

**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ADMINISTRAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO**

MARCELO OLESKOVICZ

**SISTEMAS DE AGENDAMENTO EM SERVIÇOS DE MEDICINA
DIAGNÓSTICA**

SÃO PAULO-SP

2015

Prof. Dr. Marco Antonio Zago
Reitor da Universidade de São Paulo

Prof. Dr. Adalberto Américo Fischmann
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Roberto Sbragia
Chefe do Departamento de Administração

Prof. Dr. Moacir de Miranda Oliveira Júnior
Coordenador do Programa de Pós-Graduação em Administração

MARCELO OLESKOVICZ

**SISTEMAS DE AGENDAMENTO EM SERVIÇOS DE MEDICINA
DIAGNÓSTICA**

Dissertação apresentada ao Programa de Pós-Graduação em Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, para a obtenção do título de Mestre em Ciências.

Orientador: Prof. Dr. Marcelo Caldeira Pedroso

Versão Corrigida
(versão original disponível na Faculdade de Economia, Administração e Contabilidade)

São Paulo
2015

FICHA CATALOGRÁFICA

Elaborada pela Seção de Processamento Técnico do SBD/FEA/USP

Oleskovicz, Marcelo.

Sistemas de agendamento em serviços de medicina diagnóstica /

Marcelo Oleskovicz. – São Paulo, 2015.

128 p.

Dissertação (Mestrado) – Universidade de São Paulo, 2015.

Orientador: Marcelo Caldeira Pedroso.

MARCELO OLESKOVICZ

**SISTEMAS DE AGENDAMENTO EM SERVIÇOS DE MEDICINA
DIAGNÓSTICA**

Dissertação de mestrado em Administração para a obtenção do título de Mestre em Ciências: Administração, Universidade de São Paulo, Faculdade de Economia, Administração e Contabilidade – Departamento de Administração – Programa de Pós-Graduação em Administração.

Banca Examinadora

.....

Prof. Dr. Marcelo Caldeira Pedroso

.....

Profa. Dra. Ana Maria Malik

.....

Prof. Dr. Jorge Luiz Biazzi

Conceito:

São Paulo, de de 2015

DEDICATÓRIA

À Claudia e ao André, que com amor e compreensão me deram forças para esta conquista.

AGRADECIMENTOS

Ao Professor Marcelo Caldeira Pedroso, pela orientação e convívio durante este trabalho.

Aos professores e equipe do Departamento de Administração da Universidade de São Paulo que, de maneira direta ou indireta, colaboraram para viabilizar a realização desta pesquisa.

RESUMO

Objetivos: este estudo busca propor um modelo de sistema de agendamento para serviços ambulatoriais, tendo como foco os prestadores de serviços de medicina diagnóstica com diversidade de portfólio de serviços agendados. **Método:** a pesquisa foi conduzida na forma de simulação, tendo como unidade de análise um laboratório de medicina diagnóstica de grande porte. Utilizando modelagem quantitativa, os cenários do caso analisado foram reproduzidos para aplicação de modelos de sistemas de agendamento buscados na literatura. **Resultados:** tendo como critérios a compatibilidade com os objetivos da pesquisa e o contexto do caso analisado, foram selecionados modelos de sistemas de agendamento com potencial de aplicação. Adotando indicadores de custos identificados na revisão teórica, o modelo de maior potencial foi avaliado, tendo como referência modelos considerados tradicionais para fins comparativos. Os resultados apontam para desempenho superior do modelo proposto na grande maioria dos cenários analisados. **Conclusões:** a flexibilidade do modelo proposto, no sentido de abranger uma extensa gama de cenários, aliada a desempenho superior e uma baixa complexidade de implantação, o configuram como uma interessante alternativa para aplicação prática.

Palavras-Chave: Serviços em Saúde, Saúde, Sistemas de Agendamento, Medicina Diagnóstica

ABSTRACT

Objectives: This study aims to propose an appointment system model for outpatient services, focusing on diagnostic medicine providers with diverse portfolio of scheduled services. **Method:** this research was conducted in the form of a simulation in which the analysis unit was a large diagnostic medicine laboratory. Using quantitative modeling, the scenarios of the case analyzed were reproduced for application of appointment systems found in the literature. **Results:** in line with the research objectives and the context of the case analyzed, appointment system models with potential for application were selected. Adopting cost-based performance indicators identified in the literature review, the model with the highest potential was evaluated by reference to traditional models for comparative purposes. The results showed superior performance of the proposed model in most scenarios analyzed. **Conclusions:** The flexibility of the model, in order to cover a wide range of scenarios, combined with superior performance and low implementation complexity, make it an interesting alternative for practical application.

Keywords: Health Services, Health Care, Appointment Systems, Diagnostic Medicine

LISTA DE FIGURAS

Figura 1 Regras de Agendamento Encontradas na Literatura (Cayirli e Veral, 2003)	36
Figura 2 Seleção de Modelos da Literatura Compatíveis com os Objetivos do Estudo.	57
Figura 3 Processo de Desenvolvimento de um Estudo de Caso.	60
Figura 4 Etapas Metodológicas	66
Figura 5 Comparativo de Intervalos de Serviço entre as Regras IBFI, Yang e Dome	80
Figura 6 Valores da Constante k e do Parâmetro CR.....	83
Figura 7 Comparativo entre Regras de Agendamento com e sem Ajustes	85
Figura 8 Impactos de Diferentes Valores de Pn e Pw.....	86
Figura 9 Impactos de Diferentes Valores de Cv	87
Figura 10 Variabilidade dos Tempos de Espera para N=40, Cv=0,4 e Pn=0,05. ...	104
Figura 11 Variabilidade dos Tempos de Espera para N=40, Cv=0,8 e Pn=0,4.	105
Figura 12 Geração de Dados Aleatórios para Tempos de Serviço Simulados	114
Figura 13 Geração de Dados Aleatórios para Comparecimentos.....	115
Figura 14 Simulação de Cenário	115
Figura 15 Formulação para os Dados de Entrada	116
Figura 16 Formulação para os Horários de Agendamento (Regra Dome 1).....	116
Figura 17 Formulação para os Horários de Agendamento (Regra IBFI)	117
Figura 18 Formulação para os Horários de Agendamento (Regra 2BEG)	117
Figura 19 Formulação para as Variáveis por Cliente	118
Figura 20 Formulação para os Indicadores das Sessões	118
Figura 21 Formulação para os Indicadores do Cenário	118

LISTA DE QUADROS

Quadro 1 Gestão de Capacidade em Saúde – Classificação (Smith-Daniels et al., 1988)	18
Quadro 2 Fatores Ambientais Relevantes em Agendamento de Serviços (Cayirli e Veral, 2003).....	30
Quadro 3 Indicadores de Desempenho Identificados na Literatura (Cayirli e Veral, 2003).....	34
Quadro 4 Desenvolvimento de um Sistema de Agendamento (Cayirli e Veral, 2003)	41
Quadro 5 Metodologias de Análise (Cayirli e Veral, 2003).....	52
Quadro 6 Principais Características das Pesquisas Analisadas	53
Quadro 7 Sistemas de Agendamento Encontrados na Literatura.	55
Quadro 8 Potencias Modelos da Literatura para Aplicação em Medicina Diagnóstica com Ampla Diversidade de Serviços Agendados.....	56
Quadro 9 Estratégias e Métodos de Pesquisa.	59
Quadro 10 Variáveis dos Modelos.	62
Quadro 11 Entrevistas para Coleta de Dados.	63
Quadro 12 Variáveis do Modelo Selecionado.	77
Quadro 13 Modelagem para Aplicação da Ferramenta Solver.....	106
Quadro 14 Sistemas de Agendamento Encontrados na Literatura	109
Quadro 15 Potencias Modelos da Literatura para Aplicação em Medicina Diagnóstica com Ampla Diversidade de Serviços Agendados.....	110
Quadro 16 Indicadores de Desempenho Identificados na Literatura (Cayirli e Veral, 2003)	111

LISTA DE GRÁFICOS

Gráfico 1 Distribuição de Frequência das Taxas de <i>No-Show</i> para o Grupo de Exames US.....	70
Gráfico 2 Distribuição de Frequência das Taxas de <i>No-Show</i> para o Grupo de Exames COVV1.....	70
Gráfico 3 Distribuição de Frequência das Taxas de <i>No-Show</i> para o Grupo de Exames MM1.....	71
Gráfico 4 Fronteira da Eficiência para $N=10$, $Cv=0,8$ e $Pw=0,05$	101
Gráfico 5 Fronteira da Eficiência para $N=40$, $Cv=0,4$ e $Pw=0,05$	102

LISTA DE TABELAS

Tabela 1 Grupos de Exames de Maior Frequência – 01/01 a 09/04/15	69
Tabela 2 Taxas de <i>No-Show</i> por Grupo de Exames	72
Tabela 3 Valores de N (Tamanho da Sessão).	73
Tabela 4 Preços de Exames.	73
Tabela 5 Parâmetros do Modelo Selecionado.....	82
Tabela 6 Ganhos de Desempenho Decorrentes do Ajuste para <i>No-Shows</i> e <i>Walk-Ins</i>	88
Tabela 7 Parâmetros para Comparativo entre Regras Dome e Tradicionais	89
Tabela 8 Ganhos de Desempenho da Regra Dome 1 em Relação à Melhor Regra Tradicional.....	89
Tabela 9 Ganhos de Desempenho da Regra Dome 2 em Relação à Melhor Regra Tradicional.....	90
Tabela 10 Fatores Ambientais do Caso Estudado	91
Tabela 11 Parâmetros para o Experimento.....	93
Tabela 12 Ganhos de Desempenho Decorrentes do Ajuste para <i>No-Shows</i> e <i>Walk-Ins</i> para N=40	98
Tabela 13 Ganhos de Desempenho da Regra Dome 1 em Relação à Melhor Regra Tradicional.....	99
Tabela 14 Comparativo de Desempenho da Regra 2BEG em Relação à Regra IBFI.	103
Tabela 15 Ganhos de Desempenho da Regra Dome 1 com Aplicação da Ferramenta Solver	106
Tabela 16 Ganhos de Desempenho da Regra Dome 1 em Relação à Melhor Regra Tradicional.....	112

SUMÁRIO

1	INTRODUÇÃO.....	17
1.1	OBJETIVOS DA PESQUISA	20
1.2	RELEVÂNCIA DO TEMA.....	21
1.3	ESTRUTURA DO TRABALHO	22
2	REVISÃO DA LITERATURA.....	23
2.1	FATORES AMBIENTAIS	23
2.1.1	Natureza.....	24
2.1.2	Número de Serviços.....	24
2.1.3	Número de Servidores	24
2.1.4	Número de Atendimentos por Sessão	25
2.1.5	Processo de Chegadas.....	25
2.1.6	Tempo de Serviço	27
2.1.7	Atrasos e Interrupções do Servidor.....	28
2.1.8	Regras de Filas.....	29
2.2	INDICADORES DE DESEMPENHO	29
2.2.1	Indicadores de Custos	30
2.2.2	Indicadores de Tempo	31
2.2.3	Indicadores de Congestionamento.....	32
2.2.4	Indicadores de Equidade.....	32

2.2.5	Outros Indicadores	33
2.3	DESENVOLVIMENTO DE UM SISTEMA DE AGENDAMENTO	33
2.3.1	Regras de Agendamento	33
2.3.2	Classes de Pacientes	38
2.3.3	Ajustes para <i>No-Shows</i> , <i>Walk-Ins</i> e/ou Retornos	39
2.4	MÉTODOS DE PESQUISA EM SISTEMAS DE AGENDAMENTO.....	42
2.4.1	Estudos Analíticos	42
2.4.2	Simulações	45
2.4.3	Estudos de Caso	49
2.5	SELEÇÃO DE POTENCIAIS MODELOS PARA APLICAÇÃO NO CASO.....	53
3	MÉTODO DE PESQUISA	58
3.1	PREMISSAS DO ESTUDO	58
3.2	ETAPAS METODOLÓGICAS	58
3.2.1	Plano de Pesquisa.....	60
3.2.2	Projeto de Pesquisa.....	60
3.2.3	Preparação para Coleta de Dados	61
3.2.4	Coleta de Dados	62
3.2.5	Análise.....	63
3.2.6	Compartilhamento	65
4	RESULTADOS.....	67
4.1	COLETA DE DADOS	67
4.2	TRATAMENTO DE DADOS.....	68

4.3	SELEÇÃO DE MODELOS DE SISTEMAS DE AGENDAMENTO PARA APLICAÇÃO NO CASO ESTUDADO	74
4.4	DESCRIÇÃO DO MODELO SELECIONADO.....	76
4.4.1	Ajuste para <i>No-Shows</i> e <i>Walk-Ins</i>	78
4.4.2	Tamanho da Sessão	78
4.4.3	Regra de Agendamento	78
4.4.4	Indicadores de Desempenho	80
4.4.5	Metodologia do Modelo Seleccionado	81
4.4.6	Resultados do Modelo Seleccionado.....	84
4.5	DEFINIÇÃO DE CENÁRIOS PARA APLICAÇÃO DO MODELO SELECIONADO	91
4.6	APLICAÇÃO DO MODELO SELECIONADO AO CASO	92
4.6.1	Formulação	93
4.6.2	Geração de dados para simulação	96
4.6.3	Obtenção da Constante k para a Regra Dome 1	97
4.6.4	Ajustes para <i>No-Shows</i>	98
4.6.5	Comparativos entre Regras de Agendamento	99
4.7	APLICAÇÃO DO MODELO PROPOSTO PELO AUTOR	105
5	CONCLUSÕES.....	108
	ANEXO 1 - SIMULAÇÕES	114

1 INTRODUÇÃO

A eficiência em serviços de saúde tem sido um tema cada vez mais relevante, na medida em que a demanda aumenta e os recursos tornam-se mais escassos (TAI; WILLIAMS, 2012; GARG et al., 2012, VISSERS; BERTRAND; DE VRIES, 2001). Da mesma forma, os crescentes investimentos e custos de operação são fatores que demandam atenção na gestão neste setor (MCDERMOTT; STOCK, 2007; BOWERS, 2014; ROMERO et al., 2013; MA; DEMEULEMEESTER, 2012).

Além do aumento, Roche et al. (2012) destacam o envelhecimento da população, o qual acarreta um maior número médio de consultas por paciente (RAU et al., 2013).

McDermott e Stock (2007) apontam a pressão de custos como desafio e consideram a gestão de recursos hospitalares uma importante questão política e social.

A forte competição, aliada às limitações de recursos médicos tem levado os administradores a dar atenção a oportunidades de reduções de custo e melhoria de qualidade, com o objetivo de aumentar receitas e melhorar os serviços (JEANG; CHIANG, 2012)

Entre os componentes essenciais de um sistema de saúde, a Organização Mundial da Saúde destaca o acesso a serviços (HONG et al., 2013). A restrição de um recurso em qualquer estágio do serviço de saúde pode levar a queda do nível de serviço (KC; TERWIESCH, 2012). Van Dijk e Kortbeek (2009) e Asaduzzaman et al. (2010) enfatizam a busca por altas taxas de ocupação em recursos de maior custo, o que muitas vezes leva a limitações no acesso aos serviços e conseqüente perda na qualidade do atendimento. Deste modo, as instituições de saúde devem encontrar maneiras criativas para lidar com o *trade-off* entre custos e nível de serviço.

A variabilidade de curto prazo presente em parâmetros como duração de tratamento, sazonalidade e emergências são apontadas por Gupta et al. (2007) como um motivo para listas de espera excessivamente longas para atendimento, mesmo em situações em que a oferta e a demanda se encontram equilibradas.

A complexidade do setor de saúde também é um componente relevante neste cenário (TAI; WILLIAMS, 2012; RAMWADHDOEBE, 2009). Simplesmente conciliar capacidade e demanda não é o suficiente (BOWERS, 2011). É preciso lidar com

incertezas e heterogeneidades (GARG et al., 2012), além da natureza dinâmica inerente a esta indústria (GUPTA et al., 2007).

Dentro deste contexto, a gestão da capacidade em serviços de saúde surge como um elemento importante na busca por eficiência e qualidade.

Pedroso e Malik (2012), ao proporem um modelo de cadeia de valor para o sistema de saúde brasileiro, classificam os serviços de saúde em atividades de promoção de saúde, proteção da saúde e prevenção de doenças, diagnóstico, tratamento, reabilitação, monitoramento e gestão de doenças crônicas. Essas atividades podem ser realizadas em diferentes ambientes, tais como:

- Atendimento ambulatorial (ex. consultas, procedimentos ambulatoriais, exames de medicina diagnóstica);
- Atendimento hospitalar (ex. internações, cirurgias);
- Atendimento residencial (ex.: *home care*);
- atendimentos externos (ex.: programas de promoção de saúde em empresas ou instituições de educação física).

Smith-Daniels et al. (1988) estabeleceram uma segmentação para análise da gestão da capacidade em saúde, cruzando decisões de aquisição e alocação com instalações e mão-de-obra, conforme o Quadro 1.

	Instalações	Recursos Humanos
Decisões de Aquisição	Localização e dimensionamento de capacidade agregada Dimensionamento de unidades hospitalares Dimensionamento de unidades ambulatoriais	Equipes hospitalares Equipes ambulatoriais
Decisões de Alocação	Agendamento de internações Agendamento de unidades cirúrgicas Agendamento ambulatorial	Alocação em dias e turnos Alocação em unidades Alocação em atividades

Quadro 1 – Gestão de Capacidade em Saúde – Classificação (Smith-Daniels et al., 1988)

Dentro das classificações previamente apresentadas, este estudo se enquadra nas decisões de alocação em instalações, considerando, portanto, a utilização de recursos disponíveis (sejam eles existentes ou previamente planejados). Mais precisamente, o

estudo aborda os sistemas de agendamento ambulatorial com foco na aplicação em serviços de medicina diagnóstica.

Conforme Hong et al. (2013), um agendamento eficiente é uma etapa crítica em direção à prestação de serviços com alta qualidade. Trata-se de um problema desafiador por várias razões, incluindo incertezas nos comparecimentos (ex.: atrasos e faltas), variabilidades nos tempos de atendimento, diferentes preferências de pacientes e profissionais e a presença de múltiplos critérios, por vezes conflitantes (ERDOGAN; DENTON, 2013).

Sistemas de agendamento são a conexão entre a eficiência na utilização de recursos e a qualidade no atendimento (CAYIRLI; VERAL, 2003; GUPTA; DENTON, 2008).

Os estudos acerca de sistemas de agendamento de serviços, em sua maioria, segmentam o tipo de paciente como ambulatorial ou internação.

Para Smith-Daniels et al. (1988), atendimentos ambulatoriais e internações têm diferentes especificidades do ponto de vista de agendamento. Enquanto no primeiro existem desafios como altas taxas de *no-shows* e atrasos, no segundo a ênfase é no tempo de internação e no mix de serviços e recursos consumidos pelo paciente.

Gupta e Denton (2008) fazem distinção entre agendamentos ambulatoriais e cirúrgicos, embora também mencione casos de cirurgia em clínicas ambulatoriais. Para os autores, diferentemente de atendimentos ambulatoriais, nos casos de cirurgia os pacientes são admitidos antes da execução do serviço e permanecem após este, consumindo recursos.

Dadas estas diferenças, grande parte da literatura acerca de sistemas de agendamento em serviços de saúde foca em atendimentos ambulatoriais ou internações, havendo pouca incidência, portanto, de estudos abrangendo ambos os casos. Hong et al. (2013), por exemplo, realizaram um estudo sobre uso de simulações direcionadas especificamente para atendimentos ambulatoriais.

Segundo Berg et al. (2014), Hong et al. (2013), Rechel et al. (2010) e Gupta e Denton (2008), além do aumento da demanda há uma tendência de conversão de internações em atendimentos ambulatoriais, atribuindo parte deste fenômeno ao desenvolvimento recente de novas tecnologias e procedimentos cirúrgicos menos invasivos.

Como consequência do aumento da demanda por atendimentos ambulatoriais e relativa falta de capacidade de atendimento desta natureza, Hong et al. (2013) apontam clínicas lotadas, longas filas de espera e profissionais trabalhando em horas extras. Destacam ainda a insatisfação dos clientes e o risco de desmotivação das equipes e perda de eficiência na utilização de recursos, atribuindo estes problemas não somente ao aumento da demanda, mas também à ineficiência de sistemas de agendamento de serviços ambulatoriais.

Esta pesquisa consiste em uma modelagem quantitativa baseada em simulação, na qual será proposto um sistema de agendamento para serviços de saúde ambulatoriais, com foco na aplicação em serviços de medicina diagnóstica. Este sistema deverá ser selecionado a partir de modelos pesquisados na literatura sobre do tema.

1.1 OBJETIVOS DA PESQUISA

Este estudo tem como principal objetivo propor um sistema de agendamento para serviços ambulatoriais, tendo como foco os prestadores de serviços de medicina diagnóstica com diversidade de portfólio de serviços agendados.

Como objetivos secundários, tendo como foco estes prestadores de serviços, a pesquisa busca:

- Analisar a compatibilidade dos modelos de sistema de agendamento encontrados na literatura; e
- Avaliar o desempenho do modelo de maior potencial.

Conforme Green et al. (2006), instalações e equipamentos de medicina diagnóstica são componentes críticos de um sistema de saúde, sendo que tanto seus custos como a expectativa dos clientes têm aumentado rapidamente. Ainda segundo o autor, os gestores destes recursos, muitos de alto valor de aquisição, têm recebido pressão cada vez maior por ganhos de eficiência.

Cayirli e Yang (2014) afirmam que, além de não cobrirem todas as situações existentes, a maior parte dos sistemas de agendamento de serviços ambulatoriais, individualmente, tem desempenho satisfatório em apenas uma faixa restrita de cenários.

