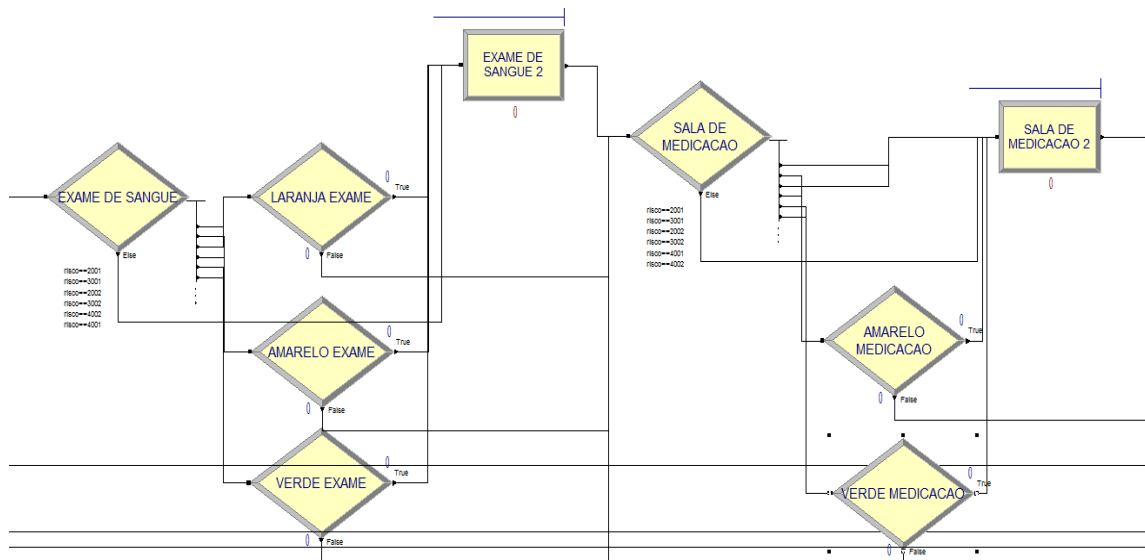
Nos horários de 7h e 19h há trocas de turno e mudança na capacidade de atendimento das consultas e da sala de medicação. Essa mudança foi inserida baseando a capacidade de atendimento dos médicos e dos funcionários da medicação na sua escala. Foram adicionadas também paradas programadas, através da função *failure*, que simbolizam o horário de almoço e pausas eventuais para descanso ou ir ao banheiro.

O regime da fila de consulta é por prioridade, o que significa que a ordem de atendimento se dá por um atributo dos indivíduos, neste caso, o paciente com classificação mais grave é atendido primeiro. Além disso, segue-se a regra de que o médico precisa esperar o atendimento ser encerrado antes de começar um novo ou fazer uma pausa.

A classificação de gravidade é feita durante a triagem, onde os pacientes recebem o atributo “risco”, excluído os casos mais graves que recebem o atributo logo após sua chegada, pois são atendidos diretamente. Além disso, são divididos pela faixa etária que ocupam, pois, os menores de 18 anos devem ser atendidos por um pediatra, enquanto os maiores por um clínico geral.

A Figura 3 mostra a parte do modelo que representa as filas para exame de sangue e medicação.

**Figura 3:** Modelo Computacional – Exame de sangue e medicação.



A quantidade de pessoas em cada faixa etária, em cada nível de risco, que realizam exame de sangue, raio x e medicação e a quantidade de desistentes é determinada por porcentagens. Todos os pacientes recebem o atributo correspondente à sua cor de classificação de risco, à sua faixa etária, ao tempo de duração de consulta e ao número de funcionários que o atenderão na consulta.

Esse último atributo se justifica pelo fato de os pacientes emergenciais requererem a atenção de todos os médicos disponíveis até sua estabilização, desse modo, o número de recursos utilizados se dá pela seguinte equação: número de (médico) escalados – número de (médico) ocupados. Porém, pode acontecer de o resultado da equação ser menor ou igual a 0. Dessa maneira, se adicionou o bloco *hold*, que "segura" o paciente até que o resultado da equação seja maior que 0.

Apesar de o paciente emergencial passar por essa espera virtual, ela não acontece na

realidade, descaracterizando a espera de pacientes "vermelhos" como uma variável de interesse do estudo. Porém, o tempo no qual esses pacientes estão sendo atendidos deve ser considerado para o cálculo da espera dos demais pacientes, por isso foram incluídos ao modelo.

