

SIMULAÇÃO DE POLÍTICAS DE AGENDAMENTO DE PACIENTES NO AMBULATÓRIO DE OFTALMOLOGIA DE UM HOSPITAL PÚBLICO

Igor Tona Peres

Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225 - Gávea, Rio de Janeiro - RJ, 22451-900
igor.peres@aluno.puc-rio.br

Suzana Gonzaga da Veiga

Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225 - Gávea, Rio de Janeiro - RJ, 22451-900
suzana.gonzaga@aluno.puc-rio.br

Silvio Hamacher

Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225 - Gávea, Rio de Janeiro - RJ, 22451-900
hamacher@puc-rio.br

Fernando Luiz Cyrino Oliveira

Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225 - Gávea, Rio de Janeiro - RJ, 22451-900
cyrino@puc-rio.br

Thaís Spiegel

Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225 - Gávea, Rio de Janeiro - RJ, 22451-900
thaisspiegel@gmail.com

RESUMO

O setor de saúde brasileiro vem crescendo nos últimos anos e já representa cerca de 8% do PIB. Entretanto, observa-se que este crescimento não se traduziu em melhora real do sistema. Em alguns hospitais públicos brasileiros, como o Hospital Universitário Pedro Ernesto (HUPE), o agendamento de pacientes é feito alocando todos no início do turno. Esse agendamento busca maximizar a utilização dos médicos, mas costuma gerar grandes filas. Nesse contexto, o artigo tem como objetivo estudar um caso real de agendamento no ambulatório de oftalmologia do HUPE e propor uma política que mantenha a utilização dos recursos alta, mas que reduza o tempo de fila. O estudo testou diferentes políticas por meio de simulação de eventos discretos, propondo a mais adequada. As simulações sugerem que a política proposta pode ocasionar redução de 55% no número máximo de pessoas na fila e 60% no tempo médio de fila.

PALAVRAS CHAVE. Agendamento de Consultas, Simulação de Eventos Discretos, Saúde. Tópicos (PO na Área de Saúde, Simulação)

ABSTRACT

The Brazilian health care sector has been growing in recent years and now accounts for about 8% of GDP. However, it is observed that this growth has not been translated into real improvement of the system. In some Brazilian public hospitals, such as Hospital Universitário Pedro Ernesto (HUPE), patient appointment-scheduling is done allocating all patients at the beginning of the day. This schedule seeks to maximize the utilization of doctors, but usually generates long queues. In this context, the article aims to study a real case in the ophthalmology ambulatory of HUPE and to propose a policy that maintains the resource utilization high, but reduces waiting time. The study tested different scheduling policies through discrete event simulation, proposing the most appropriate policy. The simulations indicate that the proposed policy can lead to a reduction of 55% in the maximum queue size and 60% in the average waiting time.

KEYWORDS. Appointment Scheduling, Discrete Event Simulation, Health Care. Paper topics (OR in Health, Simulation)

1. Introdução

Segundo a *World Health Organization* [WHO 2014], as despesas de saúde no Brasil vêm crescendo nos últimos anos, de forma que, em 2014, os gastos somaram 8,3% do total do PIB brasileiro. Entretanto, observa-se que este crescimento não se traduziu em melhora real do sistema, que vem sofrendo uma série de problemas nas mais diversas áreas da saúde.

O tempo de espera dos pacientes e o congestionamento da sala de espera são dois dos elementos de qualidade tangíveis em gestão de serviços de saúde [Cayirli e Veral 2003]. Dessa maneira, sistemas de agendamento de pacientes bem projetados têm o potencial de aumentar a utilização de pessoas e recursos de saúde caros e reduzir o tempo de espera dos pacientes. Pesquisas indicam que a espera excessiva é, muitas vezes, a principal razão da insatisfação dos pacientes em saúde [Huang 1994].

Em alguns hospitais públicos brasileiros, como o Hospital Universitário Pedro Ernesto (HUPE) da Universidade do Estado do Rio de Janeiro (UERJ), o agendamento de pacientes é feito alocando todos os pacientes no início dos turnos. Esse tipo de agendamento busca maximizar a utilização dos médicos do turno, e se proteger das altas taxas de absenteísmo. Entretanto, o nível de serviço dessa regra costuma ser muito baixo, visto que são geradas grandes filas no início do turno, fazendo com que o tempo de fila dos pacientes seja muito elevado.

Isto posto, os principais objetivos deste artigo são: (i) estudar um caso real de agendamento de pacientes no ambulatório de oftalmologia do HUPE; (ii) analisar cenários com diferentes políticas de agendamento de pacientes existentes na literatura por meio de um modelo de simulação de eventos discretos; e (iii) propor uma política para o ambulatório que mantenha uma alta utilização dos recursos e que possua o melhor resultado em termos de indicadores de fila. Além disso, busca-se analisar a sensibilidade do modelo com cenários variando o número de médicos e a proporção entre classes de pacientes (primeira vez e retorno).