Deste modo, sistemas para prestadores de serviços com amplos portfólios de serviços constituem um desafio.

Nesse contexto, a pergunta de pesquisa principal formulada para este estudo é a seguinte:

- Qual modelo de sistema de agendamento é o mais adequado para um prestador de serviços de medicina diagnóstica com ampla variedade de serviços oferecidos?

Esta pergunta de pesquisa pode ser desdobrada nas seguintes questões:

- Quais são os sistemas de agendamento existentes na literatura e direcionados para aplicação em serviços ambulatoriais?
- Quais modelos de sistemas de agendamento encontrados na literatura apresentam maior potencial para aplicação em prestadores de serviços de medicina diagnóstica com ampla diversidade de serviços agendados?
- Quais indicadores de desempenho podem ser utilizados para avaliar a adequação desses modelos?
- Qual o desempenho do modelo de sistema de agendamento de maior potencial, quando aplicado especificamente em um prestador de serviços de medicina diagnóstica com ampla variedade de serviços oferecidos?

1.2 RELEVÂNCIA DO TEMA

Serviços ambulatoriais apresentam crescente relevância à medida que uma parcela significativa de procedimentos médicos deixam de demandar internações (HONG et al., 2013).

Segundo Cayirli e Veral (2003), a área de saúde vem se tornando cada vez mais competitiva. As expectativas dos clientes têm mudado e pesquisas mostram que a escolha de um prestador de serviços de saúde é feita tanto com base na competência técnica quanto na capacidade de se garantir um atendimento adequado. Deste modo, administradores do setor devem evitar sistemas de agendamento que sejam orientados somente para custos.

Gupta e Denton (2008) afirmam que os atuais sistemas de agendamento beneficiam mais as redes de saúde e os profissionais do que os pacientes.

De um modo geral, para Hong et al. (2013), atualmente os principais problemas em clínicas ambulatoriais são os longos tempos de espera e as horas-extras dos profissionais, como consequência de deficiências em sistemas de agendamento, alocação de recursos e o fluxo de pacientes.

Sistemas de agendamento eficientes têm o potencial de aumentar a utilização de recursos ao mesmo tempo em que reduzem o tempo de espera de pacientes (CAYIRLI; VERAL, 2003).

Em contexto nacional, Oleskovicz et al. (2014) mencionam um grande número de exemplos de deficiência de atendimento no setor de saúde.

Falhas de um sistema de agendamento podem incorrer em necessidade de redução dos tempos de atendimento para recuperação de atrasos de agenda. Como consequência, gera-se o risco de erros de diagnóstico, sob a responsabilidade civil do prestador de serviços. Uma abordagem desta questão pode ser encontrada em Oliveira et al. (2011).

Contribui para a relevância do tema, ainda, o espaço existente para maior aplicação de técnicas de administração de operações e métodos quantitativos, mais precisamente pesquisa operacional, na área de saúde. Para Gupta e Denton (2008), este é um dos gargalos na busca de benefícios em eficiência e nível de serviço no setor.

Finalmente, embora existam vários trabalhos teóricos em agendamento de serviços ambulatoriais, os impactos práticos ainda são limitados (CAYIRLI; VERAL, 2003). Espera-se, com este estudo, reduzir esta lacuna.

1.3 ESTRUTURA DO TRABALHO

Esta pesquisa segue a seguinte estrutura: no Capítulo 2 é feita uma revisão da literatura sobre sistemas de agendamento de serviços ambulatoriais. No Capítulo 3 é descrita a metodologia de pesquisa empregada para seleção e simulação de modelos de sistemas de agendamento para o caso analisado. No Capítulo 4 os resultados das simulações são descritos e analisados. Por fim, o Capítulo 5 contém as conclusões, limitações e proposições para contribuições futuras.

2 REVISÃO DA LITERATURA

Os problemas de agendamento, programação e sequenciamento ocorrem em vários contextos operacionais, incluindo transporte, manufatura e serviços (VANDEN BOSCH; DIETZ, 2000), sendo que muitas décadas de pesquisa têm sido dedicadas a este tema (LEUNG, 2004). Entretanto, devido às características únicas de agendamento de serviços de saúde, estas técnicas não podem ser simplesmente replicadas para este contexto.

Cayirli et al. (2012) concluem que, mesmo em se tratando de serviços ambulatoriais, não existem sistemas de agendamento que contemplem todas as circunstâncias existentes, sendo que uma gama de situações ainda não foi suficientemente explorada.

A grande maioria da literatura acerca de agendamento em serviços ambulatoriais está concentrada nas áreas de gestão em saúde, engenharia de produção e pesquisa operacional.

Uma abrangente revisão da literatura sobre o tema foi elaborada por Cayirli e Veral (2003), a qual tem sido largamente utilizada como referência para o estado da arte em pesquisas sobre agendamentos de serviços em saúde. Esta revisão norteará a estrutura deste capítulo, sendo que os estudos nela contemplados serão revisitados, assim como serão acrescentados estudos posteriores, incluindo outras revisões de literatura.

Segundo Cayirli e Veral (2003), o objetivo do agendamento de serviços ambulatoriais é estabelecer um sistema para o qual uma determinada medida de desempenho é otimizada – uma aplicação de planejamento de recursos sob incerteza.

2.1 FATORES AMBIENTAIS

Tanto os serviços quanto os pacientes atendidos por um prestador apresentam características que impactam a configuração de um sistema de agendamento. Ao longo deste trabalho, estas características serão denominadas **fatores ambientais**.

Serviços ambulatoriais podem ser considerados como sistemas de filas, onde os parâmetros devem ser configurados conforme as especificidades de cada cenário. O caso mais simples é o de um sistema onde não há atrasos dos pacientes, com um único profissional para atendimento e tempos de serviço estocásticos (CAYIRLI;

VERAL; 2003). Fatores ambientais mais complexos surgem à medida que são inseridas variáveis como *no-shows*, múltiplos servidores e múltiplos serviços, conforme será abordado nos tópicos seguintes.

2.1.1 Natureza

Os casos de agendamento são divididos em duas naturezas: estáticos e dinâmicos. Wang (1993) define o agendamento estático como aquele em que um número finito de pacientes deve ser agendado, assumindo que ainda não haja nenhum agendamento no sistema, enquanto que no agendamento dinâmico os pacientes são agendados um a um à medida que novas demandas surgem. Já para Cayirli e Veral (2003), o agendamento estático é aquele definido antes do início da sessão de atendimento, enquanto que no agendamento dinâmico alterações podem ocorrer durante a sessão. A maioria dos estudos contempla agendamento estático. Casos de agendamento dinâmico podem ser encontrados em Fries e Marathe (1981), Liao et al. (1993), Liu e Liu (1998b), Green et al. (2006), Cayirli et al. (2012) e Erdogan e Denton (2013). Em seu trabalho, Wang (1993) contempla ambos os casos.

2.1.2 Número de Serviços

Grande parte dos autores modela um sistema de um estágio, onde os pacientes agendam um único serviço. Nos poucos estudos com simulações em cenários multiestágio (RISING et al.1973; COX et al. 1985; SWISHER et al.2001, TSAI; TENG, 2014), as probabilidades associadas a cada estágio demandam a aplicação do Processo de Markov. Neste processo o comportamento futuro de uma variável independe do seu comportamento passado.

2.1.3 Número de Servidores

Quanto ao número de servidores, na literatura há predomínio de estudos com um único servidor (RISING et al., 1973; COX et al., 1985; HASSIN; MENDEL, 2008). Neste caso, segundo Cayirli e Veral (2003) a motivação reside no fato de os pacientes terem preferência quanto a atendimento por um determinado médico, quando se trata de consultas particulares. Já para o caso de consultas em clínicas públicas normalmente não há a possibilidade de se optar por um profissional, o que configura um ambiente multiservidor, conforme pode ser verificado nos trabalhos de Babes e Sarma (1991) e Liu e Liu (1998a, 1998b).

2.1.4 Número de Atendimentos por Sessão

Os trabalhos de Vissers (1979), Heaney et al. (1991) e Meza (1998) apontam para uma correlação positiva entre os tempos de espera e o número de consultas em uma sessão de atendimento (N). Welch e Bailey (1952), Vissers e Wijngaard (1979) e Ho e Lau (1992) enfatizam a necessidade de se contemplar este parâmetro em comparativos de desempenho de regras de agendamento. Além disto, Ho e Lau (1992) afirmam que o efeito do número de consultas por sessão pode ser mitigado pela presença de *no-shows* e a variabilidade dos tempos de serviço e, portanto, não pode ser generalizado. Em modelos como os de Yang et al. (1998), Cayirli et al. (2012) e Cayirli e Yang (2014), os intervalos entre agendamentos são definidos em função de N.

2.1.5 Processo de Chegadas

Segundo Cayirli e Veral (2003), um conjunto de fatores compõe as características do processo de chegadas de pacientes e estão relacionados ao desempenho do sistema de agendamento. O primeiro fator é o atraso do paciente, o qual pode ser definido como a diferença entre o horário agendado e o horário real da sua chegada. Evidências empíricas sugerem que a frequência de pacientes que chegam antes do horário agendado é maior do que a de atrasos (FETTER; THOMPSON, 1966; VILLEGAS, 1967; BABES; SARMA, 1991; O'KEEFE, 1985; BRAHIMI; WORTHINGTON, 1991b; KLASSEN; ROHLER, 1996; LEHANEY et al., 1999). Entretanto, Welch e Bailey (1952) destacam que chegadas antecipadas nem sempre são desejáveis, uma vez que aumentam a ocupação das salas de espera. Algumas pesquisas utilizam distribuições de probabilidades teóricas para representar a pontualidade dos pacientes (BLANCO WHITE; PIKE, 1964; FETTER; THOMPSON, 1966; SWARTZMAN, 1970; COX et al., 1985; CAYIRLI et al., 2006). Com objetivo de reduzir o número de variáveis no desenvolvimento de sistemas de agendamento, Vissers e Wijngaard (1979) propõem a junção das pontualidades do paciente e do prestador de serviços, compondo uma única variável. Os modelos de filas desenhados por Mercer (1960, 1973) estabelecem um limite para o atraso do paciente. Em todos estes estudos o atraso dos pacientes é tratado como uma variável independente dos horários de seus agendamentos.

Outro fator relevante é o absenteísmo dos pacientes. Em sua revisão de literatura, Cayirli e Veral (2003) consideram como moderada a inclusão de *no-shows* em pesquisas, com probabilidades (p) variando entre 5 e 30%. Em um estudo mais recente sobre desafios e oportunidades em agendamentos, Gupta e Denton (2008) destacam este fenômeno ainda como um ponto a ser melhor explorado. Considerando *no-shows*, variabilidade do tempo de serviço e número de pacientes por sessão como fatores ambientais, Ho e Lau (1992) avaliam o primeiro como o de maior impacto na seleção e desempenho de sistemas de agendamento. Diversos estudos investigam as causas de *no-shows* na tentativa de correlacioná-lo com possíveis variáveis, como idade e nível socioeconômico, e assim definir políticas para reduzi-lo (DEYO; INUI, 1980; BARRON, 1980; SCHAFFER, 1986). Oleskovicz et al. (2014) destacam que, mesmo com a implantação de ações para mitigação do *no-show*, este dificilmente será eliminado. Deste modo, a redução da incidência de *no-shows* e a consideração desta variável em sistemas de agendamento são medidas complementares. Vale ressaltar que cancelamentos de serviços com antecedência insuficiente têm o mesmo efeito do *no-show*, sendo que a literatura usualmente trata estes parâmetros como uma única variável (GUPTA; DENTON, 2008). Vissers e Wijngaard (1979) propõem uma abordagem de aproximação, tratando *no-shows* como serviços de duração igual a zero. Exemplos de modelos de otimização e simulação contemplando *no-shows* podem ser encontrados em Kaandorp e Koole (2007), Hassin e Mendel (2008), Berg et al. (2014), Tsai e Teng (2014) e Vuyst et al. (2014).

Pacientes não-agendados também são um fator de impacto no processo de chegada de pacientes. De acordo com Cayirli e Veral (2003), grande parte das pesquisas sobre agendamentos em serviços ambulatoriais negligenciam a presença destes casos, caracterizados como emergências, urgências ou simplesmente pacientes que se dirigem a uma instituição de saúde sem ter feito um agendamento. Alguns estudos analisam as distribuições de probabilidades de pacientes não-agendados (Swartzman, 1970; Rising et al., 1973; Swisher et al., 2001). O estudo de Vissers (1979) converte o efeito dos pacientes não-agendados em variação do tempo de serviço. Cayirli et al. (2006), Cayirli et al. (2012) e Tsai e Teng (2004) são exemplos de estudos que contemplam pacientes não-agendados.

2.1.6 Tempo de Serviço

Bailey (1952) define o tempo de serviço como a soma de todos os tempos em que o paciente está demandando atenção de um médico, impedindo este de atender outros pacientes. Este tempo poder ser um valor estocástico, com uma distribuição estatística, ou um valor determinístico (fixo).

Conforme Cayirli e Veral (2003), a maioria dos estudos considera pacientes com uma distribuição de tempo de serviço homogênea e independente, enquanto que algumas pesquisas agrupam classes de pacientes com distribuições distintas, também independentes. A modelagem de problemas de agendamento com servidor estocástico é amplamente explorada na literatura (ERDOGAN; DENTON, 2013).

Em relação à variabilidade do tempo de serviço associada ao prestador de serviços, Cayirli e Veral (2003) consideram questionável a premissa de independência em relação ao horário de agendamento, destacando que vários estudos observam maior velocidade no atendimento em horários de pico devido a congestionamentos em salas de espera.

Diversos tipos de distribuição do tempo de serviço são contemplados nos trabalhos pesquisados. Alguns utilizam dados empíricos coletados em clínicas, observando distribuições unimodais e assimétricas (WELCH; BAILEY, 1952; JACKSON 1964; RISING et al. 1973; BUCHAN; RICHARDSON, 1973; COX et al. 1985; BRAHIMI; WORTHINGTON, 1991b; MEZA 1998). A maior parte dos estudos analíticos adotam distribuições de Erlang ou exponenciais para viabilizar seus modelos.

O coeficiente de variação do tempo de serviço é uma métrica comumente utilizada para análise de variabilidade (BAILEY, 1952; BLANCO WHITE; PIKE, 1964; RISING et al. 1973; O'KEEFE, 1985; BRAHIMI; WORTHINGTON, 1991b; MEZA 1998).

Os modelos de otimização de Denton e Gupta (2003) mostram influência da assimetria na distribuição de probabilidades. Por outro lado, alguns autores afirmam que o desempenho de sistemas de agendamento não é afetado pela assimetria ou curtose, mas somente pela média e a variância (HO; LAU, 1992; YANG et al., 1998). No modelo de Robinson e Chen (2001), a média dos tempos de serviço é removida sem influência no problema estudado.

Tanto o tempo de espera dos pacientes como a ociosidade dos prestadores de serviços são negativamente afetadas pela alta variabilidade dos tempos de serviço (BAILEY 1952; BLANCO WHITE; PIKE, 1964; VISSERS; WIJNGAARD, 1979; HO; LAU, 1992; KLASSEN; ROHLEDER, 1996; DENTON; GUPTA, 2003; CAYIRLI et al, 2012). De maneira análoga, Wang (1997) conclui que quanto maior o coeficiente de variação, menor o intervalo ótimo entre serviços e maior o custo criado pela incerteza no sistema.

De um modo geral, estudos que avaliam o efeito da duração do tempo de serviço apontam para uma redução do tempo de espera de pacientes para tempos de serviço menores (BAILEY, 1952; BLANCO WHITE; PIKE, 1964; WALTER, 1973).

2.1.7 Atrasos e Interrupções do Servidor

Blanco White e Pike (1964), Fetter e Thompson (1966), Vissers (1979), Mahachek e Knabe (1984), Babes e Sarma (1991), e Liu e Liu (1998a, 1998b) contemplam em seus estudos o atraso de médicos, definido como sendo o atraso na primeira consulta de uma sessão.

É consenso entre as pesquisas que o tempo de espera dos pacientes é significativamente sensível a este fator. Caso o primeiro serviço não seja iniciado pontualmente, constata-se um acúmulo de atrasos que se estende por toda a sessão de atendimento.

Outro fator relacionado ao servidor é a interrupção no atendimento, a qual, no caso de médicos, por exemplo, inclui todas as atividades durante uma sessão que podem requerer a atenção do profissional, como interações com a equipe e telefonemas.

Rising et al. (1973) e Lehaney et al. (1999) incluem interrupções em seus modelos, assumindo que estas ocorrem somente entre serviços.

Cayirli e Veral (2003) sugerem a aplicação da Teoria dos Jogos para a modelagem da pontualidade de médicos e pacientes, levando em conta o conflito de interesses de ambas as partes. O autor considera provável que pacientes cheguem mais tarde, considerando que médicos atrasam, assim como médicos possam chegar mais tarde, por conta de eventuais atrasos de pacientes.

2.1.8 Regras de Filas

Praticamente todos os estudos pesquisados assumem o atendimento de pacientes por ordem de chegada (regra FCFS - *first-come, first-served*). Quando a premissa é de que os pacientes são pontuais esta regra significa que o atendimento se dará na ordem dos serviços agendados. Entretanto, atrasos podem implicar alterações na sequência de atendimento, uma vez que o prestador de serviços provavelmente não ficará ocioso se houver outros pacientes em espera (CAYIRLI; VERAL, 2003).

Em um cenário em que existam pacientes não-agendados e atendimentos de retorno é necessário o estabelecimento de regras de prioridade, as quais determinam a sequência em que os pacientes devem ser atendidos (CAYIRLI; VERAL, 2003). Em geral, a prioridade é dada a emergências, seguidas de consultas de retorno e atendimentos agendados. A última prioridade é dada a demais pacientes não agendados, seguindo a regra FCFS (RISING et al. 1973; COX et al. 1985). Na prática não são raros os casos em que todos os pacientes são atendidos com base em FCFS mesmo com a existência de um sistema de agendamento, com o objetivo de simplificar o trabalho administrativo. No entanto, isto prejudica o todo o propósito do sistema e pode levar os pacientes a desprezar os horários agendados e comparecer para o serviço mais cedo do que o necessário.

O Quadro 2 resume os principais fatores ambientais em agendamento de serviços de saúde.

2.2 INDICADORES DE DESEMPENHO

Existe uma gama de critérios utilizada na literatura para avaliar sistemas de agendamento. Os resultados de pesquisa são medidos frequentemente em termos de tempo médio de espera de pacientes $E(W)$, tempo médio de ociosidade do servidor $E(I)$ e/ou tempo médio de horas-extras do servidor $E(O)$, cabendo a um decisor estabelecer um equilíbrio subjetivo entre estes indicadores (CAYIRLI; VERAL, 2003).

Ainda para Cayirli e Veral (2003), uma alternativa consiste em atribuir pesos aos custos relacionados a espera do paciente (C_p), ociosidade e horas-extras do servidor (C_d , C_o). Deste modo, o objetivo seria atingir um custo total esperado mínimo, representado na Equação 1.

$$\text{Mín } E(TC) = E(W)C_p + E(I)C_d + E(O)C_o \quad (1)$$

Quadro 2 – Fatores ambientais relevantes em agendamento de serviços (Cayirli e Veral, 2003)

Natureza
Estática
Dinâmica
Modelagem de Ambientes da Sessão
Número de serviços (estágio único ou multiestágio)
Número de servidores (servidor único ou multiservidor)
Número de atendimentos por sessão
Processo de chegada (determinístico ou estocástico)
Pontualidade dos pacientes
Existência de <i>no-shows</i>
Existência de <i>walk-ins</i>
Presença de acompanhantes
Tempos de serviço (distribuição empírica ou teórica)
Atrasos e interrupções do servidor
Regras de filas (FCFS, por horário agendado, por prioridade)

2.2.1 Indicadores de Custos

As pesquisas utilizam diferentes fórmulas ou variações da função de custos da Equação 1. A maior parte contempla apenas o tempo de espera do paciente e o tempo ocioso e de horas-extras do servidor. Outros utilizam o tempo de fluxo do paciente em vez do seu tempo de espera.

A função geral de custos tem como premissa a linearidade entre o custo e o tempo de espera do paciente. Entretanto, como ilustram Klassen e Rohleder (1996), não seria plausível afirmar que um sistema onde um paciente aguarda 40 minutos para um atendimento tem o mesmo desempenho de outro para o qual 20 pacientes aguardam 2 minutos cada para que sejam atendidos. Há ainda um aumento na complexidade se considerarmos que os custos relativos podem mudar de um paciente para outro. Na literatura, os trabalhos assumem custos idênticos para todos os pacientes. Para Cayirli e Veral (2003), quando da modelagem com pacientes não-pontuais e não-agendados, a premissa de custos de espera homogêneos deve ser desconsiderada e estes casos devem ter uma tolerância maior para atrasos. Para os casos usuais (pacientes

agendados e pontuais), deveria haver um limite de tempo de espera a partir do qual a tolerância do paciente diminui acentuadamente. Alguns estudos indicam que esta tolerância diminui após cerca de 30 minutos de espera (WESTMAN et al., 1987; HUANG 1994).

Conforme Cayirli e Veral (2003), em vez dos valores monetários seria suficiente, do ponto de vista de tomada de decisão, considerar valores relativos para custos, como, por exemplo, a razão entre custos de espera do prestador de serviços e do paciente (Cd/Cp). Esta razão varia entre 1 e 100 nos estudos analisados.