Quando não há médicos disponíveis, o indivíduo de alto grau de risco utiliza, inicialmente, enfermeiros e técnicos de enfermagem em seu atendimento, até que haja disponibilidade de médicos, ou até que eles sejam chamados, interrompendo a consulta em andamento e a retomando após estabilização do paciente emergencial.

Foi calculada a necessidade de 109 replicações. O modelo foi simulado e, dentre as técnicas de validação de modelos apresentadas pelos autores, a considerada mais adequada ao trabalho em questão foi o teste de Turing. Esse teste consiste em apresentar os resultados do modelo a funcionários conhecedores do sistema e, os mesmos identificam se os valores são reais ou virtuais. Se essa identificação não puder ser feita, o modelo é validado.

## 6. Resultados

### 5.1 Cenário Atual

Na Tabela 1, estão os resultados da simulação com relação à espera nas filas de atendimento, e, a tabela 2, mostra os dados de espera de pacientes, divididos por grupo de risco. A primeira variável foi considerada importante para encontrar o gargalo do sistema, onde a melhoria deve ser implementada, e a segunda para observar se as metas do Protocolo de Manchester (tempo de espera/gravidade/paciente) estão sendo atendidas. Portanto, essas são as incógnitas que serão avaliadas neste estudo.

**Tabela 1:** Filas no cenário atual

Fila	Tempo de espera (min.)
Cadastro	14,98
Triagem	27,24
Consulta Adulto	64,08
Consulta Infantil	20,91
Exame de Sangue	16,61
Raio X	9,47
Medicação	4,04

**Tabela 2:** Filas por grupo de risco no cenário atual.

Paciente	Tempo de Espera (min.)	Meta (min.)	Espera total (triagem + consulta) (min.)
Vermelho	0	0	0
Laranja Adulto	5,65	10	32,89
Amarelo Adulto	16,82	60	44,06
Verde Adulto	97,46	120	124,70
Laranja Infantil	5,77	10	33,01
Amarelo Infantil	10,23	60	37,47
Verde Infantil	29,62	120	56,86

Analisando a Tabela 2, a princípio seus valores parecem corresponder às metas estabelecidas no protocolo de Manchester, porém, quando se observa a média de espera na fila de triagem na Tabela 1 e, a considera na espera total, nota-se que o tempo para os pacientes adultos da classe laranja e verde e, para os pacientes infantis da classe laranja, ultrapassam os valores desejados.

A organização justificou que, muitas vezes, os pacientes se prolongam na descrição de sintomas ou mesmo conversando sobre assuntos aleatórios com os médicos. Em entrevista,



funcionários da administração informaram que há um treinamento para os responsáveis pela triagem que diminui seu tempo máximo de 10 minutos para 5 minutos, ou até menos. Além disso, o tempo de consulta é encurtado, em média, de 20%, pois as informações na ficha do paciente são mais completas.

Considera-se então, a opção de treinamento de enfermeiros e técnicos como exploração da restrição e segue-se para a análise desse novo cenário, com tempos menores.

## 5.2 Cenário proposto 01 - Treinamento de funcionários de triagem

Os tempos de atendimento inicialmente coletados foram alterados correspondentemente à realidade alternativa apresentada. Os valores para atendimento foram decrescidos de 20% como justificado - exceto para casos emergenciais - e o tempo máximo de triagem passou a ser de 5 minutos.

Não houve alteração na distribuição de atendimento de pacientes emergenciais porque, uma vez que não passam por triagem não há como reduzir seu tempo de atendimento baseado nessa alternativa. As distribuições foram inseridas no modelo e a simulação demonstrou resultados como nas Tabelas 3 e 4.

**Tabela 3:** Filas após treinamento.

Fila	Tempo de Espera (min)
Cadastro	12,14
Triagem	10,08
Consulta Adulto	38,45
Consulta Infantil	33,74
Exame de Sangue	40,82
Raio X	8,08
Medicação	4,10

**Tabela 4:** Filas por grupo de risco após treinamento.

Paciente	Tempo de Espera Total (triagem + consulta) (min)	Meta (min)
Vermelho	0	0
Laranja Adulto	14,84	10
Amarelo Adulto	26,82	60
Verde Adulto	87,25	120
Laranja Infantil	14,83	10
Amarelo Infantil	17,86	60
Verde Infantil	37,69	120

Analisando as tabelas, vê-se uma redução dos tempos de espera. Apesar de a meta de espera não ter sido alcançada para os pacientes de classificação laranja, a melhora na eficiência é significativa. No entanto, a fila para exames de sangue apresenta uma espera média de 40 minutos. Pode-se dizer que, a restrição foi transferida. Ao invés de externa, a limitação agora é interna, pois os recursos, materiais e humanos desse processo se sobrecarregaram pela agilidade ganha nos processos anteriores.