A estrutura do presente artigo se organizará como se segue. A Seção 2 tratará das questões abordadas na literatura relacionadas às políticas de agendamento. A Seção 3 analisará os dados do estudo de caso. O método utilizado será apresentado na Seção 4. Os resultados serão tratados na Seção 5 e as conclusões na Seção 6.

2. Revisão da literatura

O estudo do agendamento de pacientes em serviços ambulatoriais tem atraído o interesse de muitos acadêmicos e profissionais ao longo dos últimos 60 anos, começando com os trabalhos pioneiros de [Bailey 1952] e [Lindley 1952]. Para consolidar os trabalhos feitos até então, [Cayirli e Veral 2003] fizeram uma extensiva revisão da literatura nesse tema. Segundo os autores, o objetivo de estudar agendamento de pacientes é encontrar uma política que otimize determinada medida de desempenho, considerando a incerteza do ambiente analisado.

Segundo esses autores, no que tange à definição e à formulação do problema de agendamento, pode-se dizer que os ambulatórios são considerados sistemas de filas. O caso mais simples é quando todos os pacientes agendados chegam pontualmente no horário marcado e apenas um único médico está atendendo com tempos de processamento estocásticos. A formulação é mais complexa quando vários médicos e serviços são considerados. A presença de pacientes impontuais, *no-shows* (não comparecimento) e pacientes emergenciais pode acabar perturbando o desempenho do sistema.

Ainda segundo os autores, no que tange às medidas de desempenho, foi verificado que existe uma grande variedade de critérios utilizados na literatura para avaliar as políticas de agendamento. Alguns estudos listam os resultados em termos da média do tempo de espera dos pacientes $E(W)$, da média do tempo em que o médico está ocioso $E(I)$ e/ou da média do tempo em que o médico está fazendo horas extras $E(O)$. Ao dar pesos relativos em termos de custo do tempo de espera dos pacientes (C_p), de custo da ociosidade (C_d) e das horas extras do médico (C_o), a maioria dos estudos analíticos utiliza a seguinte função objetivo que minimiza o custo total esperado do sistema: $Min E(TC) = E(W)C_p + E(I)C_d + E(O)C_o$.

Segundo [Cayirli et al. 2006], há dois fatores de decisão em uma política de agendamento, quais sejam: (1) a regra de sequenciamento, ou seja, a ordem em que os pacientes são agendados com base em um sistema de classificação dos pacientes e (2) a regra de agendamento, ou seja, o número de pacientes agendados para cada período e a duração dos intervalos de agendamento. As regras de sequenciamento se tornam necessárias quando há grupos de pacientes conhecidos por serem distintos em termos de vários atributos, como, por exemplo, em termos de tempo de serviço. No caso em estudo, o ambulatório possui dois tipos de pacientes (primeira vez e retorno) e, por isso, algumas regras serão testadas utilizando o sequenciamento dessas classes.

Os autores [Cayirli et al. 2006] analisaram sete regras de agendamento por meio de simulação de eventos discretos em diversos cenários e perceberam que as regras de intervalo fixo IBFI e 2BEG tiveram melhores desempenhos gerais. Essas duas melhores regras propostas pelos autores, além da regra 2BGDM, serão utilizadas neste trabalho e estão apresentadas na Figura 1.

Regra	Descrição	Exemplo de funcionamento
IBFI	Agenda pacientes individualmente em intervalos iguais ao tempo médio de atendimento.	
2BEG	Igual a regra IBFI, com bloco inicial de dois pacientes.	
2BGDM	Combinação da 2BEG com a regra DOME, que agenda os pacientes do início do turno em intervalos mais curtos, os do meio em intervalos maiores, e o restante em intervalos mais curtos, comparado com a IBFI.	

Figura 1 – Regras de agendamento IBFI, 2BEG e 2BGDM.

Fonte: Adaptada de [Cayirli et al. 2006].

Ainda, [Ho e Lau 1992] testaram sete regras de agendamento (2BEG, 3BEG, 4BEG, entre outras) em 27 ambientes diferentes. As regras 3BEG e 4BEG são variações da 2BEG utilizando, respectivamente, três e quatro pacientes no primeiro horário, ao invés de dois pacientes, conforme apresentado na Figura 2. Estas duas regras têm como objetivo aumentar a utilização do médico no início do turno, já que este é o horário onde ele está mais ocioso devido ao fato do sistema começar vazio. [Ho e Lau 1992] concluíram que as regras 2BEG e 4BEG tiveram o melhor desempenho na maioria dos cenários. Apesar das regras 3BEG e 4BEG não estarem presentes no estudo de [Cayirli et al. 2006], essas são interessantes de serem aplicadas em locais onde se deseja garantir uma alta utilização dos médicos, sendo este um indicador mais forte do que o próprio tempo de fila, como é o caso do hospital em estudo.

Regra	Descrição	Exemplo de funcionamento
3BEG	Igual a regra IBFI, com bloco inicial de três pacientes.	
4BEG	Igual a regra IBFI, com bloco inicial de quatro pacientes.	

Figura 2 – Regras de Agendamento 3BEG e 4BEG.

Fonte: Elaboração própria.