Conforme destacam Fries e Marathe (1981), é mais simples estimar os custos relativos ao servidor, uma vez que para os pacientes há dificuldade em se atribuir um valor ao seu tempo, o que envolve fatores como boa vontade e “custos para a sociedade”.

Keller e Laughhunn (1973) dividem o salário anual do médico pelas horas trabalhadas para estimar o Cd e consideram o salário mínimo para quantificar o custo de oportunidade do tempo de espera do paciente. Em geral assume-se que $Cd > Cp$, uma vez que o Cd abrange não somente o custo do médico, mas também das instalações (YANG et al., 1998; CAYIRLI et al., 2012).

2.2.2 Indicadores de Tempo

Usualmente é desejável que os tempos de espera, ociosidade e horas-extras sejam medidos separadamente, uma vez que isto permite estabelecer limites para cada um destes componentes. Uma abordagem comum é o cálculo do tempo “real” de espera do paciente, subtraindo-se o maior valor entre a hora agendada e a hora da chegada do horário efetivo de início do serviço, uma vez que uma chegada antecipada voluntária por parte do paciente não pode ser computada como tempo de espera atribuído ao sistema de agendamento (CAYIRLI; VERAL, 2003).

A maioria dos estudos utiliza os tempos para cálculos de medidas de custos. Kaandorp e Koole (2007) é um dos poucos trabalhos que apresenta um modelo em que a função objetivo é expressa em termos de unidades de tempo.

O tempo real de espera pode ser negativo, caso o paciente seja atendido antes do horário agendado, o que pode auxiliar o decisor a capturar informação sobre benefícios percebidos pelos pacientes em função da antecipação no atendimento.

Entretanto, caso o foco seja somente em valores positivos, os valores devem ser truncados em zero (CAYIRLI; VERAL, 2003).

O tempo de fluxo de pacientes é outra medida baseada em tempo, sendo definido como o tempo total em que o paciente permanece no prestador, incluindo o tempo de serviço. Uma vez que os pacientes geralmente não consideram relevante este componente, a maior parte da literatura foca o tempo de espera em vez do tempo de fluxo.

Segundo Cayirli e Veral (2003), o tempo ocioso do prestador de serviços é o tempo total que este não é utilizado durante uma sessão, devido à ausência de pacientes para serem atendidos.

Horas-extras são computadas com base na diferença positiva entre o horário desejado de término de uma sessão e o horário real do final do atendimento do último paciente. Em geral, valores negativos de horas-extras são considerados parte do tempo ocioso, embora possam ser considerados como tempo para execução de tarefas administrativas (CAYIRLI; VERAL, 2003).

Em seu estudo, Vissers e Wijngaard (1979) propõem uma aproximação para agrupar os tempos de espera de pacientes e de servidores em uma única variável.

2.2.3 Indicadores de Congestionamento

O congestionamento em um prestador prejudica a qualidade dos serviços sob várias perspectivas. Além de aumentar a ocupação de espaço valioso com longas filas, profissionais podem ser levados a aumentar sua velocidade de atendimento ou serem forçados a adiar atendimentos. A principal medida de congestionamento é o número médio de pacientes no sistema ou fila (CAYIRLI; VERAL, 2003).

2.2.4 Indicadores de Equidade

Alguns estudos dão atenção à questão da “justiça”, a qual é representada pela uniformidade do desempenho de um sistema de agendamento entre pacientes.

Em sistemas de agendamento com intervalos fixos, os sucessivos pacientes tendem a ter, em média, um tempo de espera crescente em função do congestionamento que se acumula ao longo de uma sessão. Não somente os tempos de espera aumentam, mas também os tempos de serviço tendem a redução na medida em que o prestador

procura recuperar atrasos (HEANEY et al., 1991). Assim sendo, pacientes atendidos ao final de uma sessão tendem a experimentar a combinação menos favorável de tempo de espera e tempo de serviço, a menos que algum ajuste no sistema de agendamento seja feito para considerar este fenômeno.

Bailey (1952) mede o tempo médio de espera de pacientes de acordo com a sua sequência em uma sessão, enquanto que Yang et al. (1998) medem a uniformidade dos tempos de espera e Cox et al. (1985) comparam sistemas de agendamento baseados na variância do tamanho da fila ao longo da sessão de atendimento.

2.2.5 Outros Indicadores

Demais indicadores para avaliar o desempenho de sistemas de agendamento incluem produtividade do servidor (ou seja, o número de pacientes atendidos em uma sessão), utilização média do servidor, tempo entre solicitações e agendamento de serviços, percentual de pacientes atendidos em regime de urgência e probabilidade de pacientes serem agendados nos horários desejados.

Swisher et al. (2001) empregam uma medida chamada “efetividade”, a qual compreende tanto os lucros de uma clínica quanto o tempo de espera do paciente em valores monetários.

O Quadro 3 classifica os indicadores de desempenho mais comumente encontrados na literatura.

2.3 DESENVOLVIMENTO DE UM SISTEMA DE AGENDAMENTO

Conforme Cayirli e Veral (2003), o desenvolvimento de um sistema de agendamento pode ser dividido em uma série de decisões quanto a: (1) regras de agendamento, (2) uso de classificação de pacientes (se houver), e (3) ajustes para mitigar os efeitos negativos de *no-shows* e atendimentos não-agendados.

2.3.1 Regras de Agendamento

As sessões de atendimento são normalmente divididas em blocos, sendo que cada bloco equivale a um horário agendado. Uma regra para agendamento de pacientes é composta de três variáveis (CAYIRLI; VERAL, 2003):

- Dimensão do bloco (ni) é o número de pacientes agendados no i ésimo bloco;

- Bloco inicial ($n1$) é o número de pacientes agendados para um mesmo horário no início de uma sessão (1º bloco);
- Intervalo entre agendamentos (a_i) é o intervalo entre dois agendamentos consecutivos. Intervalos entre agendamentos podem ser constantes ou variáveis. Uma prática comum é a adoção de intervalos em função de alguma média (e por vezes do desvio padrão) dos tempos de consulta.

Quadro 3 – Indicadores de desempenho identificados na literatura (Cayirli e Veral, 2003)

Indicadores Baseados em Custos
Custo médio total calculado com combinações de:
Tempo de espera de pacientes
Tempo de fluxo de pacientes
Tempo ocioso do provedor
Horas-extras do provedor
Indicadores Baseados em Tempo
Média, máximo e distribuição de frequência do tempo de espera de pacientes
Média, variância e distribuição de frequência do tempo ocioso do provedor
Média, máximo e desvio padrão de horas-extras do provedor
Média e distribuição de frequência do tempo de fluxo de pacientes
Percentual de pacientes atendidos em até 30 min após o horário agendado
Indicadores de Congestionamento
Média e distribuição de frequência do número de pacientes na fila
Média e distribuição de frequência do número de pacientes no sistema
Indicadores de Equidade
Tempo médio de espera de pacientes conforme o estágio da sessão
Variância do tempo de espera de pacientes
Variância do tamanho da fila
Outros Indicadores
Produtividade do provedor
Ocupação média do provedor
Tempo entre a solicitação do serviço e o agendamento
Percentual de pacientes atendidos em urgência
Probabilidade de os pacientes serem agendados nos horários requisitados
Eficiência da sessão

Qualquer combinação destas três variáveis (n_i , $n1$ e a_i) forma uma possível regra de atendimento.

A Figura 1 traz as regras de agendamento mais comumente investigadas na literatura.

1. Bloco simples: esta regra consiste em agendar todos os pacientes para que cheguem como um bloco no início da sessão, sendo atendidos com base em FCFS. Esta é a forma mais elementar de sistema de agendamento, onde apenas uma data é agendada, sem definição de horário. Obviamente esta regra implica tempos de espera excessivos para pacientes, enquanto que garante a ocupação do servidor. Esta era a prática comum nos anos 1950, quando tiveram início as pesquisas em agendamento de serviços ambulatoriais. Deste modo, observa-se que estes primeiros estudos disseminam as vantagens do agendamento individual, sendo os pioneiros na transição de sistemas de bloco simples para sistemas de bloco individual (LINDLEY, 1952; BAILEY, 1952; WELCH, 1964; FRY, 1964; JOHNSON; ROSENFELD, 1968; ROCKART; HOFMANN, 1969). Sistemas de bloco simples ainda são utilizados, principalmente em clínicas públicas, provavelmente em função da simplicidade administrativa.

2. Bloco individual/intervalo fixo: nesta regra os pacientes são agendados em horários específicos, em intervalos fixos ao longo da sessão de atendimento. Fetter e Thompson (1966); Klassen e Rohleder (1996); Rohleder e Klassen (2000) e Anderson et al. (2013) realizaram estudos com este tipo de regra.

3. Bloco individual/intervalo fixo com um bloco inicial: trata-se de uma combinação da regra anterior com um grupo inicial de n_1 pacientes ($n_1 > 1$) agendados para o início de uma sessão. O objetivo desta regra é manter um “estoque” de pacientes de modo que o servidor não fique ocioso caso o primeiro paciente atrase ou não compareça. Bailey (1952) foi o primeiro pesquisador a sugerir um sistema de bloco individual com dois pacientes agendados no início da sessão e o restante dos agendamentos com intervalos fixos baseados em tempos médios de consulta ($n_1 = 2$, $n_i = 1$, $a_i = \mu$). Os estudos comparativos de Jansson (1966), Blanco White e Pike (1964), Brahimí e Worthington (1991b), Ho e Lau (1992) e Klassen e Rohleder (1996) consideram esta regra.

4. Bloco múltiplo/intervalo fixo: nesta regra grupos de m pacientes são agendados em cada horário com intervalos constantes. Soriano (1966) apresenta um estudo com um sistema de agendamento em que os pacientes são agendados em pares com intervalos equivalentes a duas vezes o tempo médio de consulta ($n_i = 2$, $a_i = 2\mu$). Para os cenários pesquisados nos trabalhos de Blanco White e Pike (1964) e Cox et al. (1985) os blocos múltiplos apresentaram melhor desempenho. Cayirli e Veral (2003)

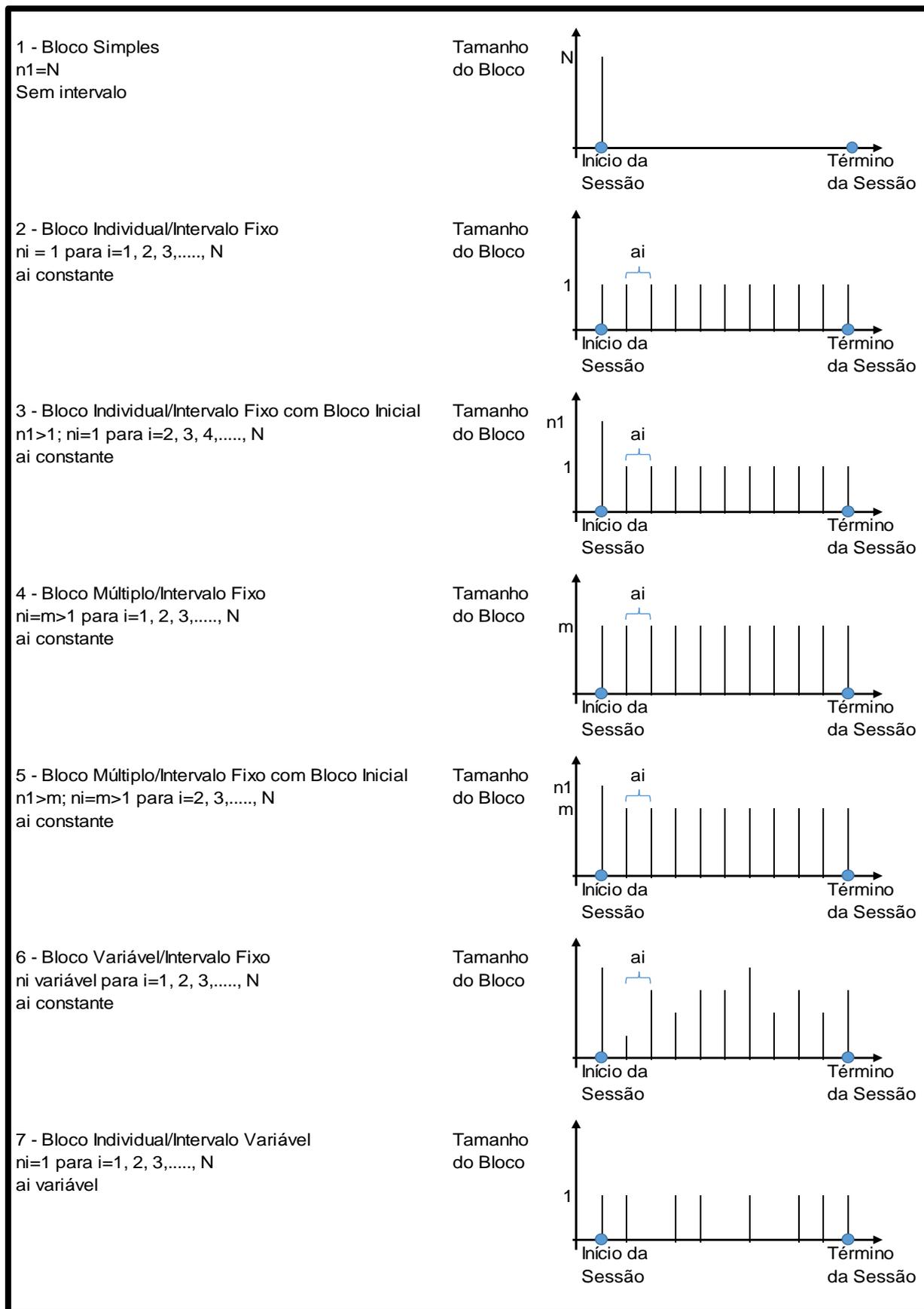


Figura 1 - Regras de agendamento encontradas na literatura (Cayirli e Veral, 2003)

ressaltam a necessidade de estudos mais rigorosos para que se identifique em que circunstâncias os blocos múltiplos são vantajosos em relação às regras de blocos individuais. É possível que agendamentos em bloco sejam mais adequados quando os tempos médios de serviço são menores, de modo que os pacientes agendados para um mesmo bloco não tenham tempos de espera excessivos. Existe ainda a vantagem prática de se agendar horários “arredondados” como, por exemplo, agendar quatro pacientes a cada 15 minutos em vez de um paciente a cada 3,75 minutos (WALTER, 1973).

5. Bloco múltiplo/intervalo fixo com um bloco inicial: trata-se de simplesmente de uma variação da regra anterior, com um bloco inicial ($n1 > m$). Cox et al. (1985) é o único estudo a investigar este tipo particular de regra.

6. Bloco variável/intervalo fixo: esta regra permite que os blocos de uma mesma sessão possuam tamanhos diferentes, enquanto que os intervalos entre horários permanecem constantes. Esta regra é objeto das pesquisas de Villegas (1967), Rising et al. (1973), Fries e Marathe (1981), Liao et al. (1993), Liu e Liu (1998a, 1998b) e Vanden Bosch et al. (1999).

7. Bloco individual/intervalo variável: nesta regra os pacientes são agendados individualmente, porém os intervalos entre horários podem variar. As simulações de Ho e Lau (1992) introduzem uma série de regras para variação dos intervalos, comparando os resultados com regras usuais. Foi constatado, entre as regras testadas, que intervalos crescentes durante a sessão de atendimento representam a regra de melhor desempenho. Alguns estudos analíticos demonstram que, para distribuições idênticas e independentes de tempo de serviço e custos uniformes para todos os pacientes, intervalos ótimos entre serviços exibem um padrão em que estes crescem entre o início e o meio de uma sessão e então decrescem em direção ao seu final. Este comportamento é conhecido como “formato de domo” e é contemplado nos estudos de Wang (1997), Robinson e Chen (2001), Denton e Gupta (2003), Cayirli et al. (2012) e Cayirli e Yang (2014). Pesquisas com blocos individuais e intervalos variáveis também podem ser encontradas em Pegden e Rosenshine (1990), Yang et al. (1998) e Vanden Bosch e Dietz (2000).

Existem ainda variações da regra 7, como as pesquisadas por Ho e Lau (1992, 1999), as quais podem ser consideradas como intermediárias em relação às regras 2 e 7.

Nestes estudos, parte da sessão tem um determinado intervalo entre agendamentos fixado, sendo que para o restante adota-se um intervalo diferente, também fixo.

2.3.2 Classes de Pacientes

Na maioria dos estudos os pacientes são considerados homogêneos e agendados por ordem de solicitação (regra FCFA - *first-call, first-appointment*). Entretanto, quando atributos como tempo de serviço, padrões de chegada e custos de espera permitem agrupá-los em classes, levanta-se a possibilidade de aprimorar o sistema de agendamento com o uso destas características.

Segundo Cayirli e Veral (2003), a classificação de pacientes em agendamentos ambulatoriais pode ser utilizada com dois objetivos: sequenciar pacientes no momento do agendamento e/ou ajustar os intervalos entre agendamentos com base nas diferentes características de tempo de serviço das diferentes classes de pacientes.

Como a agenda precisa ser feita de modo antecipado e as solicitações de agendamentos são dinâmicas, o uso da classificação de pacientes é um tanto limitada. Uma aplicação realista da classificação demanda que esta seja feita com um número de grupos gerenciável e que estes sejam alocados em janelas de tempo pré-determinadas ao solicitar um agendamento.

Na literatura, são encontradas algumas classificações para fins de agendamento com base em primeira consulta/retorno, variabilidade do tempo de serviço (baixa/alta) e tipo de procedimento. Estes fatores estão presentes nos trabalhos de Cox et al. (1985), Klassen e Rohleder (1996), Rohleder e Klassen (2000), Lehaney et al. (1999), Lau e Lau (2000) e Vanden Bosch e Dietz (2000). Tsai e Teng (2014) utilizam a classificação de pacientes quanto à probabilidade de *no-show*. Cayirli e Yang (2014) classificam os pacientes de acordo com a média e o desvio-padrão dos tempos de atendimento, a probabilidades de *no-shows* e a probabilidade de atendimentos não-agendados (*walk-ins*)

Em seu estudo em um departamento de radiologia, Walter (1973) investiga a possibilidade de aprimoramento de um sistema de agendamento mediante o agrupamento de pacientes com tempos de exame semelhantes em sessões específicas. Constatou-se que as durações de exames dependem de fatores como idade do paciente, mobilidade física e tipo de procedimento. Existe ainda literatura

relevante sobre agendamento de cirurgias, a qual contempla pacientes heterogêneos quanto ao uso de instalações. Nestes casos, pode-se estimar as durações para cada procedimento em particular. Diferentemente de agendamento ambulatorial, existe uma lista completa de todas as solicitações feitas para um determinado dia, sendo que há garantia de disponibilidade de pacientes.

Quando da avaliação de sistemas de agendamento com a aplicação de classes de pacientes, várias questões precisam ser consideradas: primeiramente, estes sistemas são menos flexíveis do que os que alocam pacientes pela regra FCFS, uma vez que limitam os horários disponibilizados. Por exemplo, uma clínica pode restringir agendamentos para novos pacientes antes das 10:00 hs e alocar consultas de retorno entre 10:00 e 12:00 hs. Neste caso, é perfeitamente possível que um paciente de retorno insista pelo horário das 9:00 hs. Além disso, mesmo que a proporção entre novas consultas e retornos seja conhecida no longo prazo, as proporções diárias podem variar, trazendo dificuldades para a agenda em determinados dias. Quando isto acontece, existem duas opções: alocar um paciente em outro horário, como exceção à regra, ou adiar o serviço. Rohleder e Klassen (2000) endereçam estas questões mediante o uso de medidas secundárias de desempenho, como a probabilidade de o paciente ser alocado no horário de preferência. Vanden Bosch e Dietz (2000) avaliam um sistema de agendamento com base no número de horários não ocupados, quando se busca minimizar o tempo entre a solicitação e a data do serviço.

2.3.3 Ajustes para *No-Shows*, *Walk-Ins* e/ou Retornos

Sempre que relevantes, *no-shows*, pacientes não-agendados (*walk-ins*) e retornos devem ser considerados na elaboração de um sistema de agendamento. Em clínicas onde as consultas de retorno ocorrem com frequência, como ortopedia, alguma reserva deve ser planejada para estas demandas adicionais (CAYIRLI; VERAL, 2003).

Embora muitos mecanismos administrativos sejam empregados para reduzir a probabilidade de *no-shows*, como *e-mails*, telefonemas e multas, não é possível eliminá-los (BARRON, 1980). Por outro lado, foram encontradas fortes correlações entre a tendência a procurar serviços sem agendamento prévio e classes sociais mais baixas e a percepção de urgência por Taylor (1984) e Virji (1990). Estes resultados sugerem que um prestador de serviços que nega acesso a pacientes não-agendados

pode desfavorecer estes grupos. De maneira geral, portanto, uma abordagem mais adequada seria antecipar-se aos *no-shows* e pacientes não-agendados e ajustar o sistema de agendamento para reduzir seus efeitos negativos.

Blanco White e Pike (1964) consideram apenas ajustes para *no-shows*, utilizando simulações para analisar os impactos em adicionar pacientes extras para compensar o número médio de *no-shows*, concluindo que estes ajustes podem melhorar consideravelmente o desempenho do sistema.

Fetter e Thompson (1966) apontam os riscos em se assumir que *no-shows* e serviços não-agendados se compensam entre si, dado que raramente estes ocorrem na mesma proporção em uma mesma sessão. Deste modo, os autores recomendam que a ocupação da sessão leve em conta as expectativas de *no-shows* e pacientes não-agendados separadamente. Vissers (1979) simula duas opções para lidar com *no-shows*: adicionar pacientes extras igualmente distribuídos ao longo da sessão de atendimento (*overbooking*); e reduzir proporcionalmente os intervalos entre consultas, concluindo que a segunda abordagem é ligeiramente mais eficiente.