Havendo interesse na elevação da restrição ao seu nível máximo pode ser realizada pela redução no tempo de estabilização de pacientes classificados como vermelhos. Alguns hospitais de emergência têm investido em uma gama de equipamentos portáteis para ambulâncias de maneira a enriquecer o atendimento prévio e minimizar o tempo de estabilização dos pacientes no hospital.

### 5.3 Cenário proposto 02 - Aumento da quantidade de recursos no exame de sangue

Para diminuir o tempo de espera para exame de sangue, a alternativa seria aumentar a capacidade de atendimento. Assim a empresa precisaria investir na contratação de funcionários ou diminuir o número de requisição de exames.

Considerando que a empresa consiga disponibilizar uma sala para exames e contratar mais dois auxiliares de laboratório, esse segundo cenário foi simulado. Os resultados estão demonstrados na Tabela 5.

**Tabela 5:** Filas após aumento da capacidade de exames de sangue.

Fila	Tempo de Espera (min.)
Cadastro	9,34
Triagem	8,81
Consulta Adulto	46,97
Consulta Infantil	22,57
Exame de Sangue	2,38
Raio X	14,83
Medicação	4,38

Analisando a Tabela 5, o incremento de eficiência é inquestionável. A espera na fila caiu de 40 para 2 minutos. No entanto, essa alternativa possui um alto custo, devido ao investimento em estrutura e em pessoal. Cabe à organização decidir sobre sua viabilidade.

### 5.4 Cenário proposto 03 - Diminuição de requisição de exames

Como informado pela Unidade, um grande número de pacientes que chegam ao sistema são pacientes que poderiam agendar consulta em um Posto de Saúde. Alguns médicos afirmam que, às vezes, passam o exame para poder dispensar o paciente que não necessita de atendimento. Houve relato de pessoas que se dirigiram à Unidade para exames de rotina por comodidade.

Assim, o terceiro cenário considera que os médicos diminuem o número de requisições de exames para pacientes de classificação verde. Essa diminuição se dá na proporção de 30% dentro do conjunto de pacientes verdes. Apesar de parecer significativa, é uma redução simbólica feita para analisar a sensibilidade do sistema nesse aspecto, pois representa apenas uma parte do número de exames pedidos sem necessidade emergente no pronto atendimento.

**Tabela 6:** Filas com diminuição da requisição de exames.

Fila	Tempo de Espera (min.)
Cadastro	10,81
Triagem	10,07
Consulta Adulto	49,84
Consulta Infantil	18,95
Exame de Sangue	22,34
Raio X	12,53
Medicação	5,94

A análise da Tabela 6 mostra uma redução de quase 50% no tempo de fila dos exames de sangue. Apesar de ser um valor alto pode ser diminuído se a alternativa for tratada com maior rigor, ao sentido de não requisitar exames que não sejam realmente emergenciais.

Além da possibilidade de elevação, a presente alternativa não representa custo adicional à instituição, em contrapartida, é realizada economia com exames feitos desnecessariamente.

## 5.5 Comparação dos cenários

A Tabela 7 abaixo mostra uma comparação dos resultados adquiridos com os cenários para melhor visualização de suas vantagens e desvantagens.

**Tabela 7:** Comparação dos tempos de espera em fila de cada cenário.

Filas	Tempo de Espera (min.)			
	Cenário Atual	Cenário 01	Cenário 02	Cenário 03
Cadastro	14,98	12,14	9,34	10,81
Triagem	27,24	10,08	8,81	10,07
Consulta Adulto	64,08	38,45	46,97	49,84
Consulta Infantil	20,91	33,74	22,57	18,95
Exame de Sangue	16,61	40,82	2,38	22,34
Raio X	9,47	8,08	14,83	12,53
Medicação	4,04	4,10	4,38	5,94

## 7. Considerações Finais

A simulação do ambiente do pronto atendimento demonstrou que há uma espera por parte dos pacientes, maior que a indicada para uma Unidade de Pronto Atendimento. A análise pela Teoria das Restrições apontou o fator limitante do sistema como sendo o processo de triagem, direcionando a sugestão da primeira alternativa.