Os autores [Vanden Bosch e Dietz 2000] analisaram as políticas de agendamento para uma clínica e concluíram que não existe uma regra para a sequência ótima, sendo difícil generalizar um resultado sobre a ordenação de pacientes por meio do valor esperado ou da variância do tempo de serviço.

Segundo [Cayirli et al. 2006], a maioria dos estudos de simulação existentes na literatura analisa o ambiente específico de determinada clínica, tornando os resultados desses estudos difíceis de serem generalizados. Por isso, como sugerido por [Shafer e Smunt 2004], combinar o poder e a flexibilidade da simulação com a utilização de dados empíricos passa a ser uma das maneiras mais efetivas de solucionar essa lacuna existente entre o rigor acadêmico e a aplicabilidade gerencial. Segundo [Robinson 2005], a Simulação de Eventos Discretos (SED) é um dos métodos de modelagem mais populares utilizados para analisar sistemas complexos. Em outro trabalho, [Cayirli e Veral 2003] também afirmam que uma vantagem da simulação em relação a estudos analíticos é a capacidade de modelar sistemas complexos de filas e representar variáveis ambientais. Eles observam que os estudos encontrados na literatura simulam cenários para avaliar o desempenho de cada política de agendamento e para compreender a relação entre fatores ambientais e diferentes medidas de desempenho. Em serviços de saúde, o uso de SED ganhou força por causa de três vantagens principais: a capacidade de incluir atributos individuais, a flexibilidade para expressar a dinâmica e a lógica do processo, e o desenvolvimento avançado de pacotes de *software* [Brailsford 2007].

3. Aplicação

O estudo foi elaborado por meio da coleta de dados no ambulatório de oftalmologia do Hospital Universitário Pedro Ernesto (HUPE). O ambulatório em questão atende pacientes com demandas específicas da área de oftalmologia. Os tipos de pacientes atendidos são de primeira vez e retorno. O atendimento é feito por médicos residentes de primeiro, segundo e terceiro ano (R1, R2 e R3), supervisionados por preceptores (médicos especialistas), sendo que os residentes R1 e R2 realizam atendimento diário dos pacientes de primeira vez e de retorno e os residentes R3 atendem alguns casos de pacientes de retorno em alguns dias da semana.

Como o sistema corporativo não possuía, à época da pesquisa, as informações de horários de atendimento e chegada de pacientes, os dados foram coletados através de medições de tempo feitas *in loco*, com registro manual dos tempos das atividades e de informações relevantes para o estudo. Foram coletados os seguintes dados: comparecimento ou não comparecimento, horário agendado, horário da chegada real, horários de início e término do atendimento. Além da coleta dos dados, foram feitas entrevistas com os profissionais para auxiliar no entendimento do processo, que é apresentado na Figura 3.

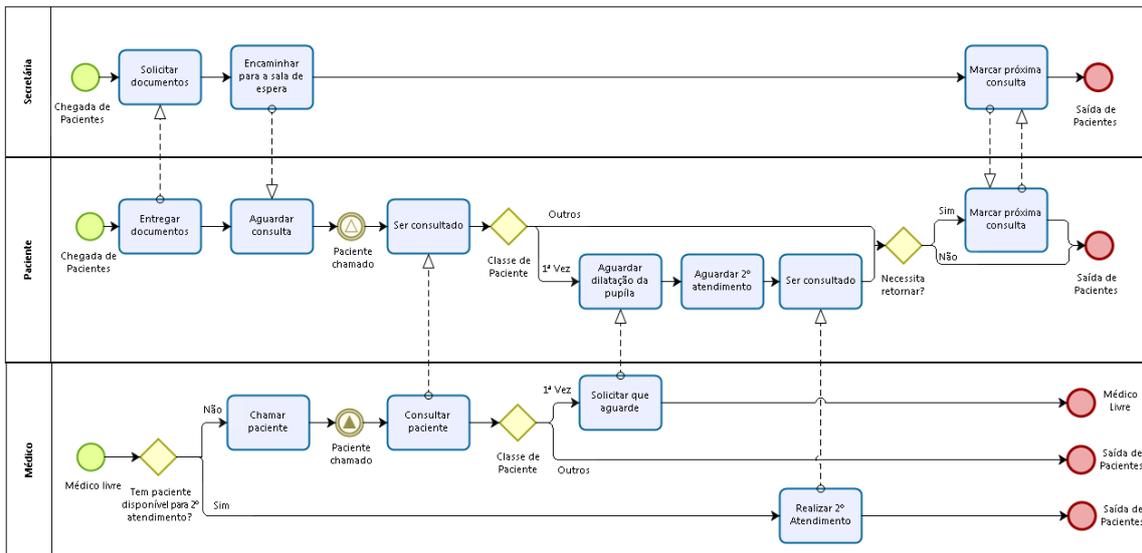


Figura 3 – Modelo conceitual do processo do ambulatório

Fonte: Elaboração própria.