Pierskalla e Brailer (1994) sugerem que um sistema de agendamento que considera a variação estocástica de pacientes não-agendados separadamente da variação de *no-shows* atinge melhores resultados.

Para os casos não-agendados, os ajustes são tanto manter horários livres como programar intervalos de serviços mais longos. Na primeira opção há uma decisão secundária que é a definição de quais horários serão deixados em aberto. Em seu estudo de caso, Rising et al. (1973) relatam que, na presença de *walk-ins*, reservar horários livres pode suavizar o fluxo de pacientes.

Fetter e Thompson (1966) incluem pausas para emergências no sistema de agendamento. Na prática, quando o médico é direcionado para uma emergência, clínicas particulares usualmente contatam os pacientes agendados para adiamento da consulta (SCHAFER, 1986).

Klassen e Rohleder (1996) buscaram identificar o melhor posicionamento de horários livres para casos de urgência ao longo de uma sessão, sem, no entanto, atingir resultados conclusivos: quando os horários livres são concentrados no início da sessão, o tempo médio de espera dos pacientes é menor e um número menor de

urgências é atendida; já a reserva de horários concentrada ao final da sessão, a ociosidade do servidor é reduzida, com um atendimento maior de casos de urgências.

Cayirli et al. (2012) e Cayirli e Yang (2014) consideram tanto *no-shows* quanto *walk-ins* em seus modelos.

O Quadro 4 resume as áreas de decisão relevantes e os fatores ambientais que devem ser considerados no desenvolvimento de sistemas de agendamento.

Quadro 4 – Desenvolvimento de um sistema de agendamento (Cayirli e Veral, 2003)

Regra de Agendamento
Tamanho do Bloco
Individual
Múltiplo
Variável
Intervalo entre Serviços
Fixo
Variável
Bloco Inicial
Com
Sem
Qualquer Combinação destes Fatores
Classificação de Pacientes
Sem Classificação (pacientes homogêneos)
Classificação para:
Sequenciamento de Pacientes
Ajuste de Intervalos entre Serviços
Qualquer Combinação destes Fatores
Ajustes
Para <i>No-Shows</i>
Nenhum
Overbooking de Pacientes em Horários Determinados
Redução de Intervalos entre Serviços
Para <i>Walk-Ins</i>, Consultas de Retorno, Urgências e Emergências
Nenhum
Reserva de Horários Determinados
Aumento de Intervalos entre Serviços
Qualquer Combinação destes Fatores

2.4 MÉTODOS DE PESQUISA EM SISTEMAS DE AGENDAMENTO

A depender do ambiente estudado e das premissas adotadas, Cayirli e Veral (2003) segmentam os estudos sobre sistemas de agendamento em três métodos de pesquisa: estudos analíticos, simulações e estudos de caso.

2.4.1 Estudos Analíticos

As abordagens analíticas de sistemas de agendamento compreendem teoria das filas e métodos de programação matemática.

A maior parte dos estudos iniciais assumem um estado contínuo, o que nunca é alcançado em ambientes com um número pequeno e finito de pacientes. Lindley (1952) utiliza um modelo de filas com um único servidor em que as distribuições dos intervalos de chegada dos pacientes e tempos de serviço são arbitradas. É estabelecida uma relação elementar entre os tempos de espera de pacientes sucessivos, o que permite derivar uma distribuição de tempos de espera. Conclui-se que os resultados são expressivamente melhores quando os pacientes são agendados em intervalos regulares, quando da comparação com chegadas aleatórias.

Jansson (1966) avalia um modelo de fila com chegadas determinísticas e um único servidor, derivando uma função de distribuição de custo total (ou seja, custos de espera e ociosidade) para o n ésimo cliente. O bloco inicial ótimo ($n1$) e o intervalo constante entre serviços ($a1$) são definidos para uma determinada razão entre o custo de ociosidade do servidor e o custo de espera do paciente (Cd/Cp) de modo que o custo total médio é minimizado.

Soriano (1996) compara funções de distribuição contínuas de tempos de espera de sistemas de agendamento de blocos individuais e múltiplos com intervalos fixos para diversos fatores de carga, tendo como premissa chegadas determinísticas e tempo de serviço com distribuição gama.

Mercer (1960) inclui a pontualidade do paciente como variável, utilizando uma distribuição geral para atrasos. Obteve, assim, uma distribuição para o tamanho de filas em regime contínuo, para um único servidor com tempos de serviço exponenciais. Assume-se a possibilidade de o paciente não comparecer. Mercer (1973) estende a pesquisa para chegadas em bloco, serviços multiestágio, e distribuição geral para tempos de serviço, considerando diferentes modelos de filas.

Fries e Marathe (1981) avaliam sistemas com blocos variáveis e intervalos fixos, comparando com sistemas de bloco simples e blocos múltiplos com intervalos fixos. Os autores aplicam programação dinâmica para determinar o tamanho ótimo dos blocos (n_i) para o período seguinte, assumindo que um saldo de agendamentos é conhecido. O estudo apresenta um método de aproximação que permite aplicar os resultados dinâmicos para geração de agendamentos na versão estática.

Conforme Cayirli e Veral (2003), o trabalho de Weiss (1990) foi o primeiro a abordar o problema da otimização simultânea de horários de início e sequenciamento de cirurgias. O autor apresenta resultados analíticos para distribuição geral de tempos de serviço para uma sessão de dois pacientes ($N = 2$) e uma solução heurística para problemas mais extensos com o objetivo de minimizar o custo ponderado do tempo de espera do cirurgião e a ociosidade da sala de cirurgia. Quanto à questão do sequenciamento, o estudo prova que o resultado ótimo define que os dois procedimentos sejam programados em ordem crescente de variabilidade quando os tempos de serviço são exponenciais ou uniformes. Para maiores valores de N , o estudo utiliza simulações para comparar diferentes regras de sequência.

Brahimi e Worthington (1991a) analisam modelos de fila com multiservidor e capacidade finita e chegadas não-homogêneas (taxas de chegadas dependentes de horários) e distribuições gerais e discretas de tempos de serviço. Nesta pesquisa, um algoritmo baseado em Cadeias de Markov calcula a distribuição dependente de horários do número de clientes no sistema, obtendo vários indicadores de desempenho. O método é estendido de modo a proporcionar resultados aproximados para distribuições contínuas de tempos de serviço. Apresenta, ainda, resultados para sistemas finitos com intervalos de chegada constantes.

Pegden e Rosenshine (1990) analisam um modelo com um número finito de agendamentos com diferentes intervalos entre chegadas e tempos de serviço exponenciais. Os autores provam que o tempo médio de espera é uma função convexa do intervalo entre chegadas para $N \leq 4$ e desenvolvem um procedimento baseado em Cadeias de Markov para cálculo dos intervalos de serviço ótimos.

Liao et al. (1993) restringem a chegada de pacientes a janelas fixas de tempo com k intervalos e N pacientes. É empregada programação dinâmica para dimensionamento ótimo de blocos para distribuição Erlang de tempos de serviço. A solução dinâmica é

então aplicada como um limite inferior para a solução de um problema estático com um algoritmo *branch-and-bound*, restrito a casos de menor escala. Vanden Bosch et al. (1999) propõem um algoritmo para o mesmo problema com clientes restritos a janelas de tempo. O algoritmo proposto demonstra ser mais eficiente que o método de Liao et al. (1993) para problemas maiores.

Wang (1993) considera tanto casos estáticos quanto dinâmicos para um sistema de servidor único com tempos de serviço exponenciais em que o objetivo é a minimização da soma ponderada do tempo de fluxo do paciente e o tempo total do sistema. O autor demonstra que o tempo de fluxo do paciente pode ser representado em uma distribuição faseada, o que permite a utilização de um método matricial para derivar os tempos de fluxo esperados. Os horários ótimos de agendamento são então calculados com o uso de um procedimento recursivo. Os resultados demonstram que os intervalos ótimos não são constantes, mas em “formato de domo”. Wang (1997) estende a pesquisa para qualquer distribuição de tempo de serviço que possa ser aproximada com uma distribuição faseada.

Liu e Liu (1998b) estudam um sistema de filas com múltiplos servidores em que os horários de chegada dos médicos são aleatórios. O estudo desenvolve uma formulação de programação dinâmica para otimizar os tamanhos de blocos e usam os resultados do caso dinâmico para resolução de problema estático. É feito, então, um comparativo de desempenho, utilizando o método de aproximação para os melhores agendamentos por meio de simulação exaustiva.

Lau e Lau (2000) endereçam dois problemas relevantes para agendamento ambulatorial e cirúrgico: como determinar o custo total do sistema dada uma regra específica de agendamento; e como determinar o agendamento ótimo dada uma sequência específica de chegadas. Os autores apresentam um procedimento eficiente para resolver a primeira questão quando os tempos de serviço têm distribuições gerais e não idênticas. Isto leva à solução do segundo problema por meio da análise de um grande número de sistemas de agendamento e identificação do sistema ótimo mediante um procedimento de busca. A acuracidade deste método de aproximação é avaliado por meio de comparação com os resultados obtidos com simulações.

O estudo de Robinson e Chen (2001) busca otimizar o agendamento quando a sequência de N pacientes já está estabelecida. O problema é formulado como um

programa linear estocástico e resolvido com o uso de integração de Monte Carlo. Os autores utilizam o “formato de domo” da regra ótima como base para o desenvolvimento de uma heurística simples que ajusta os intervalos entre serviços com base na razão Cd/Cp . Demonstra-se que a heurística tem um desempenho robusto em relação a distribuições não especificadas.

Denton e Gupta (2003) apresentam um modelo de programação linear estocástica de dois estágios para determinar o intervalo ótimo entre serviços, aplicando uma abordagem de decomposição para a solução de casos de distribuições gerais independentes e idênticas. Assim como Wang (1993, 1997) e Robinson e Chen (2001), o estudo mostra que os intervalos ótimos apresentam “formato de domo” e destaca que isto é mais acentuado para altos valores de Cd/Cp .

Kaandorp e Koole (2007) aplicam um algoritmo de busca local para construção de uma agenda, minimizando os tempos ponderados de espera de pacientes e ociosidade do servidor.

Segundo Cayirli e Veral (2003) e Cayirli et al. (2012), abordagens analíticas assumem, em sua maioria, premissas não-realistas que limitam sua aplicação, como, por exemplo, tempos de serviço com distribuições de probabilidades exponenciais.

2.4.2 Simulações

Segundo Cayirli e Veral (2003), uma vantagem do uso de simulações em relação a abordagens analíticas é a possibilidade de se modelar sistemas de filas complexos e representar variáveis ambientais, como atributos relacionados a pacientes ou servidores. As pesquisas conduzem experimentos de simulação para avaliar o desempenho de diferentes sistemas de agendamento e/ou compreender a relação entre diversos fatores ambientais e medidas de desempenho. Existem também alguns pacotes com modelos genéricos de simulação, os quais permitem que programadores e administradores avaliem a efetividade de diferentes sistemas de agendamento (KATZ, 1969; PAUL E KULJIS, 1995).

Segundo Cayirli e Veral (2003), o estudo de Bailey (1952) foi o primeiro a analisar um sistema de agendamento de blocos individuais em um tempo em que hospitais ainda utilizavam sistemas de bloco simples. O pesquisador utilizou uma técnica de simulação manual de Monte Carlo para otimizar o tamanho do bloco inicial ($n1$) e o

intervalo entre consultas em clínicas com diferentes valores de N . Como conclusão, um sistema com blocos individuais, intervalos fixos e um bloco inicial com dois pacientes proporciona o melhor equilíbrio entre o tempo de espera dos pacientes e o tempo ocioso dos médicos. Isto é conhecido como a Regra de Bailey, a qual ainda é amplamente estudada na literatura.

Blanco White e Pike (1964) relaxam a premissa de pontualidade tanto de pacientes quanto médicos ao examinar os efeitos do bloco inicial ($n1$), do número de pacientes por bloco (ni) e dos intervalos entre consultas (ai). Foi constatado que, nas duas clínicas investigadas, diferentes sistemas de agendamento têm desempenho melhor, dados os diferentes níveis de pontualidade dos pacientes.

Fetter e Thompson (1996) conduziram experimentos de simulação para analisar os efeitos de diversas variáveis, como atrasos de pacientes, atrasos de médicos, taxas de *no-show*, taxas de pacientes não-agendados, intervalos entre consultas e ocupação. Os resultados confirmam a relevância da pontualidade do médico e exploram o papel de uma ocupação realista da clínica na eficiência da sua operação.

Vissers e Wijngaard (1979) propõem a redução das variáveis de um sistema de agendamento em cinco: tempo médio de serviço, coeficiente de variação do tempo de serviço, desvio padrão da pontualidade do paciente, número de agendamentos por sessão e antecipação média do sistema. A antecipação média engloba todos os fatores que reduzem o risco de ociosidade do servidor, como chegada antecipada dos pacientes, atraso de médicos, agendamento em bloco ($ni > 1$), agendamento de bloco inicial ($n1 > 1$) e programação de intervalos de serviços inferiores ao tempo médio de consulta. Em outro estudo relacionado, Vissers (1979) estende a análise para vários valores de N , desenvolvendo heurísticas para a seleção de sistemas de agendamento considerando estas cinco variáveis e um equilíbrio entre tempos de espera e ociosidades.

Charnetski (1984) emprega simulação para estudar o problema da alocação de janelas de tempo para cirurgias utilizando a regra FCFS com o objetivo de balancear o custo de espera do cirurgião e o custo de instalações e equipes. A heurística proposta demonstra que diferentes tipos de procedimentos têm diferentes distribuições de tempo de serviço e define janelas de atendimento baseadas na média e desvio-padrão dos tempos *individuais* de serviço.

Quando comparados a estudos anteriores de avaliação de ambientes, Ho e Lau (1992, 1999) e Ho et al. (1995) são os mais abrangentes, contemplando 50 regras de agendamento sob vários ambientes de operação. Estas pesquisas trazem sistemas de agendamento de blocos individuais e intervalos variáveis e testam seu desempenho em relação a regras tradicionais. A melhor regra de intervalo variável identificada permite que pacientes cheguem em intervalos menores no início de uma sessão e em intervalos maiores ao seu final. Conclui-se que não existe uma regra que tenha bom desempenho sob todas as circunstâncias e propõe-se uma heurística simples para seleção de uma regra de atendimento conforme os valores de p , CV , N , e Cp/Cd . O procedimento propõe uma “fronteira de eficiência” para cada ambiente (combinações de p , CV e N) simulado, construída com diferentes valores de Cp/Cd , proporcionando um *framework* único para comparativos de desempenho de sistemas de agendamento. Os autores concluem que p , CV e N afetam um sistema de agendamento em ordem decrescente de importância.

Klassen e Rohleder (1996) apresentam sistemas de agendamento que classificam pacientes com base em suas variabilidades de tempo de serviço esperadas e usam simulação para comparar alternativas para sequenciamento de pacientes com variâncias baixas e altas para intervalos de serviços constantes. Constata-se que sistemas que agendam pacientes de baixa variância no início de uma sessão (regra LVBEG) apresentam resultado superior às regras de melhor desempenho de Ho e Lau. A pesquisa modela, ainda, casos de urgência. Em outro estudo, Rohleder e Klassen (2000) consideram a possibilidade de o programador cometer algum erro ao classificar pacientes e, principalmente, a possibilidade de não adotar a sequência perfeita quando eventualmente os pacientes insistem em ser atendidos em horários específicos. Mesmo contemplando estas premissas mais realistas, a regra LVBEG ainda apresenta desempenho superior.

Liu e Liu (1998a) pesquisam um sistema de agendamento de blocos variáveis e intervalos fixos para um sistema de filas com multiservidor em que os médicos podem atrasar. Um procedimento de simulação de busca é desenvolvido para determinação do número de pacientes a ser agendado em cada bloco (n) que minimizará o custo total do fluxo de pacientes e ociosidade dos médicos. Usando os resultados das melhores regras, derivadas após a simulação de vários fatores ambientais (número de médicos, probabilidade de *no-shows*, número de blocos e Cd/Cp), é proposto um

procedimento simples para encontrar uma regra de agendamento para um determinado ambiente.

Yang et al. (1998) propõem uma heurística que é apresentada como uma função matemática da média e do desvio padrão dos tempos de serviço e uma “constante de planejamento k ”, onde k é calculado para um ambiente particular (ou seja, combinações de CV , p , N e Cd/Cp) usando um modelo de regressão. Esta regra procura explicitamente ser mais “justa” mediante o aumento do intervalo entre consultas ao longo da sessão para que sejam evitados tempos de serviço desuniformes. O estudo emprega simulação para comparar o desempenho da heurística com as melhores regras propostas por Ho e Lau (1992) e, da mesma maneira que estes, constrói uma “fronteira da eficiência” com diferentes valores de Cp/Cd .

Swisher et al. (2001) desenvolvem um modelo de simulação de eventos discretos, o qual pode ser utilizado para tomada de decisão em serviços ambulatoriais. O modelo é aplicado em clínicas particulares e demonstra que os resultados são expressivamente sensíveis a mudanças no mix de pacientes, nos agendamentos e no nível das equipes. Quanto ao agendamento, o estudo considera apenas o efeito da alteração de horário em que uma certa categoria de paciente está agendada, em vez de comparar diferentes regras de agendamento.

Cayirli et al. (2012), com o objetivo de propor uma “regra universal” para sistemas de agendamento, simulam seu desempenho em vários ambientes. Além dos parâmetros CV , p , N e Cd/Cp , acrescentam a probabilidade de pacientes não agendados (*walk-ins*). De modo similar ao trabalho de Yang et al. (1998), uma constante de planejamento é proposta. Entretanto, a formulação matemática proporciona um “formato de domo” aos intervalos entre serviços. Assim como nos estudos de Ho e Lau (1992, 1999), Ho et al. (1995) e Yang et al. (1998), é montada uma fronteira da eficiência.

Cayirli e Yang (2014) estendem este último estudo com a inclusão de classificação de pacientes e uma análise de sensibilidade para o caso de pacientes não-pontuais.

Cayirli e Veral (2003) destacam que as pesquisas que utilizam simulação geralmente falham em reportar as técnicas de redução de variância empregadas e/ou a significância estatística dos resultados.

2.4.3 Estudos de Caso

Em estudos de caso, os pesquisadores analisam um ambiente em particular, fazem recomendações para aprimoramento do sistema existente e, em alguns casos, avaliam os resultados de implantações (CAYIRLI; VERAL, 2003). Muito embora os estudos de caso ofereçam recomendações valiosas com base em situações reais, sua maior desvantagem é a falta de generalização.

Villegas (1967) realiza um estudo em uma clínica de um departamento ambulatorial vivenciando práticas reais. A pesquisa compara o desempenho de um certo número de sistemas de agendamento de blocos variáveis e intervalos fixos com relação a tempos de espera de pacientes e ociosidade de médicos.

Williams et al. (1967) empregam simulação para analisar uma clínica universitária com o objetivo de melhorar o agendamento de pacientes e médicos. Os autores demonstram que, quando blocos múltiplos são utilizados em oposição a blocos simples, os tempos de espera de pacientes são reduzidos substancialmente, sem prejuízo para a utilização das equipes.

Johnson e Rosenfeld (1968) analisaram fatores que afetam o tempo de espera de pacientes em oito hospitais utilizando uma abordagem de observação. Concluem que o sistema de agendamento em uso é o fator que mais influencia os tempos de espera e tanto sistemas de agendamento de blocos individuais quanto múltiplos apresentam desempenho superior aos sistemas de bloco simples.

Em estudo realizado em um hospital, Rockart e Hofmann (1969) observam que quando as clínicas migram para sistemas de agendamento de blocos individuais, em que os pacientes são alocados em um determinado horário para um médico específico, tanto pacientes quanto médicos são mais pontuais e as taxas de *no-show* se reduzem. Com base em dados coletados no mesmo hospital, Hofmann e Rockart (1969) analisam as variáveis que afetam as taxas de *no-show* em clínicas ambulatoriais, constatando que o intervalo de tempo entre a solicitação do serviço e a data agendada é um fator relevante. Quanto maior este intervalo, maior o percentual de *no-shows*.

Walter (1973) emprega simulação para modelagem de um sistema de filas em um departamento de radiologia e explora os efeitos da variação de n , m , CV e a relação de pacientes com agendamento (r). O estudo também investiga o efeito de se dividir

as sessões em grupos mais homogêneos, concluindo que mesmo um simples agrupamento de pacientes ambulatoriais e de internação resulta em uma redução significativa na ociosidade dos médicos.

Rising et al. (1973) utilizam modelagem de simulação para analisar uma clínica ambulatorial em uma universidade, com o objetivo de aprimorar o agendamento de médicos e pacientes. Os autores sugerem agendar pacientes de modo a complementar os padrões de chegada de pacientes não-agendados, resultando em uma suavização da taxa real de chegadas. Os resultados da implantação sugerem evolução em termos de redução de horas-extras e tempo de espera de pacientes não-agendados.

Em uma simulação para uma clínica de otorrinolaringologia, Cox et al. (1985) avaliam uma série de sistemas de agendamento com variações nos parâmetros a_i , n_1 e n_i , bem como várias formas de sequenciamento de pacientes novos e de retorno no momento do agendamento. Os pesquisadores validam o modelo mediante comparativos com situações reais. Quando implantadas, as regras melhoram o fluxo de pacientes, tamanhos uniformes de filas e cargas uniformes de trabalho para os médicos.