Após a simulação do primeiro cenário, percebeu-se um aumento de mais de 100% na espera para realização de exames de sangue quando comparado com o cenário atual, o que demonstra uma transferência de restrição da triagem para o exame de sangue. Dois novos cenários foram propostos em complemento. O segundo cenário apresentou uma redução de 94% no tempo de fila do laboratório de exame de sangue em comparação com o primeiro, porém, pode não ser a melhor opção devido ao seu alto investimento. O terceiro cenário, apesar de demonstrar redução de apenas 45% no tempo de espera desta fila, apresenta-se como uma alternativa sem custo.

É comum que após a implantação de uma melhoria em um determinado setor, que outro passe a limitar os processos. Cabe à organização estabelecer um ciclo com os passos aqui aplicados para que haja melhoria contínua.

O fato de a amostra para a simulação ter sido coletada em um único pronto Atendimento de Vitória pode não refletir a realidade média da cidade. Além disso, é relevante considerar para próximos estudos a qualificação e quantificação da amostra antes da construção do modelo, para a geração de resultados mais acurados.

A simulação computacional representou uma valiosa ferramenta de análise de sistemas, pois, quando devidamente implementada proporciona a visualização do funcionamento dos mesmos. Além disso, permite que se analise o impacto de decisões gerenciais sobre o sistema, sem que haja o dispêndio da implementação real, onde os resultados são difíceis de se prever instintivamente.

Anexa à essa ferramenta, a Teoria de Restrições representou um importante complemento à análise e gestão de processos, fazendo dessa união um método mais econômico, e, razoavelmente confiável de gerenciamento de decisões.

## 8. Referências

BEST, A. M. et al. (2014). Using Discrete Event Computer Simulation to Improve Patient Flow in a Ghanaian Acute Care Hospital. American Journal of Emergency Medicine. pg 917-922.

CHWIF, L.; MEDINA, A. C. (2010). Modelagem e Simulação de Eventos Discretos: Teoria e Aplicações. 3 ed. Bravarte. São Paulo.

GOLDRATT, E. M. (2002). A Meta: um processo de melhoria contínua. 2 ed. Editora Nobel.

GONÇALVES, A. A. (2004). Gestão da capacidade de atendimento em hospitais de câncer. Dissertação de Doutorado apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal do Rio de Janeiro.

IBOPE. (1998). Promoção da Saúde - pesquisa de opinião pública realizada sob encomenda do CONASS / FNS.

LIMA, M. E. (1999). A produção e a estrutura de custos dos hospitais públicos: uma aplicação de um modelo translogarítmico. Universidade do Minho; Portugal.

NEVES, M. (2009). Introdução à Estatística e Probabilidade.

PACHECO, R. F.; CARDOSO, T. de M. (2005). Utilização de simulação na análise de sistemas de saúde: um estudo de caso no ambulatório da santa casa de misericórdia de Goiânia. *XII Simpósio de Engenharia de Produção - SIMPEP* - Bauru, SP, Brasil.

PEGORARO, F. (2012). Aplicação da simulação computacional e teoria das Restrições (TOC) para a redução do tempo de espera por atendimento de urgência e emergência em um hospital da região sul do estado do Tocantins. Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia de Produção e Sistemas da Pontifícia Universidade Católica de Goiás.

RAMOS, D.D., LIMA, M. A. D. S. (2003). Acesso e acolhimento aos usuários em uma unidade de saúde de Porto Alegre. Rio Grande do Sul, Brasil. *Cad. Saúde Pública*. Rio de Janeiro, 19(1):27-34. Disponível em: <http://www.scielo.org/pdf/csp/v19n1/14902.pdf/>.

ROCHA, S. A. S. et al. (2005). Gestão Estratégica de Serviços na Área de Saúde. *XII Simpósio de Engenharia de Produção - SIMPEP* – Bauru, SP, Brasil.

SABBADINI, F. et al. (2006). Gestão da capacidade de atendimento e simulação computacional para a melhoria na alocação de recursos e no nível de serviço em hospitais. *III Simpósio de Excelência em Gestão e Tecnologia*.

SAKURADA, N.; MIYAKE, D. I. (2009). Aplicação de simuladores de eventos discretos no processo de modelagem de sistemas de operações de serviços. *Gest. Prod.*, São Carlos, v. 16. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0104-530X2009000100004&lng=en&nrm=iso/](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-530X2009000100004&lng=en&nrm=iso/).

YANG, X.; KOZIEL, S.; LEIFSSON, L. (2013). Computational Optimization, Modelling and Simulation: Recent Trends and Challenges. *Procedia Computer Science*. 18, p. 855 – 860.