Após análise das informações obtidas nas entrevistas e tratamento dos dados, foram propostas distribuições de probabilidade para representar o funcionamento do sistema. Como não existiam dados de atendimento no sistema corporativo, buscou-se propor distribuições que representassem a realidade de acordo com as medições *in loco*. Como o número de dados coletados manualmente era insuficiente para testes estatísticos de aderência, foram sugeridas distribuições triangulares para representar de forma estocástica os tempos de atendimento de pacientes, de acordo com o valor mínimo, mais provável e máximo. Dessas análises, chegou-se a uma série de características relacionadas ao ambulatório em questão, listadas a seguir:

- O número de médicos em atendimento varia entre três e quatro por dia, e os pacientes são agendados conforme a disponibilidade de médicos por dia.
- O horário de trabalho é de 8h às 17h, com parada de almoço de 12h às 13h.
- Os pacientes são agendados através de marcação única de horário para todos os pacientes, sendo os do turno da manhã agendados para as 8h e os do turno da tarde agendados para as 13h.
- O atendimento inicial do paciente ao entrar no ambulatório é feito por uma secretária, com tempo *Triangular*(1; 1.5; 2), em minutos;
- O processo de serviço dos médicos difere para cada classe de paciente. Os pacientes de primeira vez devem ter um atendimento completo, tendo que passar por duas etapas, com tempos de atendimento, em minutos, *Triangular* (12; 16; 30) e *Triangular* (7; 11; 32), respectivamente. Após o primeiro atendimento, os pacientes devem aguardar 20 minutos para ocorrer a dilatação da pupila e, enquanto isso, o médico pode atender outro paciente. Os pacientes de retorno possuem apenas um atendimento, sendo este mais simples e com distribuição *Triangular* (5; 11; 27), em minutos.
- A taxa média de *no-show* foi de 46,51% para pacientes de primeira vez e de 15% para pacientes de retorno, obtida pela relação entre o número de pacientes que não compareceram à consulta agendada e o número total de pacientes agendados.
- O atendimento é feito por ordem de chegada (disciplina *First In First Out* – FIFO).
- A prioridade de atendimento pertence aos pacientes de primeira vez, sendo a proporção aproximada analisada de um paciente de primeira vez para três pacientes de retorno.
- Não foi possível verificar a distribuição de impontualidade do paciente, uma vez que, como todos são agendados no primeiro horário, eles acabam não respeitando o horário agendado, pois sabem que não serão atendidos pontualmente. Para os cenários propostos, o valor da impontualidade foi assumido como uma distribuição *Normal* (-15; 25), sendo este o valor

encontrado na maioria dos estudos da literatura que analisaram a impontualidade do paciente [Cayirli et al. 2006], [Cayirli et al. 2008], [Sepúlveda e Berroeta 2012].

4. Método

O método utilizado para modelar esse sistema de fila e suas variáveis ambientais foi a Simulação de Eventos Discretos, conforme as orientações de [Cayirli e Veral 2003], [Robinson 2005] e [Brailsford 2007]. O modelo de simulação desenvolvido teve como objetivo avaliar o desempenho de cada regra de agendamento no processo do ambulatório. O método compreende a simulação terminal com dois turnos, sendo o primeiro turno de 8h às 12h e o segundo de 13h às 17h.

4.1 Modelo de Simulação

A partir das características observadas na seção anterior, construiu-se um modelo de simulação no *software* ProModel® versão 8.6 para representar o processo do ambulatório em questão (Figura 4). Este tem como característica ser um modelo de simulação terminal, contemplando os dois turnos de trabalho. Como o sistema começa vazio em ambos os turnos, ou seja, fora do horário de trabalho não há pacientes na sala de espera nem há pacientes sendo atendidos, logo, não há necessidade da utilização de um tempo para o aquecimento da simulação.

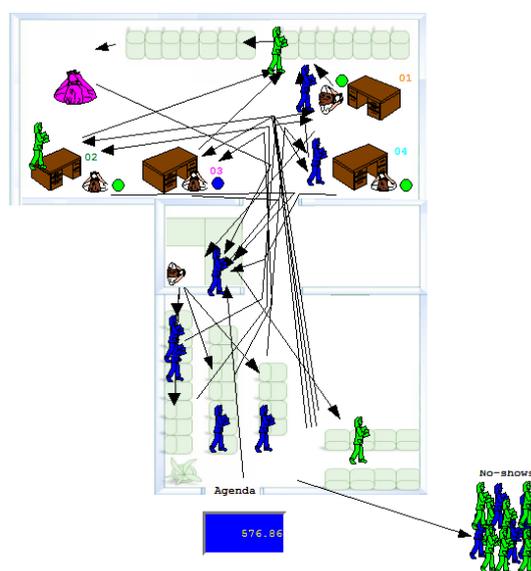


Figura 4 – Modelo de Simulação

Dessa maneira, procurou-se fazer com que o modelo representasse da melhor forma possível o caso em estudo e, para que fosse atribuída maior confiabilidade a ele, cada simulação teve 500 replicações de um dia de funcionamento para cada um dos cenários. Para verificar se o número de replicações estava adequado, foi feito um estudo variando esse número entre 100 e 999 replicações e percebeu-se que a variação entre os resultados de 500 e 999 replicações era muito pequena, conforme Tabela 1. Portanto, escolheu-se utilizar 500 replicações para facilitar a geração dos resultados, dado que foram simuladas diversas combinações de regras e cenários diferentes.