Mahacheck e Knabe (1984) empregam simulação para analisar alternativas de decisões operacionais com respeito a agendamento de pacientes, demanda de equipes, mix de pacientes e dimensionamento de instalações. As regras avaliadas envolvem um esquema de classificação de pacientes novos e de retorno.

O'Keefe (1985) usa uma abordagem predominantemente qualitativa em sua análise das operações de três departamentos ambulatoriais. O sistema de agendamento proposto, o qual considera uma classificação de pacientes novos *versus* pacientes de retorno, é rejeitado pela equipe, que prefere manter o sistema simples e uniforme em toda a instituição. De modo similar, o autor enfrenta enorme resistência por parte dos médicos, os quais se recusam a mudar seus hábitos. Este estudo de caso ilustra bem o fato de um sistema de agendamento ser, antes de tudo, um problema "político".

Babes e Sarma (1991) investigam uma clínica que adota um sistema de bloco simples onde os pacientes são agendados para um determinado dia, sem um horário especificado. Sob estas condições, o problema é reduzido a determinar o número de pacientes por sessão (N) e o número de médicos (S) que irão otimizar o desempenho em custos do sistema. Inicialmente os autores aplicam modelos contínuos de teoria

das filas. No entanto, quando os resultados se mostram muito diferentes dos observados na operação real, passam a utilizar modelagem de simulação. São avaliadas as sensibilidades de indicadores de desempenho para N , S , o atraso de médicos e o tempo médio de serviço.

Brahimi e Worthington (1991b) aplicam seu trabalho anterior em modelo de filas dependentes de duração (Brahimi e Worthington 1991a) ao problema de aprimoramento de sistemas de agendamento em sete clínicas do Reino Unido. São comparados diferentes sistemas com base em indicadores de desempenho e, como resultado, os autores sugerem um sistema com blocos individuais, intervalos fixos e um bloco inicial de três pacientes, observando melhora no tempo de espera dos pacientes após a implantação.

Huang e Lee (1996) usam simulação para modelar o departamento ambulatorial de um hospital que não utiliza sistema de agendamento, com o objetivo de aprimorar o tempo de espera dos pacientes e a utilização dos médicos. Os pesquisadores reportam que não foi possível implantar um sistema de blocos individuais devido à resistência da equipe. Em vez disto, recomendaram estender o horário de trabalho dos médicos para um melhor equilíbrio entre oferta e demanda.

Bennet e Worthington (1998) aplicam uma abordagem de sistemas em que consideram a interação do departamento ambulatorial com outras unidades de um hospital. Os autores observam que o *overbooking* e o agendamento excessivo de consultas de retorno criam o principal problema de capacidade. As recomendações não puderam ser implantadas, uma vez que demandavam mudanças comportamentais por parte dos médicos.

Lehaney et al. (1999) propõem um sistema de agendamento que classifica pacientes em ordem crescente de tempos de serviço, similar ao sequenciamento de tarefas pela regra do menor tempo de processamento (SPT) em ambientes de produção. Muito embora a pesquisa se dê em um contexto de serviço ambulatorial, em uma aplicação real o programador precisa alocar uma janela de tempo para um paciente sem saber se o próximo irá demandar um tempo menor ou maior de serviço. Os autores reconhecem que também não seria prática a tentativa de atribuir estimativas precisas para durações individuais de consultas.

Vanden Bosch e Dietz (2000) avaliam políticas de agendamento e sequenciamento em uma clínica, a qual utiliza um esquema de classificação baseado no histórico de agendamentos dos pacientes ou tipos de procedimento. Esta é a primeira tentativa de se estudar o melhor mix e sequência de pacientes ao longo de vários dias. É apresentada uma abordagem analítica para resolução do problema estático onde todos os pacientes que necessitam de um agendamento para o dia são conhecidos antecipadamente. Constata-se que não há uma regra simples para a sequência ótima, sendo difícil generalizar quaisquer resultados de sequenciamento por ordem de médias ou variâncias de tempos de serviço. Além disso, devido à dimensão do problema, a solução ótima só pode ser atingida para situações de pequena escala.

Para os casos mais complexos de agendamento e sequenciamento dinâmicos os autores desenvolvem heurística para uma política e testam seu desempenho mediante uso de simulação.

Cayirli e Gunes (2013) comparam o desempenho de diferentes sistemas de agendamento em clínicas de diversas especialidades com foco em sazonalidade de *walk-ins*.

O Quadro 5 resume as várias metodologias de análise que endereçam problemas de agendamento conforme observado na literatura.

Quadro 5 – Metodologias de análise (Cayirli e Veral, 2003)

Estudos Analíticos
Teoria das Filas
Programação Matemática
Programação Dinâmica
Programação Não-Linear
Programação Linear Estocástica
Estudos com Simulações
Análise Ambiental (quais fatores afetam os indicadores de desempenho?)
Comparativos de Desempenho de Alternativas de Sistemas de Agendamento
Estudos de Caso
Observacionais
Experimentos com Alternativas de Sistemas de Agendamento em Contextos Reais
Modelagem Quantitativa (simulação, modelagem de filas etc.) para Construção de Alternativas de Sistemas de Agendamento, com ou sem Análise Pós-Implementação

Um resumo dos principais fatores que diferenciam as características das pesquisas analisadas pode ser encontrado no Quadro 6.

Pode-se perceber o predomínio de sistemas de agendamento estáticos, premissa de pontualidade dos pacientes, um único servidor e abordagens analíticas, sendo que para os demais fatores observa-se uma distribuição mais uniforme das suas características.

Quadro 6 – Principais características das pesquisas analisadas

Pesquisa	Natureza do Sistema	Inclusão de No-Shows	Intervalo entre Serviços	Pontualidade do Paciente	Multiservidor	Metodologia de Análise
Anderson et al. (2013)	Estática	N	Variável	S	N	Simulação
Babes e Sarma (1991)	Estática	N	Não se aplica	N	S	Estudo de Caso
Bailey (1952); Welch e Bailey (1952)	Estática	N	Fixo	S	N	Estudo Analítico
Berg et al. (2014)	Estática	S	Variável	S	S	Estudo de Caso
Blanco White e Pike (1964)	Estática	S	Fixo	N	N	Estudo Analítico
Brahimi e Worthington (1991a)	Estática	N	Não-especificado	S	S	Estudo Analítico
Brahimi e Worthington (1991b)	Estática	S	Fixo	S	N	Estudo de Caso
Cayirli et al. (2006)	Estática	S	Variável	N	N	Simulação
Cayirli et al. (2012)	Estática	S	Variável	S	N	Simulação
Cayirli e Yang (2014)	Estática	S	Variável	S	N	Simulação
Cayirli e Gunes (2013)	Estática	N	Fixo	S	N	Estudo de Caso
Cox et al. (1985)	Estática	S	Fixo	N	S	Estudo de Caso
Denton e Gupta (2001)	Estática	N	Variável	S	N	Estudo Analítico
Erdogan e Denton (2013)	Dinâmica	S	Variável	S	N	Estudo Analítico
Fetter e Thompson (1966)	Estática	S	Fixo	N	S	Simulação
Fries e Marathe (1981)	Dinâmica	N	Fixo	S	N	Estudo Analítico
Green et al. (2006)	Dinâmica	S	Fixo	S	N	Estudo Analítico
Ho e Lau (1992, 1999); Ho et al. (1995)	Estática	S	Variável	S	N	Simulação
Huang e Lee (1996)	Estática	N	Variável	S	S	Estudo de Caso
Jansson (1966)	Estática	N	Variável	S	N	Estudo Analítico
Kaandorp e Koole (2007)	Estática	S	Fixo	S	N	Estudo Analítico
Klassen e Rohleder (1996); Rohleder e Klassen (2000)	Estática	S	Fixo	S	N	Simulação
Lau e Lau (2000)	Estática	N	Variável	S	N	Estudo Analítico
Lehaney et al. (1999)	Estática	N	Variável	S	S	Estudo de Caso
Liao et al. (1993)	Dinâmica	N	Fixo	S	N	Estudo Analítico
Lindley (1952)	Estática	N	Fixo	S	N	Estudo Analítico
Liu e Liu (1998a)	Estática	S	Fixo	S	S	Simulação
Liu e Liu (1998b)	Estática/Dinâmica	S	Fixo/Variável	S	S	Estudo Analítico
Mahacheck e Knabe (1984)	Estática	N	Fixo	S	S	Estudo de Caso
Mercer (1960, 1973)	Estática	S	Fixo	N	S/N	Estudo Analítico
Muthuraman e Lawley (2008)	Estática	S	Fixo	S	N	Simulação
O'Keefe (1985)	Estática	S	Fixo	N	S	Estudo de Caso
Pegden e Rosenshine (1990)	Estática	N	Fixo	S	N	Estudo Analítico
Rising et al. (1973)	Estática	N	Fixo	S	S	Estudo de Caso
Robinson e Chen (2001)	Estática	N	Variável	S	N	Estudo Analítico
Soriano (1966)	Estática	N	Variável	S	N	Estudo Analítico
Swisher et al. (2001)	Estática	N	Não-especificado	S	S	Simulação
Tsai e Teng (2014)	Estática	S	Fixo	S	N	Simulação
Vanden Bosch e Dietz (2000)	Estática	S	Variável	S	N	Estudo de Caso
Vanden Bosch et al. (1999)	Estática	N	Fixo	S	N	Estudo Analítico
Vissers e Wijngaard (1979); Vissers (1979)	Estática	S	Fixo	N	N	Simulação
Walter (1973)	Estática	N	Fixo	S	N	Estudo de Caso
Wang (1993)	Estática/Dinâmica	N	Variável	S	N	Estudo Analítico
Wang (1997)	Estática	N	Variável	N	N	Estudo Analítico
Yang et al. (1998)	Estática	S	Variável	S	N	Simulação

2.5 SELEÇÃO DE POTENCIAIS MODELOS PARA APLICAÇÃO NO CASO

Uma vez revisada a literatura, foram obtidos os elementos para resposta à primeira questão de pesquisa:

- Quais são os sistemas de agendamento existentes na literatura e direcionados para aplicação em serviços ambulatoriais?

Conforme descrito no item 2.3 deste capítulo, um sistema de agendamento é composto de três elementos: regras de agendamento, classificação de pacientes e ajustes para *no-shows* e *walk-ins*, sendo que as principais configurações para estes fatores foram resumidas no Quadro 4.

Dentre os modelos analisados na literatura, vários podem ser parametrizados de modo a representar mais de um sistema de agendamento. Como exemplo, o modelo de Cayirli et al. (2012), o qual contempla ajustes para *no-shows* e *walk-ins*, permite parametrizar as probabilidades destes fatores como nulos, convertendo o modelo para ser aplicado como um sistema sem estes ajustes. Da mesma forma, se o modelo de Bailey (1952), o qual considera dois pacientes no primeiro horário, for adaptado para apenas um paciente inicial, o sistema passa a ser de bloco individual sem bloco inicial, sem que isso demande o desenho de um novo modelo.

Tendo como princípio esta possibilidade de se configurar modelos mais “completos” para cenários menos complexos, esta análise considera as 24 combinações entre os fatores: intervalos fixos ou variáveis, blocos individuais, múltiplos ou variáveis, com ou sem ajustes para *no-shows* e *walk-ins* e com ou sem classificação de pacientes.

Destas 24 combinações, 18 foram encontradas na literatura, compondo a resposta para a questão de pesquisa formulada, conforme apontado no Quadro 7. Neste quadro pode-se verificar, ainda, que seis combinações não foram identificadas na revisão teórica, as quais estão concentradas em blocos múltiplos e variáveis.

Ainda com os elementos deste capítulo, pode-se responder à segunda questão de pesquisa:

- Quais modelos de sistemas de agendamento encontrados na literatura apresentam maior potencial para aplicação em prestadores de serviços de medicina diagnóstica com ampla diversidade de serviços agendados?

Quadro 7 – Sistemas de Agendamento encontrados na literatura

Intervalo entre Serviços	Bloco	Ajustes No-Shows/Walk-Ins	Classificação Pacientes	Encontrado na Literatura?
Fixo	Individual	Sim	Sim	Sim
Fixo	Individual	Sim	Não	Sim
Fixo	Individual	Não	Sim	Sim
Fixo	Individual	Não	Não	Sim
Fixo	Múltiplo	Sim	Sim	Sim
Fixo	Múltiplo	Sim	Não	Sim
Fixo	Múltiplo	Não	Sim	Sim
Fixo	Múltiplo	Não	Não	Sim
Fixo	Variável	Sim	Sim	Não
Fixo	Variável	Sim	Não	Sim
Fixo	Variável	Não	Sim	Sim
Fixo	Variável	Não	Não	Sim
Variável	Individual	Sim	Sim	Sim
Variável	Individual	Sim	Não	Sim
Variável	Individual	Não	Sim	Sim
Variável	Individual	Não	Não	Sim
Variável	Múltiplo	Sim	Sim	Não
Variável	Múltiplo	Sim	Não	Não
Variável	Múltiplo	Não	Sim	Não
Variável	Múltiplo	Não	Não	Sim
Variável	Variável	Sim	Sim	Não
Variável	Variável	Sim	Não	Sim
Variável	Variável	Não	Sim	Não
Variável	Variável	Não	Não	Sim

Conforme já mencionado, um modelo pode representar vários sistemas de agendamento. Da mesma forma, um estudo da literatura pode conter mais de um modelo ou sistema de agendamento. Deste modo, um primeiro filtro foi aplicado aos estudos, eliminando os que não contemplam ajustes para *no-shows*, com base nos dados do Quadro 6. Conforme consta na literatura pesquisada, não se deve considerar a possibilidade de eliminar *no-shows*, mas apenas reduzi-los (DEYO; INUI, 1980; BARRON, 1980; SCHAFFER, 1986; OLESKOVICZ et al., 2014). Com este critério inicial, restringiu-se as 51 pesquisas da literatura a 27.

Deste total de 27 pesquisas foram excluídos os estudos analíticos, uma vez que geralmente utilizam premissas não-realistas que limitam sua aplicação, como, por exemplo, tempos de serviço com distribuições de probabilidades exponenciais, além de permitirem um número restrito de variáveis (CAYIRLI; VERAL, 2013; CAYIRLI et al., 2012). Percebe-se na literatura pesquisada, ainda, que vários modelos analíticos

têm apenas o propósito de servir como base para uma posterior modelagem com uso de simulação. Finalmente, em alguns casos, os modelos analíticos têm limitação quanto às faixas de valores para as variáveis. Os 19 estudos restantes deste filtro foram, então, limitados a nove trabalhos, com a exclusão de pesquisas com intervalos fixos entre serviços.

Conforme constatado nesta revisão de literatura, os estudos de Ho e Lau (1992), Ho e Lau (1999), Ho et al. (1995), Cayirli et al. (2012) e Cayirli e Yang (2014) fazem diversos comparativos entre regras de agendamento de intervalos fixos e variáveis entre serviços, sendo que o desempenho de intervalos variáveis se mostrou consistentemente superior. Além disso, diferentemente dos modelos de intervalos fixos que apresentam desempenho satisfatório somente em determinadas situações, os modelos de intervalo variável mostraram maior versatilidade quanto a mudança de fatores ambientais, condição presente nos objetivos deste estudo.

Analisando-se as propostas destes nove estudos, percebeu-se que as pesquisas de Ho e Lau (1992), Ho e Lau (1999), Ho et al. (1995) podem ser consideradas como um único modelo, dado que consideram conjuntos de regras de agendamento semelhantes, com uma mesma abordagem, havendo extensões e aprimoramentos de um estudo para outro. O mesmo pode ser constatado com os estudos de Cayirli et al. (2012) e Cayirli e Yang (2014), uma vez que este último pode ser parametrizado para ser aplicado como o primeiro. Assim sendo, nos 51 estudos revisados, foram identificados 6 modelos com potencial para atender aos objetivos deste trabalho, conforme apresentado no Quadro 8.

Quadro 8 – Potencias modelos da literatura para aplicação em medicina diagnóstica com ampla diversidade de serviços agendados

Modelo
Berg et al. (2014)
Cayirli et al. (2006)
Cayirli et al. (2012); Cayirli e Yang (2014)
Ho e Lau (1992, 1999); Ho et al. (1995)
Vanden Bosch e Dietz (2000)
Yang et al. (1998)

A Figura 2 resume a aplicação das premissas e critérios nesta seleção.

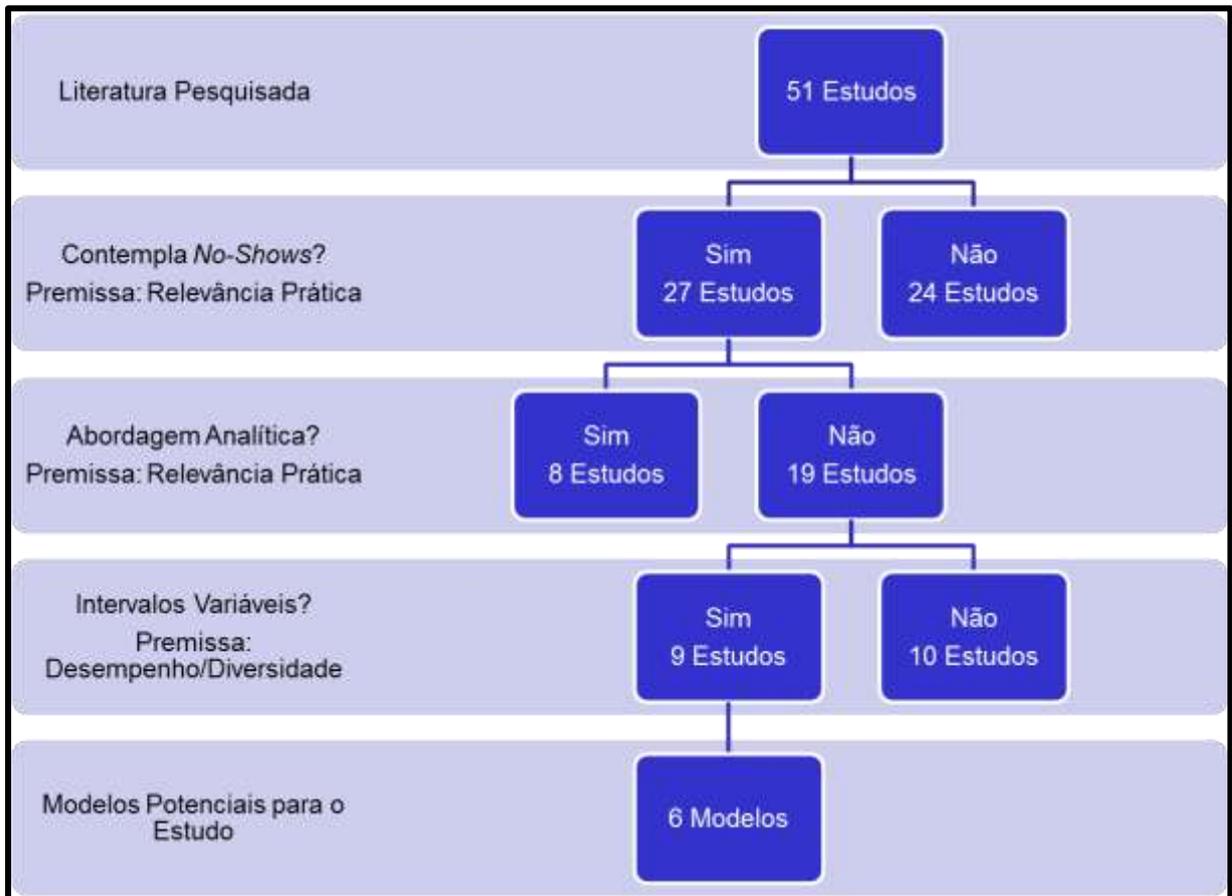


Figura 2 - Seleção de modelos da literatura compatíveis com os objetivos do estudo

A revisão de literatura deste trabalho permite, ainda, responder à terceira questão de pesquisa:

- Quais indicadores de desempenho podem ser utilizados para avaliar a adequação dos modelos de sistemas de agendamento encontrados na literatura com maior potencial para aplicação em prestadores de serviços de medicina diagnóstica com ampla diversidade de serviços agendados?

A avaliação desses modelos pode ser realizada com o uso dos indicadores da literatura levantados neste capítulo, os quais estão resumidos no Quadro 3.

3 MÉTODO DE PESQUISA

Este capítulo tem por objetivo descrever as premissas adotadas para o desenvolvimento do trabalho, bem como os conceitos utilizados e as etapas metodológicas empregadas, estabelecendo a sua relação com os objetivos e questões de pesquisa.

3.1 PREMISSAS DO ESTUDO

As análises do sistema de agendamento a ser proposto foram realizadas com a premissa de independência entre sessões de atendimento. Deste modo, assumiu-se que tanto o desempenho quanto as decisões referentes a uma sessão não interferem em outras sessões, mesmo sendo para um mesmo serviço. Conforme exposto na literatura revisada, esta premissa é predominante nos estudos sobre sistemas de agendamento.

Os tempos de serviço e as taxas de *no-shows* (absenteísmo) e *walk-ins* (demanda de pacientes não agendados) foram considerados dados de entrada para o estudo, com base nas médias históricas coletadas e tratadas. Assim como nos sistemas de agendamento identificados na revisão teórica, modelos de predição para estas variáveis não fizeram parte do escopo deste trabalho.

3.2 ETAPAS METODOLÓGICAS

Para Bryman (1989), as estratégias de pesquisa são classificadas em: experimento, questionário, pesquisa qualitativa, estudo de caso e pesquisa-ação. Quanto aos métodos, as pesquisas podem empregar: questionário auto-administrado, entrevista estruturada, observação participante, entrevista não-estruturada, observação estruturada, simulação e pesquisa documental. Estas classificações podem ser observadas no Quadro 9.