Tabela 1 – Verificação do Modelo

Nº de Replicações (n)	Tempo Médio de Fila (min)	Tempo Máximo de Fila (min)	Primeiro Atendimento		Segundo Atendimento		Horário de fim do dia (h)	Nº de pacientes atendidos
			Tempo Médio de Fila (min)	Tempo Máximo de Fila (min)	Tempo Médio de Fila (min)	Tempo Máximo de Fila (min)		
100	104.03	202.92	100.10	198.41	20.69	62.44	16.75	49.03
200	102.89	198.02	99.10	194.36	20.06	60.59	16.76	48.87
300	103.19	200.02	99.52	195.94	19.39	59.52	16.78	49.03
400	103.43	201.03	99.63	196.59	19.86	61.82	16.78	49.04
500	103.53	200.85	99.83	196.64	19.52	60.31	16.79	49.18
999	103.31	200.71	99.64	196.36	19.69	60.79	16.76	49.22

No cenário inicial, a simulação caracterizou-se por representar o funcionamento atual do consultório, conforme descrito na seção 3. Foram testadas cinco regras de agendamento da literatura, a IBFI, que é a regra usualmente utilizada em consultórios privados, e outras quatro regras que têm como objetivo manter alta a utilização do médico ao longo do dia, sendo estas a 2BEG, 3BEG, 4BEG e três configurações da 2BGDM, além da regra atual que agenda todos os pacientes no início do turno. Além disso, para avaliar o impacto da regra de prioridade de atendimento no início do dia específica do HUPE que é dada aos pacientes de primeira vez, as regras 2BEG e 2BGDM foram testadas alocando no início do dia os pacientes de primeira vez e, em seguida, os de retorno, além dos testes com as suas configurações normais. Entretanto, para as regras de agendamento propostas, essa regra de prioridade não será obrigatória, uma vez que esta somente é necessária quando o sistema está sobrecarregado (com grandes filas), como é o caso da regra de agendamento atual, que agenda todos os pacientes no início do turno. Quando o sistema está controlado, os pacientes podem ser atendidos sem prioridade. Para todos os testes, sempre foi tido como base as mesmas características citadas anteriormente, alterando apenas a regra de agendamento.

Foram gerados dois cenários, um com três e outro com quatro médicos disponíveis, sendo o número de pacientes agendados proporcional ao número de médicos (32 agendados para o cenário com três médicos e 42 agendados para o cenário com quatro médicos). Além disso, para cada um desses cenários, foi feito também um cenário que avalia a proporção de pacientes que o hospital deseja adotar no futuro, em aderência as orientações do sistema de regulação do município do Rio de Janeiro (SISREG), que compreende a proporção de um paciente de primeira vez para dois pacientes de retorno, além do cenário atual de um para três.

A partir das 40 simulações feitas, relacionando cada um dos quatro cenários a cada uma das dez regras de agendamento, foram obtidos os resultados a partir da ferramenta *Output Viewer* do *software* ProModel® versão 8.6, sendo que alguns dados foram exportados para uma planilha para obter resultados mais específicos, visando os objetivos desse estudo.

4.2 Indicadores de desempenho

Para o cenário atual, foi escolhida como regra de agendamento para o consultório aquela que manteve o “horário de fim do dia” próximo do resultado do cenário atual e apresentar o melhor desempenho conjunto nos seguintes indicadores: tempo médio e máximo de fila de pacientes, porcentagem de pacientes esperando até 30 minutos na fila, tamanho médio da fila, tamanho máximo da fila.

4.3 Premissas básicas

Com o intuito de tornar o estudo possível, foi feita uma série de premissas. Primeiramente, o modelo parte da premissa de atendimento ininterrupto como uma realidade do consultório, ou seja, após a realização de um exame, o próximo paciente da fila é chamado, independente do horário atual, não havendo tempos de intervalo entre duas consultas consecutivas e estando os médicos e a secretária sempre disponíveis durante os dois turnos de trabalho. Assume-se que a pontualidade do paciente independe do horário de consulta. O estudo assume que os tempos de serviço são independentes e identicamente distribuídos, de forma que o tempo de um serviço não dependa do tamanho da fila no momento nem de outro atendimento. Por último, assume-se que cada paciente somente é atendido uma única vez por visita e que a probabilidade de *no-show* é igual para todos os pacientes, em cada classe.

4.4 Validação do Modelo

Ao simular o cenário atual com as características observadas na Seção 4.1 e com a regra atual do HUPE, pôde-se comparar os resultados encontrados na simulação com os resultados empíricos (Tabela 2). A fim de validar o modelo, considerou-se um atraso de 30 minutos para os médicos para os turnos da manhã e da tarde. Entretanto, para os cenários propostos, os médicos serão considerados como pontuais, já que ações estão sendo tomadas nesse sentido pela diretoria. A proporção adotada no modelo é a de um paciente de primeira vez para três pacientes de retorno, segundo observações *in loco*. Foram considerados três médicos no plantão de atendimento por turno. Além disso, para analisar o impacto da utilização das distribuições triangulares, a distribuição de tempo de atendimento foi testada também como sendo uma *Lognormal*, já que esta vem sendo utilizada em testes recentes para modelar tempos de atendimento em consultórios [Sepúlveda e Berroeta 2012], [Klassen e Yoogalingam 2014], [Bard et al. 2016].

Tabela 2 – Validação do Modelo

Política	Tempo Médio de Fila (min)	Tempo Máximo de Fila (min)	Primeiro Atendimento		2º Atendimento		Horário de fim do dia (h)	Nº de pacientes atendidos
			Tempo Médio de Fila (min)	Tempo Máximo de Fila (min)	Tempo Médio de Fila (min)	Tempo Máximo de Fila (min)		
Empírico	106.65	195.00	91.56	195	22.28	66.00	16.75	49
Atual (Triangular)	103.53	200.85	99.83	196.64	19.52	60.31	16.79	49.18
Atual (Lognormal)	102.42	201.04	98.03	198	18.73	58.28	16.78	48.97

Percebe-se que os resultados da simulação para todos os indicadores de desempenho estão coerentes com a realidade do consultório estudado (indicadores empíricos), o que valida o modelo desenvolvido. Além disso, percebe-se que a utilização da distribuição *Triangular* para os tempos de atendimento teve resultados bem próximos em relação ao modelo com distribuição *Lognormal*, mostrando que não há diferença entre utilizar esses dois tipos de distribuições nesse caso. Dessa forma, será utilizada a distribuição *Triangular* para os tempos de atendimento.

5. Resultados

Os resultados da simulação para o cenário com três médicos com proporção atual de um paciente de primeira vez para três pacientes de retorno para os indicadores de desempenho considerados são apresentados na Tabela 3. Constata-se que a regra 3BEG possui os melhores indicadores de fila dentre as regras que mantem o horário de fim do dia próximo do horário da regra Atual (16h15min). A regra 3BEG teve um tempo médio de fila de 22 minutos, enquanto que a regra atual tinha um tempo de 71 minutos. Essa redução também ocorreu para os outros

indicadores de fila, como o tempo máximo (de 161min para 85min), a fila média (de cinco para dois pacientes) e a fila máxima (de 21 para nove pacientes), além do aumento no percentual de esperas de no máximo 30min (de 23% para 70%). As regras IBFI, 2BEG e as três configurações da 2BEGDM tiveram melhores resultados em termos de indicadores de fila, mas terminaram o dia mais tarde do que a regra atual. Como nesse caso deseja-se terminar o dia próximo das 16h15min, de forma a permitir que outras atividades sejam feitas até o final do turno, a regra 3BEG será a regra escolhida. Quanto às regras IBFI-1VEZBEG e 2BEG-1VEZVEG, essas não obtiveram resultados tão bons quanto às outras, o que pode estar relacionado ao fato de, como nessa regra os pacientes de primeira vez são agendados no começo do turno, isso pode estar piorando o desempenho do sistema, uma vez que estes são os pacientes de maior taxa de *no-show*. Portanto, não se recomenda fazer qualquer tipo de priorização de agendamento de pacientes de primeira vez no início do dia.

Tabela 3 – Resultados do cenário com três médicos com proporção atual de um para três

Política	Tempo Médio de Fila (min)	Tempo Máximo de Fila (min)	Horário de fim do dia (h)	% Pacientes até 30min na Fila	Fila Média	Fila Máxima	Tamanho Máximo da Fila (com acompanhante)	Nº de pacientes atendidos
Atual	70.96	160.59	16:14	22.68%	5.01	20.89	41.78	49.03
IBFI	10.73	58.09	16:54	88.72%	1.11	5.74	11.48	49.39
2BEG	15.17	70.04	16:36	81.19%	1.46	7.19	14.38	49.28
IBFI-1VEZBEG	16.93	77.77	16:47	76.81%	1.45	7.08	14.16	49.16
2BEG-1VEZBEG	21.46	86.95	16:32	69.87%	1.80	7.91	15.82	49.28
2BGDM 1	14.46	69.35	16:39	82.42%	1.39	6.83	13.66	49.37
2BGDM 2	17.98	76.01	16:30	76.57%	1.67	8.04	16.08	49.26
2BGDM 3	20.39	78.87	16:25	72.15%	1.86	8.83	17.66	49.26
3BEG	22.02	84.17	16:21	69.70%	1.95	8.81	17.62	49.18
4BEG	30.72	100.53	16:09	56.14%	2.51	10.80	21.60	49.22

Essa mesma análise foi realizada para os outros três cenários e os resultados consolidados das melhores regras de cada cenário, comparados com o resultado da regra atual, estão apresentados na Tabela 4.

Tabela 4 – Resumo das melhores regras de agendamento em cada cenário, comparadas com a atual.