A estratégia da presente pesquisa está mais próxima de um experimento de campo, uma vez que estabelece uma relação de causalidade entre variáveis em modelos de sistemas de agendamento e utiliza dados de um caso real, provenientes de um prestador de serviços de medicina diagnóstica (BRYMAN, 1989).

Quadro 9 – Estratégias e Métodos de Pesquisa

Fonte: Bryman (1989, pg 29)

Estratégias	Métodos
Experimento (principais distinções: experimentos de laboratório e de campo; experimentos e quase-experimentos)	Questionário Auto-Administrado
Questionário	Entrevista Estruturada
Pesquisa Qualitativa	Observação Participante
Estudo de Caso	Entrevista Não-Estruturada
Pesquisa-Ação	Observação Estruturada
	Simulação
	Pesquisa Documental

Cayirli e Veral (2003) e Gupta e Denton (2008) destacam a importância de contextos reais em estudos acerca de sistemas de agendamento ambulatoriais, o que permite maior aproximação entre teoria e prática. Conforme Bryman (1989), experimentos levam ao conhecimento prático e relevante.

Para a realização deste estudo foi empregada modelagem quantitativa baseada em simulação, sendo este o principal método adotado. Os dados para as simulações foram obtidos em uma empresa de medicina diagnóstica.

Bryman (1989) descreve o método de simulação como uma tentativa de representar uma realidade, abstraindo-se as variáveis relacionadas às perguntas de pesquisa e criando um ambiente em que a operação destas variáveis possa ser observada.

Ainda quanto aos métodos, foram aplicadas entrevistas não-estruturadas com o objetivo de compreender o contexto operacional da empresa estudada. A implantação de modelos não foi contemplada no escopo deste trabalho.

Embora classificada como simulação, esta pesquisa teve as etapas metodológicas construídas conforme a estrutura de estudo de caso proposta por Yin (2015), uma vez que o contexto do campo balizou tanto a modelagem como os valores das variáveis nas simulações realizadas.

Segundo Martins e Theóphilo (2009), não há um conjunto fixo de etapas ou uma sistematização de um projeto para um estudo de caso. Contudo, é possível elaborar um plano com uma sequência lógica de ações, balizado pelos objetivos da pesquisa, incluindo coleta, tratamento e análise de dados.

Yin (2015) descreve a pesquisa de um estudo de caso como um processo linear, porém com iterações entre etapas, conforme ilustrado na Figura 3. Norteado por este processo e em linha com os objetivos definidos, este trabalho foi desenvolvido mediante realização das etapas descritas a seguir.

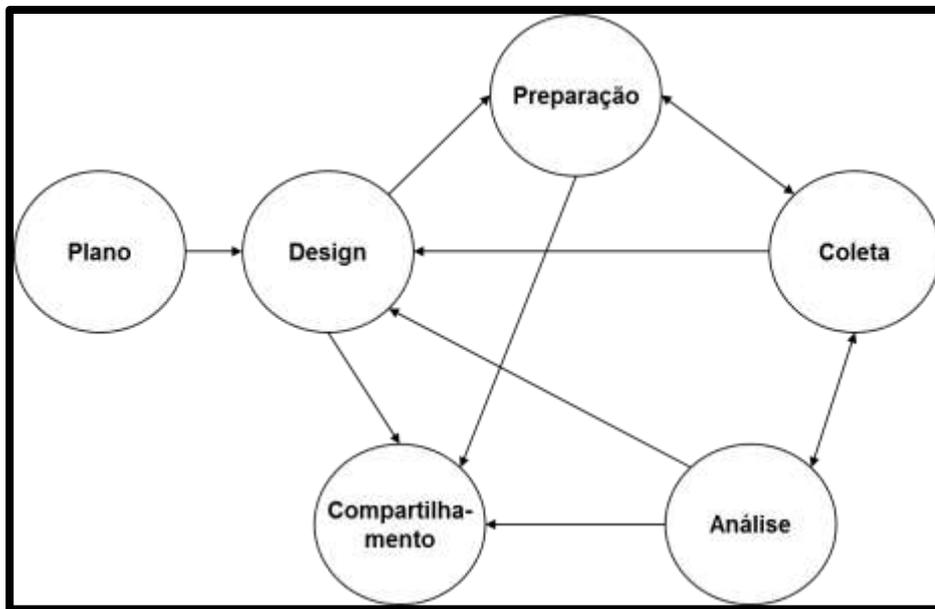


Figura 3 - Processo de Desenvolvimento de um Estudo de Caso

Fonte: Yin (2015, pg 1)

3.2.1 Plano de Pesquisa

Esta etapa teve como objetivo selecionar o método de pesquisa a ser empregado, sendo que se optou pela simulação com modelagem quantitativa, conforme já abordado neste capítulo.

Também foi classificado trabalho como sendo do tipo descritivo, uma vez que, conforme Yin (2015), seu propósito é de investigar um fenômeno contemporâneo em um contexto de mundo real.

3.2.2 Projeto de Pesquisa

O projeto de pesquisa é a fase que orienta o processo de coleta, análise e interpretação dos dados, proporcionando um modelo lógico que permite fazer inferências acerca das relações entre as variáveis investigadas (YIN, 2015).

Para o presente trabalho adotou-se um caso único para análise. Entre os motivos para esta escolha, Yin (2015) destaca a seleção de um caso crítico como meio para determinar se a teoria ou as proposições da pesquisa estão corretas ou se algum conjunto alternativo de explicações pode ser mais relevante.

Contribuíram para esta decisão as limitações de acesso a prestadores de serviços de saúde com os fatores ambientais desejados e as restrições de prazo inerentes a uma dissertação.

Tendo como base as questões formuladas neste estudo, definiu-se como unidade de análise um prestador de serviços de saúde que proporcionasse os elementos constantes no escopo desta pesquisa:

- Ser um prestador de serviços agendados de medicina diagnóstica
- Oferecer um portfólio diversificado de exames

Atendidos estes requisitos, foram consideradas a disponibilidade e a adequação dos dados para a pesquisa como critérios complementares.

Dados estes fatores, foi selecionado um laboratório de medicina diagnóstica de grande porte, com um portfólio de cerca de 200 tipos serviços agendados, com 10 unidades de atendimento e a realização de mais de 700.000 exames agendados anualmente, além de exames não-agendados.

Nesta instituição os serviços agendados apresentam taxas significativas de *no-shows*, enquanto que a admissão de clientes não-agendados (*walk-ins*) para estes exames não é praticada. Os serviços, em sua totalidade, são caracterizados como de estágio único, ou seja, não são realizados em etapas.

3.2.3 Preparação para Coleta de Dados

A seleção da unidade de análise para realização do estudo, um dos objetivos desta etapa, seguiu os critérios definidos na fase de projeto de pesquisa.

Acesso à instituição e dados históricos disponíveis, suficientes e adequados para realização da pesquisa também compuseram os requisitos para esta escolha.

A preparação para a coleta foi dividida entre obtenção de dados quantitativos, na forma de um conjunto de variáveis, e uma lista de tópicos a serem cobertos por meio de entrevistas.

As variáveis selecionadas, apresentadas no Quadro 10, foram delineadas de modo a viabilizar a análise dos resultados mediante a Equação 1, conforme especificado no projeto de pesquisa.

Quadro 10 – Variáveis dos modelos

Variáveis de Entrada	Variáveis de Decisão	Medidas de Desempenho
Variância do Tempo de Serviço	Bloco Inicial ($n1$)	Custo de Espera de Clientes (Cp)
Taxa Média No-Show	Intervalo Serviços (ai)	Custo de Ociosidade Provedor (Cd)
Variância No-Show		Custo de Horas-Extras Provedor (Co)
Taxa Média <i>Walk-Ins</i>		Custo Total Provedor (TC)
Variância <i>Walk-Ins</i>		
Tamanho das Sessões		
Atraso Médio de Pacientes		
Variância do Atraso de Pacientes		
Número Servidores		

Quanto às entrevistas, estas tiveram como base as seguintes questões:

- Existem casos de *no-shows* e *walk-ins*?
- Quais os procedimentos de ajustes para *no-shows* e *walk-ins*?
- São realizadas alterações de agenda após o início de uma sessão?
- Quais variáveis são medidas e armazenadas?

Estas questões, no entanto, serviram com um guia para condução das entrevistas, e não como um questionário a ser aplicado.

As entrevistas realizadas foram gravadas, integrando a base de dados da pesquisa.

3.2.4 Coleta de Dados

Em função dos processos e organograma da instituição pesquisada, os dados foram obtidos em três unidades de coleta: elaboração das agendas, agendamento de

clientes e especificação e envio de dados históricos. Na forma de entrevistas, os dados foram coletados seguindo a preparação planejada.

Na primeira entrevista procurou-se entender a elaboração das agendas e seus critérios, assim como identificar as restrições e premissas a serem endereçadas pelo modelo a ser proposto, focando as variáveis de decisão e de desempenho.

Nesta etapa pôde ser avaliada, também, a diversidade do portfólio de serviços.

O agendamento de clientes foi o alvo da segunda entrevista, onde foi identificado o modo como as agendas são preenchidas, buscando levantar eventuais desvios em relação ao planejamento.

Na terceira etapa foi apresentada a especificação dos dados históricos para composição das variáveis de entrada para o modelo, procedendo-se a sua coleta.

O Quadro 11 apresenta as áreas da organização em que foram realizadas as entrevistas e as respectivas questões.

Quadro 11 –Entrevistas para Coleta de Dados

Área	Atividade Relacionada	Principais Perguntas
Diagnóstico por Imagem	Elaboração das Agendas	Existem casos de <i>no-shows</i> e <i>walk-ins</i> ? Quais os procedimentos para ajustes de <i>no-shows</i> e <i>walk-ins</i> ?
<i>Call Center</i>	Agendamento de Clientes	São realizadas alterações de agenda após o início de uma sessão?
Inteligência de Mercado	Monitoramento e Armazenagem de Dados	Quais variáveis são medidas e armazenadas?

3.2.5 Análise

Esta fase compreendeu as etapas que vão da adequação dos dados coletados à obtenção dos resultados que buscam responder às questões de pesquisa, conforme descrito a seguir.

3.2.5.1 Tratamento de Dados

Inicialmente os dados coletados foram submetidos a uma adequação de tamanho para viabilização do seu tratamento nos ambientes e ferramentas utilizadas. Na sequência, novos filtros foram aplicados no sentido de reduzir a complexidade, mantendo, entretanto, uma representatividade suficiente para a pesquisa.

Em seguida buscou-se adequar a forma dos dados quantitativos às variáveis especificadas no Quadro 9, lançando mão de cálculos matemáticos e técnicas estatísticas.

Como resultado, os dados tratados ficaram compatíveis com as formas encontradas na literatura acerca do tema. Quanto aos números obtidos, estes permitiram identificar as faixas em que as variáveis se enquadravam, fator essencial para a etapa de seleção de modelos de agendamento.

3.2.5.2 Seleção de Modelos de Sistemas de Agendamento para Aplicação no Caso

Uma vez tratados os dados, foi dado andamento à seleção de modelos para aplicação no caso estudado, dentre os elencados no Quadro 8.

Tendo como critérios a compatibilidade com o caso, o potencial de desempenho e o cronograma deste estudo, foi selecionado um modelo para aplicação.

Para efeitos de comparativos de desempenho, foram acrescentados dois modelos, considerados tradicionais pela literatura, também para aplicação no caso.

3.2.5.3 Descrição do Modelo Selecionado

Uma vez selecionado, o modelo foi detalhado quanto à sua metodologia e aos seus resultados, de forma a subsidiar a construção dos comparativos previstos no estudo.

3.2.5.4 Definição de Cenários para Aplicação do Modelo Selecionado

Esta etapa metodológica buscou delimitar as faixas de valores para as variáveis de entrada a serem consideradas para o modelo proposto, de modo a garantir a reprodução do caso.

3.2.5.5 Aplicação do Modelo Selecionado no Caso

Os dados coletados e tratados foram base para a geração de números aleatórios para realização de simulações, abordagem utilizada no modelo selecionado.

Com estes dados, foram montados cenários de modo a representar da melhor maneira possível o contexto do caso analisado.

Uma modelagem matemática foi construída de modo a contemplar as variáveis delineadas, permitindo que a simulação dos cenários gerasse os indicadores especificados para medição de desempenho.

Embora o modelo escolhido contemplasse todas as variáveis previstas para o estudo, as faixas de valores identificadas no caso levantado exigiram simulações adicionais em relação às realizadas no modelo selecionado.

Quanto aos critérios para interpretação das constatações do trabalho, o principal indicador foi definido como sendo o custo total ponderado entre a ociosidade e as horas-extras do prestador e a espera dos clientes, dado pela Equação 1.

$$\text{Mín } E(TC) = E(W)C_p + E(I)C_d + E(O)C_o \quad (1)$$

Os indicadores de custos em função de tempos são os mais frequentemente encontrados na literatura, sendo este o motivo para esta seleção, uma vez que o presente estudo contempla comparativos entre modelos.

Com o objetivo de explorar outros indicadores da literatura e contribuir para proposições futuras, foram medidos os coeficientes de variação dos tempos de espera de clientes para uma amostra de cenários. Estes cenários foram selecionados de modo a representar diferentes faixas de variabilidade de tempos de serviço.

Feitas estas adequações e obtidos os indicadores, foram realizados estudos comparativos para análise de desempenho do modelo proposto no caso estudado, buscando as respostas para as perguntas de pesquisa formuladas.

Para os estudos comparativos, foi adotado o critério do modelo selecionado, o qual propõe confrontá-lo com regras de agendamento consideradas pelos seus autores como sendo tradicionais.

Um comparativo adicional foi apresentado, confrontando as regras consideradas tradicionais pelos autores do modelo selecionado, com o objetivo de analisar o comportamento destas regras frente à diversidade de fatores ambientais.

3.2.5.6 Modelo Proposto pelo Autor

Um último comparativo foi realizado com um modelo de natureza prática proposto pelo autor deste trabalho, com o objetivo de explorar alternativas para sistemas de agendamento.

3.2.6 Compartilhamento

Conforme Yin (2015), a estrutura composicional do presente trabalho pode ser classificada como analítica linear, sendo esta a mais recomendada para teses e dissertações. Este tipo de abordagem tem como início o problema do estudo e a revisão de literatura. Na sequência descrevem-se os métodos empregados, os achados dos dados coletados e analisados e, finalmente as conclusões e suas implicações.

Um esquema contendo os principais tópicos das etapas metodológicas é apresentado na Figura 4.

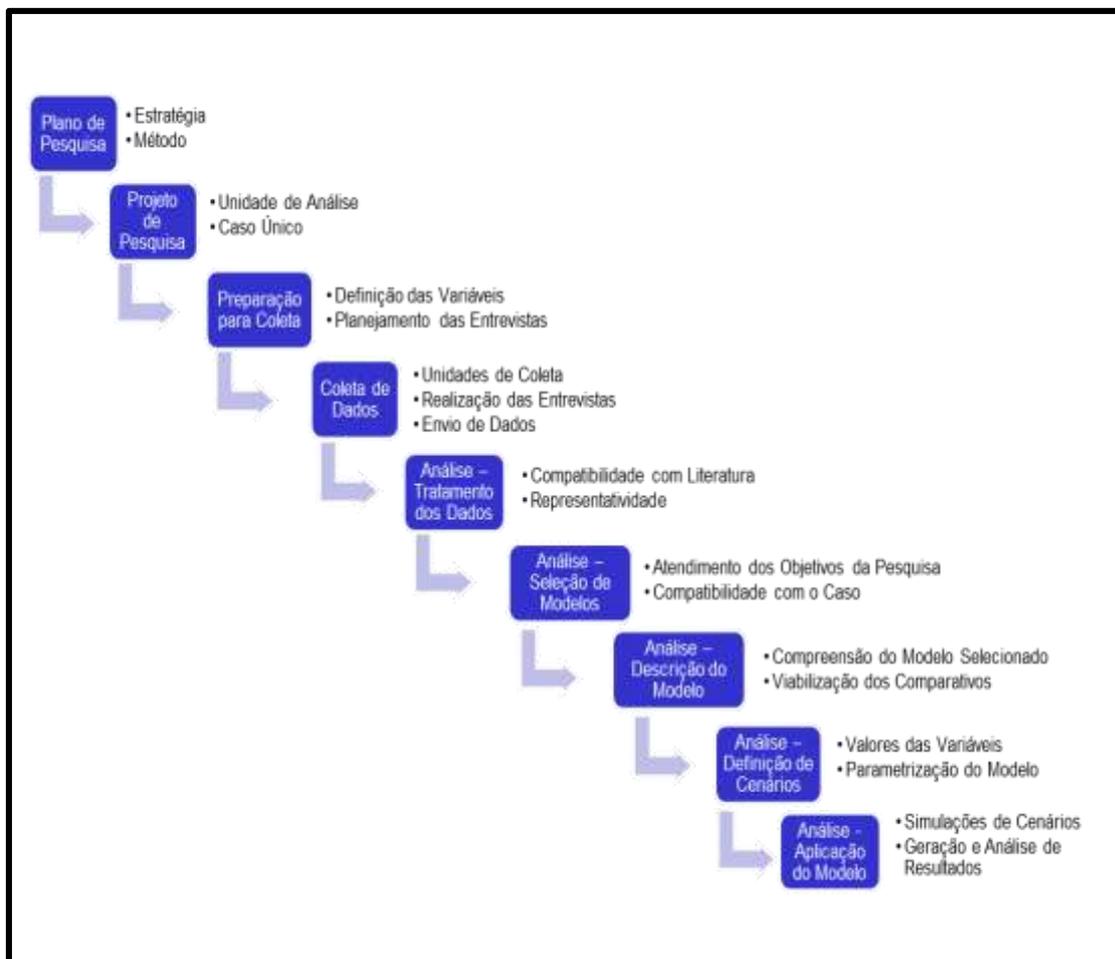


Figura 4 - Etapas metodológicas

4 RESULTADOS

4.1 COLETA DE DADOS

A obtenção de informações deu-se em três etapas, na forma de entrevistas e envio eletrônico de dados.

A primeira etapa consistiu em uma entrevista para entendimento do processo e coleta de dados referentes à geração das sessões de atendimento.

Embora constatada a prática de sessões simultâneas de um mesmo serviço, em uma mesma unidade, o sistema reportado pode ser enquadrado como de servidor único, uma vez que tanto os recursos quanto os clientes são associados a uma sessão específica.

Quanto à sua natureza, o processo corresponde a agendamento estático, dado que todas as decisões de alocação, tanto de recursos como de clientes, são tomadas antes do início da sessão, após o qual permanecem inalteradas.

Os intervalos entre serviços seguem padrões, na sua maior parte, de 15 e 20 minutos, a depender do exame, sendo que as sessões são distribuídas em três períodos do dia. As sessões da manhã e da tarde têm duração de cinco horas, enquanto que no período da noite a duração é de três horas.

Os horários dos exames são definidos pela divisão das sessões de acordo com os padrões de intervalos. Uma agenda para o período da manhã, por exemplo, com duração de cinco horas, planejada para um exame com padrão de intervalo de 15 minutos, contém 20 exames agendados. Conseqüentemente, obtém-se um padrão de número de atendimentos por sessão, por período e por exame.

As sessões agendadas não contam com um código de identificação.

Para determinados exames, aplica-se *overbooking* para a compensação de *no-shows*, mediante a inserção de horários adicionais distribuídos uniformemente ao longo da sessão. A quantidade de horários adicionais é determinada em função da taxa de *no-shows* observada na semana anterior à programada.

Ainda que haja um padrão para os intervalos entre exames, os tempos de serviço e sua variabilidade não são medidos. Da mesma forma, não há monitoramento dos

indicadores de tempos de espera dos pacientes e de ociosidade e horas-extras dos recursos do laboratório.

Os custos unitários destas medidas de tempo também não foram disponibilizados. Obteve-se, entretanto, os preços dos exames realizados para clientes particulares.

A pontualidade dos clientes em relação aos horários agendados para exames é outra variável não monitorada pelo laboratório.

Uma segunda entrevista procurou mapear o processo de preenchimento das agendas à medida em que ocorrem as demandas de clientes. Confirmou-se a natureza estática do sistema de agendamento, uma vez que a alocação de clientes ou alterações em horários de exames em sessões já iniciadas não é regra da organização.

Também não é praticada a classificação de clientes para alocação nas agendas, cabendo a estes a escolha do horário disponível de acordo com suas preferências.

Na terceira etapa da coleta de dados, também mediante entrevista, foi apresentada a especificação dos dados referentes à ocupação das sessões após a realização dos exames, os quais compõem a maior parte das variáveis de entrada para o modelo a ser proposto. Foram disponibilizadas informações em arquivo eletrônico, contendo os exames agendados, bem como elementos para identificar se foram ou não realizados (*no-shows*), abrangendo o período de 01/02/2014 a 09/04/2015. Dada a ausência de um código de identificação das sessões, os dados não permitiram estabelecer uma relação entre exames e as sessões em que foram realizados.

Conforme já mencionado, os dados realizados de tempos de serviço, espera de clientes e ociosidade dos recursos não estavam disponíveis, uma vez que não são monitorados.

4.2 TRATAMENTO DE DADOS

Inicialmente procurou-se reduzir o tamanho da amostra de exames realizados de modo a viabilizar o desempenho do seu tratamento em ambiente MS-Excel. Uma vez que o presente estudo aplica modelagem quantitativa, de acordo com a metodologia especificada, esta ferramenta foi intensamente demandada na geração e análise de dados ao longo do trabalho.