Cenário	Política	Tempo Médio de Fila (min)	Tempo Máximo de Fila (min)	Horário de fim do dia (h)	% Pacientes até 30min na Fila	Fila Média	Fila Máxima	Fila Máxima (com acompanhante)	Nº de pacientes atendidos
3 médicos e proporção 1/3	Atual	70.96	160.59	16:14	22.68%	5.01	20.89	41.78	49.03
3 médicos e proporção 1/3	3BEG	22.02	84.17	16:21	69.70%	1.95	8.81	17.62	49.18
3 médicos e proporção 1/2	Atual	75.05	167.6	16:22	22.77%	4.97	20.49	40.98	47.65
3 médicos e proporção 1/2	3BEG	24.96	91.37	16:25	65.61%	1.94	8.65	17.30	47.46
4 médicos e proporção 1/3	Atual	78.46	178.01	16:31	21.37%	7.14	27.41	54.82	64.71
4 médicos e proporção 1/3	3BEG	28.62	104.6	16:32	60.32%	3.10	12.02	24.04	64.82
4 médicos e proporção 1/2	Atual	80.87	185.89	16:39	20.83%	6.86	26.60	53.20	62.27
4 médicos e proporção 1/2	3BEG	31.45	108.86	16:41	65.01%	3.01	11.89	23.78	62.41

Percebe-se que a regra 3BEG continuou sendo a melhor regra para os outros cenários analisados. Em termos de tempo de fila, essa regra proporcionou uma redução de mais de 60% no tempo médio e 40% do tempo máximo em todos os cenários, mantendo o término do dia próximo do horário de fim atual. Percebe-se que a alteração na proporção de pacientes atual (um de primeira vez para três de retorno) para o cenário futuro (um de primeira vez para dois de retorno), fez com que o desempenho do sistema piorasse como um todo, tanto para o cenário com três quanto para o cenário com quatro médicos. Isso pode ser explicado uma vez que, como a proporção de primeira vez aumenta e como esta classe de paciente é a que possui o maior tempo de atendimento, o sistema torna-se mais congestionado de pacientes de maior complexidade de atendimento, o que piora os indicadores de fila. Além disso, essa alteração na proporção também afetou o número de pacientes atendidos, fazendo com que ocorresse uma diminuição nos atendimentos, já que os pacientes de primeira vez possuem maior *no-show*. Pode-se observar também que o sistema com quatro médicos foi pior em todos os indicadores em relação ao sistema com três médicos (com exceção do número de pacientes atendidos), mesmo tendo agendado pacientes de forma proporcional ao número de médicos disponíveis (cerca de 11 pacientes por médico). Isso pode ser explicado uma vez que o sistema fica mais congestionado, mesmo tendo o mesmo número de atendimentos a fazer por médico.

Com o intuito de entender melhor a diferença dos resultados encontrados para a regra 3BEG em relação aos resultados da regra Atual, vale mostrar também a evolução da fila média de pacientes ao longo do tempo, conforme apresentado na Figura 5.

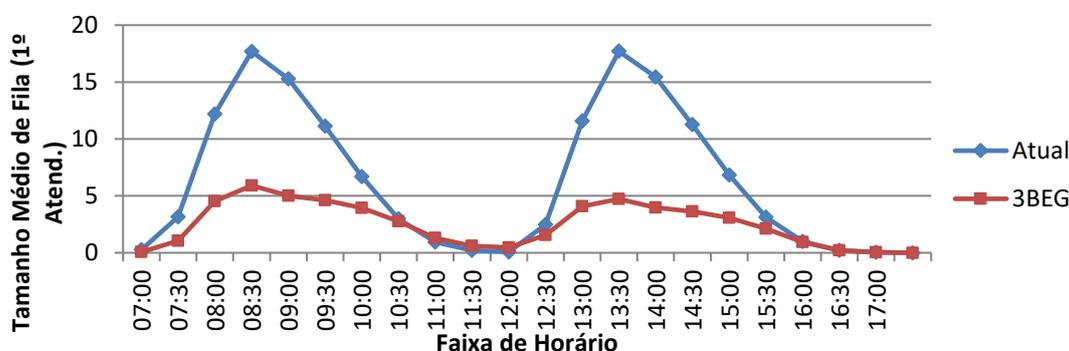


Figura 5 – Comparação da fila média de pacientes no tempo (cenário com três médicos e proporção 1/3)

Percebe-se que o número médio de pacientes aguardando na fila por horário do dia é mais baixo para a regra 3BEG em todos os horários do dia, comparado com a regra Atual. Pode-se notar que, na regra 3BEG, a fila médica não passa de seis pacientes no horário de pico, enquanto que, na regra atual, existem 18 pacientes em média em fila no horário de pico. Esse mesmo comportamento também é observado para os outros três cenários analisados.

6. Conclusões

A partir da análise dos resultados, pode-se concluir que a melhor regra para o agendamento do ambulatório é a 3BEG, visto que apresenta os melhores resultados em termos de tamanho e tempo de fila, mantendo o horário de fim do dia próximo do horário de fim atual.

De acordo com os resultados da simulação, pôde-se verificar que a regra 3BEG consegue reduzir o tempo médio de fila do paciente em mais de 60% e o tempo máximo em mais de 40% do em todos os cenários testados, mantendo a meta de horário de fim do dia e uma alta utilização dos médicos. Dessa forma, o tempo médio de espera que era maior do que 70min na regra atual, não passará de 32min, e o tempo máximo de espera, que era maior do que 2h40min na regra atual, não passará de 1h50min, para todos os cenários propostos.

Além disso, com a implementação da 3BEG, espera-se reduzir o tamanho máximo da fila (incluindo acompanhantes) de 54 para 24 pessoas (cenário com quatro médicos) e de 41 para 17 pessoas (cenário com três médicos). Essa redução de mais de 50% é um fator importante uma vez que a sala de espera do ambulatório tem capacidade para cerca de 25 pessoas, fazendo com que frequentemente os pacientes tenham que aguardar em pé ao lado de fora do ambulatório, o que afeta significativamente a satisfação do paciente (e de seus acompanhantes).

Pela análise do processo, percebeu-se que o retorno não é agendado de forma imediata, de forma que o paciente precisa retornar para remarcar a consulta. Isso acaba aumentando o fluxo de pacientes e tende a atrapalhar o processo da clínica. Recomenda-se, portanto, que o sistema de agendamento do retorno seja alterado para um agendamento imediato.

Também foi evidenciada no processo a alta taxa de *no-show* de primeira vez (46%), recomendando-se a avaliação de ações para tentar minimizar o não comparecimento dos pacientes. Soluções indicadas na literatura, como, por exemplo, efetuar ligações com o intuito de antever melhor os possíveis pacientes que não comparecerão e poder remarcar um novo paciente no lugar, além de encaminhar mensagens de texto uma hora antes da consulta; já foram testadas no HUPE, mas não apresentaram melhora na taxa de *no-show* até o momento. A duplicidade nos canais de comunicação com o paciente, tanto do hospital, como do sistema de regulação municipal, somada ao grau de instrução da população atendida, coloca desafios adicionais a intervenção.

Ressalta-se que a implementação da regra 3BEG no sistema de agendamento do ambulatório traz melhorias para o sistema de atendimento, visto que reduz o número de pacientes na fila e o tempo de espera, resultando em maior satisfação do paciente, além de reduzir o número de pessoas na sala de espera e não permitir que pacientes e acompanhantes fiquem em pé aguardando a consulta. Corrobora, desta forma, com as melhorias em implementação no HUPE no âmbito do programa de humanização.

Dentre as limitações do estudo, ressalta-se a dificuldade no acesso aos dados. Recomenda-se como estudos futuros, realizar cenários com *overbooking* e com redução na taxa de *no-show*.

Referências

- Bailey, N. (1952). A study of queues and appointment systems in hospital outpatient departments with special reference to waiting times. *Journal of the Royal Statistical Society*, 14:185–199.
- Bard, J., Shu, Z., Morrice, D., Poursani, R. e Leykum, L. (2016). Improving patient flow at a family health clinic. *Health Care Management Science*, 19(2):170–191.
- Brailsford, S. C. (2007). Advances and challenges in healthcare simulation modeling. In *Anais da 39th Winter Simulation Conference*, p. 1436–1448, Washington, DC. IEEE.
- Cayirli, T. e Veral, E. (2003). Outpatient scheduling in health care: A review of literature. *Production and Operations Management*, 12(4):519–549.
- Cayirli, T., Veral, E. e Rosen, H. (2006). Designing appointment scheduling systems for ambulatory care services. *Health Care Management Science*, 9:47–58.
- Cayirli, T., Veral, E. e Rosen, H. (2008). Assessment of patient classification in appointment system design. *Production and Operations Management*, 17(3):338–353.
- Ho, C. e Lau, H. (1992). Minimizing total cost in scheduling outpatient appointments. *Management Science*, 38(12):1750–1764.

Huang, X. (1994). Patient attitude towards waiting in an outpatient clinic and its applications. *Health Services Management Research*, 7:2–8.

Klassen, K. e Yoogalingam, R. (2014). Strategies for appointment policy design with patient unpunctuality. *Decision Sciences*, 45(5):881–911.

Lindley, D. V. (1952). The theory of queues with a single server. In *Mathematical Proceedings of the Cambridge Philosophical Society*, 48:277–289. Cambridge University Press.

Robinson, S. (2005). Discrete-event simulation: from the pioneers to the present, what next? *The Journal of the Operational Research Society*, 56(6):619–629.

Sepúlveda, J. e Berroeta, C. (2012). Reglas de decisión para la selección de sistemas de citas basadas en características del paciente y de la institución de salud. *Revista Médica de Chile*, 140(7):867–872.

Shafer, S. M. e Smunt, T. L. (2004). Empirical simulation studies in operations management: Context, trends, and research opportunities. *Journal of Operations Management*, 22:345–354.

Vanden Bosch, P. e Dietz C. (2000). Minimizing expected waiting in a medical appointment system. *IIE Transactions*, 32(9):841–848.

WHO - World Health Organization. (2014). Web page. <http://apps.who.int/gho/data/view.main.HEALTHEXPRATIOBRA>. Acessado: 2016-30-08.