Deste modo, as informações foram limitadas aos exames realizados no ano de 2015, resultando em 230.984 exames agendados distribuídos em 199 tipos de exames. Este

filtro também contribuiu para a utilização de dados mais atualizados, eliminando serviços já não presentes no portfólio.

Foi constatado que alguns serviços são realizados com os mesmos recursos humanos e instalações, sendo agendados em uma mesma sessão, formando grupos de exames. Nestes casos o serviço a ser realizado é conhecido somente quando há o agendamento de um cliente, sem que isto acarrete alterações nos parâmetros da sessão.

Deste modo, os 199 tipos de exames foram agrupados segundo este critério, gerando 130 grupos.

Como delimitação adicional o estudo prosseguiu com os grupos de maior frequência, proporcionando uma representatividade de 90% da amostra. Este novo filtro resultou em 11 grupos, totalizando 207.796 exames agendados, conforme apresentado na Tabela 1.

Tabela 1 – Grupos de Exames de Maior Frequência – 01/01 a 09/04/15

Grupo	Descrição	Exames Agendados
US	Ultrassom	132.282
COVV1	Colposcopia/Vulvoscopia	31.034
MM1	Mamografia	8.135
CA2	Teste Ergométrico	7.226
CA9	Ecocardiograma - Doppler Colorido	6.585
MM3	Mamografia Digital	5.423
CA1	Eletrocardiograma	4.723
DO2	Densitometria - 2 segmentos	4.387
RX50	Raios X - Tórax - P.A. - Lateral - Perfil	3.469
DO1	Densitometria Óssea	2.971
CA13	Monit. Amb. da PA (MAPA) - 24 hs	1.561
Total		207.796

Na sequência procedeu-se a análise do histórico de *no-shows*, segmentando os dados por grupo de exames.

A primeira análise deste parâmetro buscou identificar o tipo de distribuição de probabilidades. Uma vez que um sistema de agendamento foca o desempenho ao longo de uma sessão independente, seria mais adequado tratar amostras separadas

em sessões individuais. Entretanto, a instituição estudada não atribui um código que permita identificar a que sessão cada exame pertence. Deste modo, a análise das taxas de *no-shows* considerou a combinação do grupo de exame com o dia da realização, gerando aproximadamente 100 taxas de *no-shows* para cada grupo.

Em seguida, para cada grupo de exames, foi gerado um histograma das taxas de *no-show*. Com exceção de alguns casos de poucos exames realizados em um dia, para alguns grupos, não configurando uma sessão, observa-se um perfil bastante compatível com uma distribuição binomial. Os Gráficos 1 a 3 ilustram os histogramas com os três grupos de exames com maior representatividade no histórico.

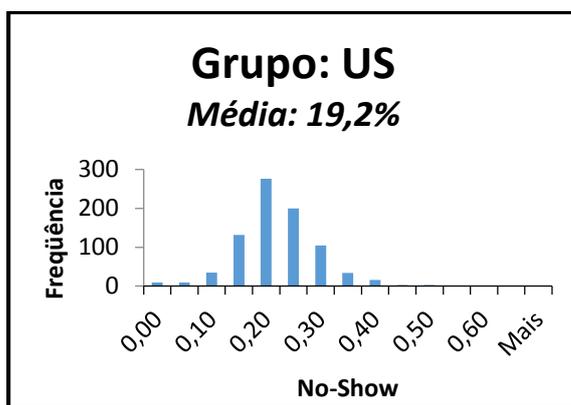


Gráfico 1 - Distribuição de Frequência das Taxas de *No-Show* para o Grupo de Exames US

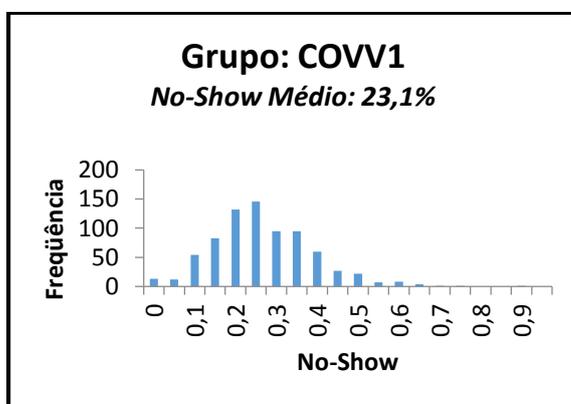


Gráfico 2 - Distribuição de Frequência das Taxas de *No-Show* para o Grupo de Exames COVV1

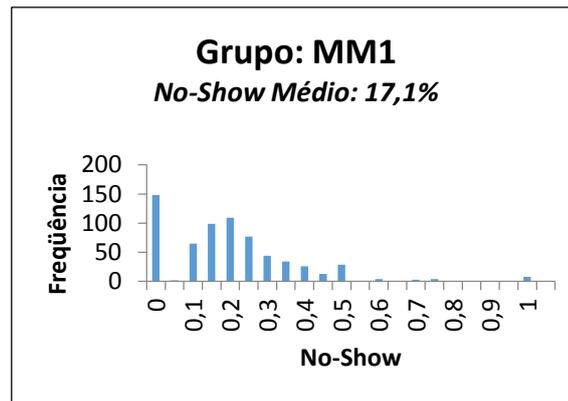


Gráfico 3 - Distribuição de Frequência das Taxas de *No-Show* para o Grupo de Exames MM1

Com o objetivo de refinar os dados e identificar a amplitude das taxas de *no-show* para cada grupo de exames, buscou-se uma maior segmentação dos dados. É relevante mencionar que esta segmentação somente tem efeito se utilizar parâmetros conhecidos antes do preenchimento da sessão, uma vez que a classificação de clientes não é aplicada na instituição. Considerar eventuais taxas diferenciadas por gênero, por exemplo, não permite aplicação prática, uma vez que esta informação não pode balizar os parâmetros de uma sessão, definidos antes de se ter os dados dos clientes.

Assim uma nova análise foi realizada utilizando-se como critério o dia da semana. Conforme pode ser observado na Tabela 2, existem variações significativas de taxas de *no-show* em relação a este critério, com valor mínimo de 10% e máximo de 37%.

Em seguida buscou-se identificar o tamanho das sessões de atendimento, correspondente ao número de pacientes agendados por sessão (parâmetro N). Conforme já mencionado, a instituição pesquisada adota um padrão para os intervalos entre exames para cada grupo. Entretanto, a inserção de horários adicionais (*overbooking*) por conta de *no-shows* não é regular, uma vez que considera a taxa de *no-shows* da semana anterior à sessão agendada. Além disso, esta prática não é adotada para a totalidade dos exames.

Tabela 2 – Taxas de *No-Show* por Grupo de Exames

Grupo	No-Show		
	Médio	Mínimo	Máximo
US	19%	17%	22%
COVV1	23%	20%	32%
MM1	17%	16%	22%
CA2	26%	24%	37%
CA9	23%	20%	26%
MM3	17%	14%	22%
CA1	22%	17%	26%
DO2	12%	10%	14%
RX50	22%	19%	25%
DO1	14%	13%	22%
CA13	20%	18%	31%

Deste modo, os valores de N para cada grupo foram estimados com base nestes padrões, corrigidos pelas taxas de *no-shows* observadas nos dados históricos, aplicados às durações das sessões, conforme o período do dia. Foram identificados os valores mínimo e máximo de N, buscando-se estimar as faixas deste parâmetro para composição dos ambientes do estudo. Para tanto, utilizou-se a Equação 2.

$$N = \frac{R}{(1-Pn)} \quad (2)$$

Sendo R o padrão para número de agendamentos por sessão (sem *overbooking*), e Pn a taxa de *no-shows* verificada. Os valores encontrados para N podem ser observados na Tabela 3, onde constata-se que os valores mínimo e máximo para N são, respectivamente, 11 e 33.

Tabela 3 – Valores de N (Tamanho da Sessão)

Grupo	Padrão - Agendamentos por Sessão			No-Show Mínimo	N Mínimo po Período			No-Show Máximo	N Máximo por Período			N Mínimo do Grupo	N Máximo do Grupo
	Manhã	Tarde	Noite		Manhã	Tarde	Noite		Manhã	Tarde	Noite		
US	20	20	12	17%	24	24	14	22%	26	26	15	14	26
COVV1	20	20	12	20%	25	25	15	32%	29	29	18	15	29
MM1	20	20	12	16%	24	24	14	22%	26	26	15	14	26
CA2	14	14	8,4	24%	18	18	11	37%	22	22	13	11	22
CA9	20	20	12	20%	25	25	15	26%	27	27	16	15	27
MM3	20	20	12	14%	23	23	14	22%	26	26	15	14	26
CA1	20	20	12	17%	24	24	14	26%	27	27	16	14	27
DO2	20	20	12	10%	22	22	13	14%	23	23	14	13	23
RX50	25	25	15	19%	31	31	19	25%	33	33	20	19	33
DO1	20	20	12	13%	23	23	14	22%	26	26	15	14	26
CA13	20	20	12	18%	24	24	15	31%	29	29	17	15	29

Uma última análise foi realizada com o objetivo de mapear as relações de custos CR (Cd/Cp), indicador de desempenho mais frequentemente utilizado nos modelos da literatura. Embora o custo de espera do cliente (Cp) não seja um dado disponível, o custo de ociosidade do servidor (Cd) pode ser considerado um indicativo das faixas de CR esperadas. Na Tabela 4 pode-se verificar a amplitude dos preços dos exames estudados, aos quais considera-se plausível associar a uma proporção, mesmo que não linear, com seus custos. Os valores mínimo e máximo são R\$ 98 e R\$ 3.360, respectivamente, com uma razão de 34,3 entre ambos.

Estes dados foram coletados em 27/07/2015, no *website* da instituição pesquisada.

Tabela 4 – Preços de Exames

Grupo	Preços	
	Mínimo	Máximo
US	R\$ 98,00	R\$ 739,20
COVV1	R\$ 264,00	R\$ 268,00
MM1	R\$ 3.360,00	
CA2	R\$ 288,00	
CA9	R\$ 400,00	
MM3	R\$ 395,00	
CA1	R\$ 160,00	
DO2	R\$ 320,00	
RX50	R\$ 180,00	
DO1	R\$ 260,00	
CA13	R\$ 388,00	
Geral	R\$ 98,00	R\$ 3.360,00
Razão Mínimo/Máximo	34,3	

4.3 SELEÇÃO DE MODELOS DE SISTEMAS DE AGENDAMENTO PARA APLICAÇÃO NO CASO ESTUDADO

Uma vez tratados os dados coletados, obtém-se a seguinte caracterização do caso estudado:

- Amplo portfólio de serviços
- Existência de *no-shows*
- Diversidade significativa de fatores ambientais, constatada pela amplitude das taxas de *no-shows* (P_n) e número de agendamentos por sessão (N)
- Diferentes proporções de custos entre clientes e serviços (CR)
- Ausência de compartilhamento de recursos entre atendimentos agendados e não-agendados (*walk-ins*)
- Um único estágio de serviços
- Arranjo de servidor único
- Agendamento de natureza estática

Devido à indisponibilidade de dados referentes aos horários reais de início dos exames, assumiu-se a premissa de pontualidade dos clientes.

Com base nestes fatores e nos modelos elencados no Quadro 8, deu-se andamento à seleção de modelos compatíveis com o caso e com os objetivos do estudo.

Uma primeira análise descartou o modelo de Vanden Bosch e Dietz (2000), uma vez que se baseia em um algoritmo limitado a seis pacientes por sessão ($N=6$), focando em otimização baseada em classificação de pacientes. Além disso, os intervalos, embora variáveis, são definidos como múltiplos de 10 minutos.

A análise do modelo de Cayirli et al. (2006) gerou uma nova exclusão. Embora o estudo contemple um comparativo entre várias regras de agendamento, os desempenhos são avaliados com base em alternativas de classificação de pacientes, flexibilidade esta não contemplada no caso estudado. Como principal resultado, a pesquisa aponta o impacto de variáveis associadas a pacientes nos indicadores de

vários sistemas de agendamento, sem focar a busca pelo sistema de melhor desempenho.

O modelo de Berg et al. (2014), embora contemple uma única regra, tendo ganhos de desempenho como objetivo, também foi desenvolvido para aplicação em casos em que existe a possibilidade de classificação de pacientes. Esta classificação subsidia a decisão de sequenciamento dos pacientes, decisão esta inexistente no caso estudado. Deste modo, este modelo também foi descartado para aplicação no presente estudo.

Entre os modelos restantes, o de Ho e Lau (1992), Ho e Lau (1999), Ho et al. (1995), apresenta um conjunto de regras com grande abrangência de fatores ambientais, embora a escolha da regra mais adequada para cada combinação de fatores consista em um procedimento de relativa complexidade, dificultando sua aplicação prática.

Já a proposta de Yang et al. (1998) consiste em uma regra única que, utilizando parte dos conceitos do modelo de Ho e Lau (1992), Ho e Lau (1999), Ho et al. (1995), supera seu desempenho em um comparativo realizado com grande diversidade de cenários.

Quanto ao modelo de Cayirli et al. (2012) e Cayirli e Yang (2014), este também obtém resultados superiores ao modelo de Ho e Lau (1992), Ho e Lau (1999), Ho et al. (1995). Da mesma forma, acumula tanto seus conceitos como os desenvolvidos por Yang et al. (1998). Como diferencial o modelo incorpora o formato de domo aos intervalos variáveis, o qual apresenta desempenho superior em vários estudos onde foi aplicado. Também acrescenta uma técnica de ajustes de médias e variâncias de tempos de serviço com o objetivo de incremento de desempenho, além de contar com simulações em uma faixa mais ampla que as testadas por Yang et al. (1998) e a possibilidade de se incluir *walk-ins*.

Conforme a literatura revisada, mesmo os modelos que se propõem a cobrir uma ampla gama de cenários não apresentam melhor desempenho em todas as situações possíveis. Deste modo, seria interessante a aplicação dos três modelos, até aqui elencados, no caso.

Pelo mesmo motivo, a inclusão de regras consideradas tradicionais pela literatura, as quais são de grande facilidade de aplicação prática, também seriam válidas, uma vez que, conforme os fatores ambientais presentes, podem apresentar desempenho satisfatório, tornando sistemas mais complexos menos atrativos.

Entretanto, as limitações inerentes a uma dissertação de mestrado impõem a necessidade de se restringir a análise ao cronograma do trabalho. Deste modo, optou-se por aplicar o modelo de Cayirli et al. (2012) e Cayirli e Yang (2014), dados os argumentos aqui descritos.

Como vantagem adicional, o modelo de Cayirli et al. (2012) e Cayirli e Yang (2014) embute comparativos com duas regras tradicionais. A primeira é a regra de Bailey (1952), a qual considera o agendamento de dois pacientes no primeiro horário da sessão e intervalos fixos equivalentes à média do tempo de serviço. A segunda regra tradicional agenda um paciente por vez, em intervalos fixos equivalentes à média do tempo de serviço. Esta última regra, classificada como de blocos individuais e intervalos fixos, não é proposta por nenhum estudo específico, embora esteja presente em várias pesquisas para fins comparativos.

Deste modo, a análise contará com três modelos a serem comparados: regra de bloco simples e intervalos fixos, regra de Bailey (1952) e o modelo de Cayirli et al. (2012) e Cayirli e Yang (2014).

O estudo de Cayirli e Yang (2014) consiste na proposta de Cayirli et al. (2012) acrescida da possibilidade de classificação de pacientes. Neste caso, mediante parametrização, a classificação pode ser desconsiderada, com os modelos gerando os mesmos resultados. Para efeitos de simplificação, a análise adotará o estudo de Cayirli et al. (2012) como referência.

4.4 DESCRIÇÃO DO MODELO SELECIONADO

O modelo de Cayirli et al. (2012) tem como principais objetivos a flexibilidade e a facilidade de implantação, ao mesmo tempo em que procura atingir um bom desempenho quanto aos indicadores de custos de um sistema de agendamento.

Esta flexibilidade se traduz na busca por uma regra universal, abrangendo uma gama ampla de fatores ambientais e custos relativos entre servidor e cliente. Já a facilidade de implantação diz respeito à possibilidade de se traduzir esta regra universal em uma única equação matemática, simplificando a aplicação prática do modelo.

O trabalho tem como base os estudos de Ho e Lau (1992), Ho e Lau (1999), Ho et al. (1995) e Yang et al. (1998), os quais também contemplam uma gama maior de situações em relação ao restante da literatura pesquisada acerca do tema.

O modelo consiste em um sistema de agendamento com blocos individuais e intervalos variáveis, sendo que a existência de um bloco inicial varia em função do ambiente onde é aplicado.

Utilizando abordagem de simulação, o modelo contempla ajustes para *no-shows* e *walk-ins*. Embora inclua um apêndice com uma adaptação do modelo para classificação de pacientes, este não é o foco do trabalho.

Os ambientes simulados são formados por combinações de N (tamanho da sessão), Cv (coeficiente de variação do tempo de serviço), Pn (probabilidade de *no-shows*) e Pw (probabilidade de *walk-ins*) para diferentes proporções entre custos de servidor e cliente (CR).

O Quadro 12 traz um resumo das variáveis contempladas na descrição do modelo selecionado.

Quadro 12 – Variáveis do Modelo Selecionado

Notação	Descrição
Cv	Coeficiente de Variação do Tempo de Serviço
Pn	Probabilidade de <i>No-Shows</i>
Pw	Probabilidade de <i>Walk-Ins</i>
N	Tamanho da Sessão
CR	Razão entre Custo Espera de Clientes e Ociosidade do Provedor
Cd	Custo Unitário de Ociosidade do Provedor
Cp	Custo Unitário de Espera do Cliente
μ	Tempo Médio de Serviço
μ'	Tempo Médio de Serviço Ajustado
σ	Desvio Padrão do Tempo de Serviço
σ'	Desvio Padrão do Tempo de Serviço Ajustado
T	Objetivo para Número de Pacientes por Sessão
Ai	Horário Agendado para o iésimo Cliente
k	Constante - Cálculo de Ai
π	Constante - Cálculo de Ai
WAIT	Tempo Médio de Espera dos Clientes em uma Sessão
IDLE	Tempo de Ociosidade do Provedor por Cliente Atendido em uma Sessão
OVER	Tempo de Horas-Extras por Cliente Atendido em uma Sessão

4.4.1 Ajuste para *No-Shows* e *Walk-Ins*

O sistema emprega a abordagem de Vissers e Wijngaard (1979), a qual considera os *no-shows* como serviços de duração zero e os *walk-ins* como serviços com duração equivalente a duas vezes o tempo médio. Com base nestas premissas os autores recalculam a média e a variância dos tempos de serviço, obtendo as Equações 3 e 4.

$$\mu' = (1 - Pn + Pw)\mu \quad (3)$$

$$\sigma'^2 = (1 - Pn + Pw)(\sigma^2 + (Pn - Pw)^2\mu^2) + Pn(1 - Pn + Pw)^2\mu^2 + Pw(2\sigma^2 + (1 + Pn - Pw)^2\mu^2) \quad (4)$$

4.4.2 Tamanho da Sessão

O tamanho da sessão é calculado com base em um objetivo (T) para o número de pacientes a serem atendidos e as probabilidades de *no-shows* e *walk-ins*, mediante a Equação 5.

$$N = \frac{T}{(1 - Pn + Pw)} \quad (5)$$

4.4.3 Regra de Agendamento

O estudo propõe uma regra universal, a qual utiliza as Equações 6 e 7.

$$A_i = \text{Máx}\{0, k(i - 1)\mu - \sigma\sqrt{i}.\pi\} \quad (6)$$

$$\text{Para } i = 1, \dots, N \text{ sendo } \pi = \frac{(N+i)}{(N-1)}$$

$$k = f(N, Cv, Pn, Pw, CR) \quad (7)$$

Onde A_i é o horário agendado para o paciente i e μ e σ são, respectivamente, a média e o desvio-padrão dos tempos de serviço. A constante k dá a característica de universalidade da regra, uma vez que, dentro de um mesmo modelo, diferentes valores de k implicam diferentes intervalos entre serviços.

Esta regra é uma versão modificada da regra de Yang et al. (1998), uma vez que introduz o parâmetro π na equação, o qual confere um formato de domo aos intervalos entre serviços. Isto significa que os intervalos são mais curtos no início da sessão de atendimento, aumentando ao longo do período e voltando a serem menores na sua parte final.

Conforme indicado na Equação 6, o valor de k é definido em função dos fatores ambientais N , C_v , P_n e P_w , além da proporção entre os custos de ociosidade do servidor e de espera do paciente (CR).

A Figura 5 traz uma representação gráfica com a evolução dos intervalos entre serviços ao longo de uma sessão, em um determinado ambiente, comparando a regra proposta (regra Domo), para dois valores de k , com a regra de Yang et al. (1998) e a regra de bloco individual e intervalo fixo (IBFI). Nesta última, o intervalo é o tempo médio de serviço.

Pode-se perceber que na regra de Yang os intervalos entre serviços são crescentes ao longo da sessão, enquanto que na regra Domo os intervalos aumentam no início da sessão, com posterior redução em direção ao seu final.

A figura permite ainda observar os efeitos de se alterar o valor do parâmetro k na regra Domo. Para um valor menor de k os intervalos entre serviços são mais curtos, o que é favorável quando o valor relativo do servidor é maior que o do paciente (CR maior), sendo que o oposto também é verdadeiro.

Deste modo, é importante ressaltar que a uma grande vantagem da regra Domo é o fato de esta ser uma função do valor de CR . Regras tradicionais, como a IBFI, são adequadas somente para uma faixa limitada de valores de CR para cada ambiente.

O modelo propõe, ainda, uma alteração na Equação 6 com base no estudo de Vissers e Wijngaard (1979), ajustando os valores de μ e σ , conforme as Equações 3 e 4, gerando a Equação 8.

$$A_i Ajustado = Máx\{0, k(i-1)\mu' - \sigma'\sqrt{i}\cdot\pi\} \quad (8)$$

$$\text{Para } i = 1, \dots, N \text{ sendo } \pi = \frac{(N+i)}{(N-1)}$$

A Equação 7, no entanto, permanece inalterada.

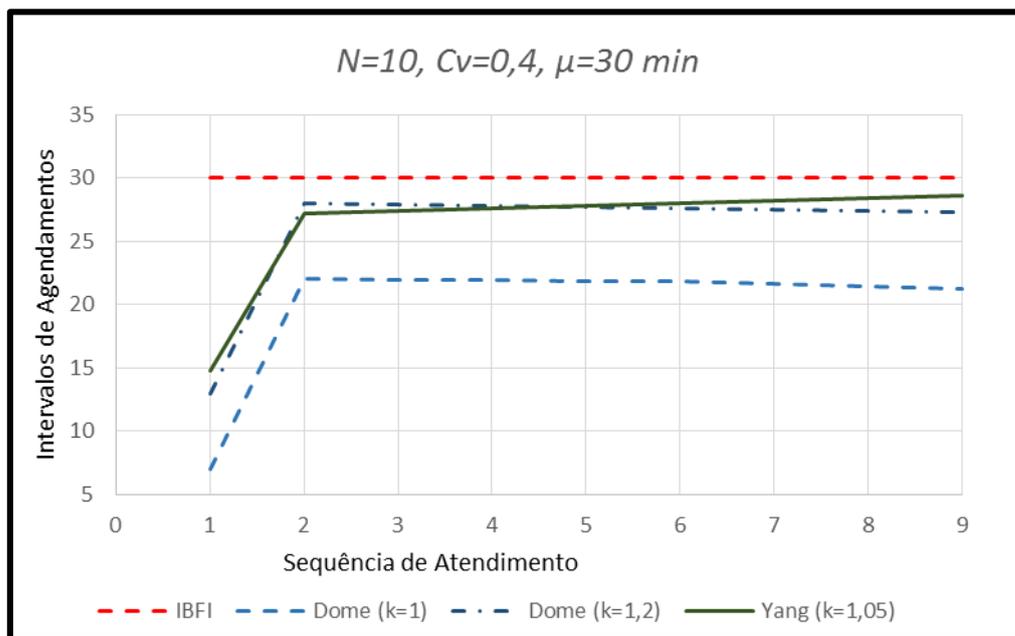


Figura 5 – Comparativo de Intervalos de Serviço entre as Regras IBFI, Yang e Dome

Fonte: Cayirli et al. (2012)

4.4.4 Indicadores de Desempenho

Os indicadores são baseados em tempo médio de espera do paciente (WAIT), tempo de ociosidade do servidor por paciente (IDLE) e tempo médio de horas-extras do servidor, também por paciente (OVER).

- WAIT: inclui o tempo médio de espera tanto dos pacientes agendados quanto não-agendados (*walk-ins*). Para um paciente agendado, é a diferença entre o horário de início do serviço e o horário do agendamento. Para os casos de *walk-ins* o cálculo é feito a partir do horário de chegada, uma vez que para estes casos não há agendamento;

- IDLE: é o tempo em que o servidor permanece ocioso ao longo da sessão, desde o seu início até o término agendado, independentemente do término de atendimento do último paciente, dividido pelo número de pacientes atendidos em toda a sessão;
- OVER: é o tempo requerido para atender pacientes, além do horário agendado para término da sessão, dividido pelo total de pacientes atendidos em toda a sessão.

Estes três indicadores de tempo são, portanto, medidos por paciente atendido.

O equilíbrio entre estes indicadores pode ser estimado subjetivamente pelo decisor com base nos valores relativos dos custos por de espera do paciente (C_p), custo de ociosidade do servidor (C_d) e custo de horas-extras do servidor (C_o). Estes três custos são medidos por unidade de tempo.

O estudo combina C_p e C_o em uma única variável por meio da relação C_o/C_d fixada em 1,5. Segundo os autores da pesquisa, este valor reflete a prática comum quanto à proporção entre horas regulares e horas-extras.

Desta maneira, o custo total esperado (TC) do sistema, por paciente, pode ser representado pela Equação 9, sendo esta a variável a ser minimizada no modelo.

$$TC = (WAIT)C_p + [(IDLE) + 1,5(OVER)]C_d \quad (9)$$

Já a razão CR (C_d/C_p) é simulada na faixa de 1 a 50. O estudo considera este espectro adequado, uma vez que em alguns casos devem ser considerados os custos de toda a instalação, e não somente o tempo dos profissionais envolvidos no atendimento.

4.4.5 Metodologia do Modelo Selecionado

Para obtenção de valores para o parâmetro k , Cayirli et al. (2012) aplicam uma abordagem de simulação. São gerados dados aleatórios representando os ambientes mais comumente encontrados na literatura, variando os parâmetros N , C_v , P_n e P_w , em um sistema composto de um único servidor, em um único estágio de serviço. Para

os tempos de serviço foi considerada uma distribuição lognormal, enquanto que para a chegada de *walk-ins* foi adotou-se a distribuição exponencial.

Os tempos médios de serviço são fixados em 1, sem perda de generalidade, sendo assumida a pontualidade dos pacientes.

O experimento conta com um total de 10.000 sessões para cada ambiente simulado, sendo que, de acordo com os autores, este número proporciona uma acuracidade de 1% em um intervalo de confiança de 95%.

Um total de 900 ambientes são simulados, decorrentes da combinação de 5 valores para Cv, 6 valores para Pn, 6 valores para Pw e 5 valores para N, conforme pode ser verificado na Tabela 5.

Tabela 5 – Parâmetros do Modelo Selecionado

Fonte: Cayirli et al. (2012)

Fator	Valores
Cv	0,2, 0,4, 0,6, 0,8, 1,0
Pn (%)	0, 5, 10, 20, 30, 40
Pw (C%)	0, 5, 10, 20, 30, 40
N	5, 10, 15, 20, 25

Para cada um destes ambientes a regra Dome (Equação 6) é aplicada, com diferentes valores de k, sendo que para cada valor de k são gerados os respectivos valores para os parâmetros WAIT e $IDLE + 1,5.OVER$, compondo uma curva. A Figura 6 ilustra esta curva para os fatores ambientais $N=20$, $Cv=0,6$, $Pn=0,1$ e $Pw=0,4$

No exemplo, a reta que tangencia a curva no ponto onde $k=1,28$, obtida por interpolação, tem uma inclinação na proporção de 10,38 entre WAIT e $IDLE+1,5.OVER$. Este é, portanto, o valor de CR (Cd/Cp) que caracteriza esta reta. Em qualquer ponto desta linha a soma dos custos, representada pela Equação 9, é constante para um $CR=10,38$. Por este motivo esta reta é denominada linha de isocusto.

É importante ressaltar que qualquer linha com esta inclinação tem o mesmo valor de CR, sendo que o objetivo do sistema de agendamento é atingir uma linha de isocusto o mais próximo possível da origem de modo a obter, para um mesmo valor de CR, a menor soma de custos (TC). No exemplo da Figura 6, o valor de $k=1,28$ configura o ponto onde se tangencia a reta de inclinação 10,38 que está mais próxima da origem.

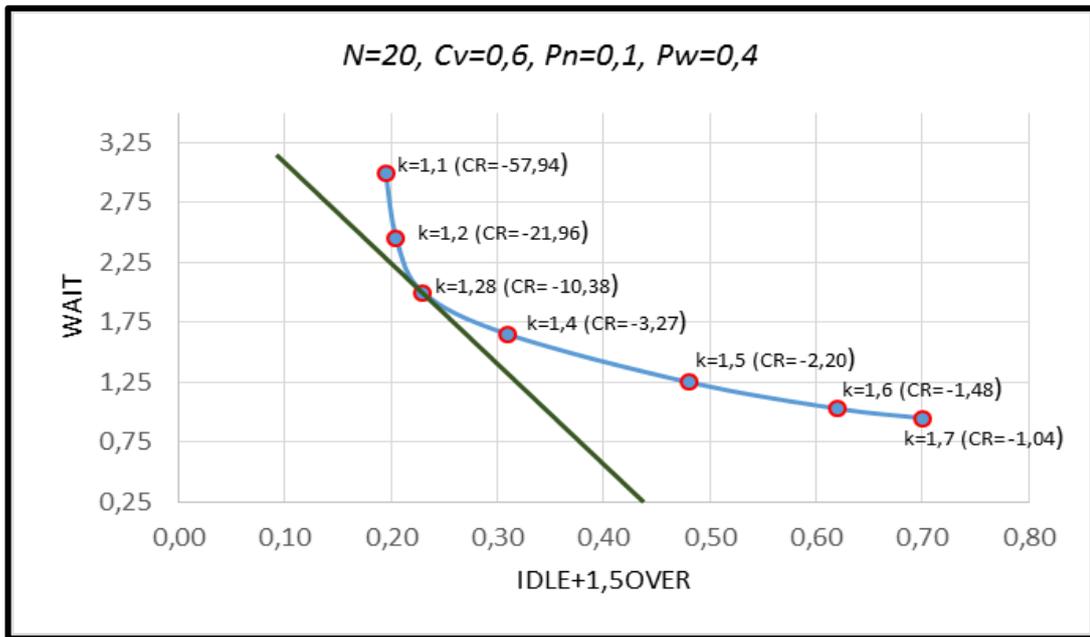


Figura 6 - Valores da Constante k e do Parâmetro CR

Fonte: Cayirli et al. (2012)

Este processo é repetido para vários valores de k, de modo a se cobrir a faixa de CR de 1 a 50 proposta no estudo. A curva então formada por estes pontos é denominada fronteira da eficiência, sendo que cada um dos 900 ambientes simulados tem a sua fronteira específica.

Uma vez obtidos os valores do parâmetro k para os diferentes ambientes (combinações de N, Cv, Pn e Pw) e valores de CR, o estudo utiliza a técnica de regressão não-linear ACE (*Alternating Conditional Expectation*) para obter uma fórmula (Equação 10) que estima o valor de k em função destas cinco variáveis.

$$k = \left\{ 0,9973 - 0,103 \left[0,005765CR(1 - Pn) + (CR(1 - Pn))^{-0,3481} \right] - 0,10699[Cv^{1,257}] - 0,627 \left[(N(1 - Pn))^{-0,8579} \right] - 0,007574 \left[(|CR(1 - Pw) - 2,143|)^{0,9682} - 0,622CR(1 - Pw) \right] + 0,004855(CR^{0,8913}) \right\}^{-1,898} \quad (10)$$

Com este método de regressão foi possível obter um coeficiente de determinação (R^2) de 0,9529, significativamente superior ao obtido com o Método dos Mínimos Quadrados, com o qual obteve-se um coeficiente de apenas 0,6891.

O uso desta equação faz parte do objetivo de simplificar a aplicação prática do modelo.

4.4.6 Resultados do Modelo Selecionado

O desempenho do modelo proposto é confrontado com três regras de agendamento denominadas tradicionais, como segue:

- 1) IBFI: formada por blocos individuais e intervalos fixos, iguais à média do tempo de serviço. Nesta regra os pacientes são agendados em horários individuais, em um mesmo intervalo entre atendimentos. Os horários de agendamento seguem a Equação 11.

$$A_1 = 0; \quad \text{para } i > 1, A_i = A_{i-1} + \mu \quad (11)$$

- 2) 2BEG (regra de Bailey): proposta por Bailey (1952), similar à regra IBFI, porém com dois pacientes agendados para o início da sessão, conforme a Equação 12.

$$A_1 = A_2 = 0; \quad \text{para } i > 2, A_i = A_{i-1} + \mu \quad (12)$$

- 3) OFFSET: é a regra de melhor desempenho do estudo de Ho e Lau (1992) e segue a Equação 13, com os parâmetros $\beta_1=0,15$, $\beta_2=0,30$ e $k_1=5$.

$$\begin{aligned} A_i &= (i - 1)\mu - \beta_1(k_1 - i)\sigma & \text{para } i \leq k_1, \\ A_i &= (i - 1)\mu + \beta_2(i - k_1)\sigma & \text{para } i > k_1 \end{aligned} \quad (13)$$

4.4.6.1 Comparativo com e sem Ajuste para *No-Shows* e *Walk-Ins*

Um primeiro estudo comparativo verifica os efeitos dos ajustes da média e do desvio-padrão do tempo de serviço em função de *no-shows* e *walk-ins* propostos na Equação 8, em relação a uma agenda sem ajustes (Equação 6). Este comparativo é realizado tanto para a regra proposta (Dome) como para as regras tradicionais. Para estas, o ajuste é realizado nas médias e no desvio-padrão (quando utilizado), conforme as Equações 11 a 13.

A Figura 7 traz um exemplo deste comparativo, para um ambiente com $N=13$, $Cv=0,2$, $Pn=0,4$ e $Pw=0,2$. Pode-se verificar que para a regra Dome o ajuste proporciona uma fronteira da eficiência mais próxima da origem quando o ajuste da Equação 8 é aplicado. Isto se traduz em retas de isocusto mais próximas da origem, tornando seu desempenho melhor, à medida que, para um mesmo valor de CR, obtém-se um custo total (TC) mais baixo.

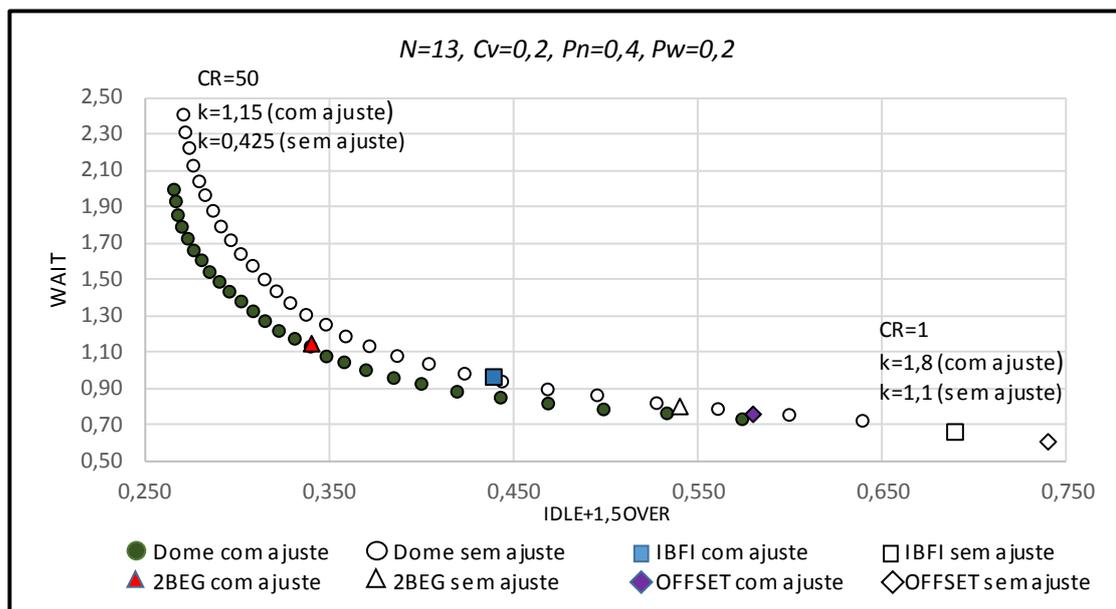


Figura 7 - Comparativo entre Regras de Agendamento com e sem Ajustes

Fonte: Cayirli et al. (2012)

Quanto às regras tradicionais, seus indicadores não são uma função da variável CR, portanto estas geram apenas pontos no gráfico, e não uma fronteira de eficiência. Deste modo, o ajuste analisado provoca apenas um deslocamento deste ponto,

tornando a regra mais favorável para um faixa diferente de valores de CR. No exemplo da Figura 7 temos $P_n > P_w$, o que leva a um ajuste que reduz os valores de μ e σ nas regras tradicionais e, portanto, reduz o intervalo entre serviços. Isto favorece o servidor, reduzindo a probabilidade de ociosidade, o que é mais interessante somente para valores maiores de CR.

Esta constatação evidencia a vantagem do modelo proposto, descrita no estudo, de se obter uma regra única atendendo um espectro amplo de situações, caracterizando-a como “universal”, diferentemente da maioria das demais regras, que possuem um desempenho satisfatório somente para uma determinada faixa de valores de CR.

Uma análise adicional compara os efeitos de diferentes valores de P_n e P_w nos indicadores para a regra Dome. Conforme o exemplo da Figura 8, nota-se que a incerteza proveniente de probabilidades mais altas de *no-shows* (P_n) e *walk-ins* (P_w) tem efeito negativo no custo total (TC), uma vez que provoca um afastamento das fronteiras de eficiência em relação à origem. Novamente a vantagem da regra com ajustes pode ser observada.

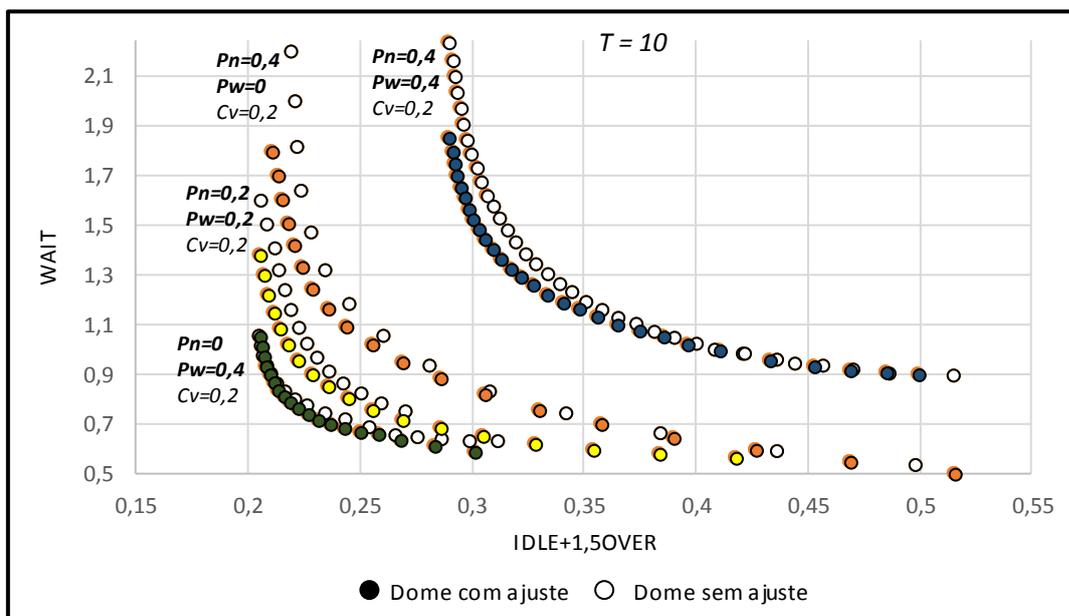


Figura 8 - Impactos de Diferentes Valores de P_n e P_w

Fonte: Cayirli et al. (2012)

Os mesmos impactos podem ser notados quando do aumento no C_v , conforme exemplo da Figura 9. Pode-se observar, ainda, que quanto menor o C_v , maiores os benefícios do ajuste.

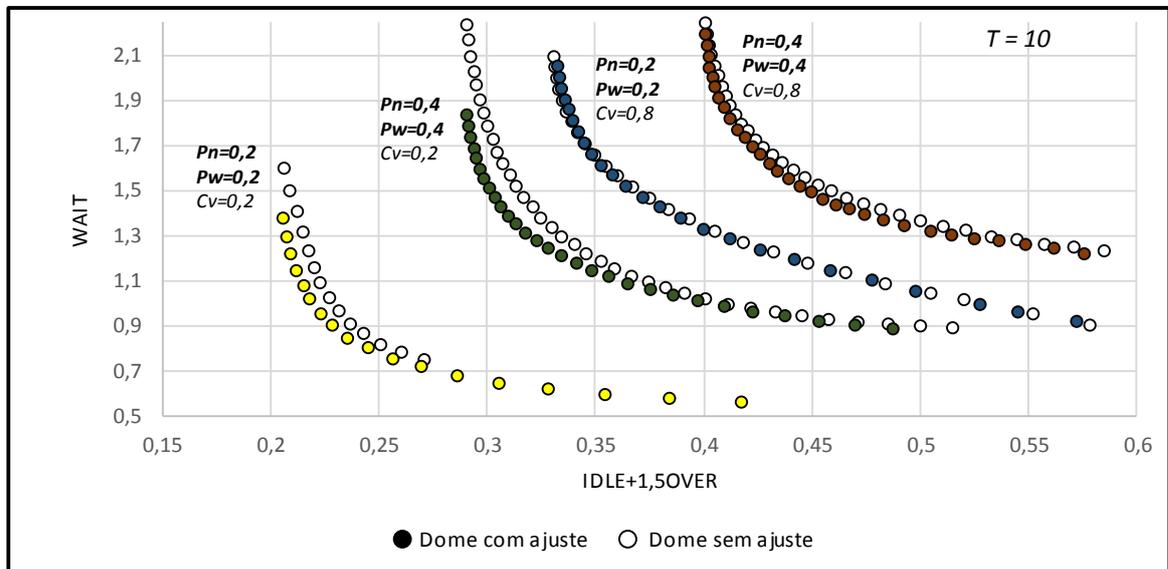


Figura 9 - Impactos de Diferentes Valores de C_v

Fonte: Cayirli et al. (2012)

A Tabela 6 contém um resumo dos efeitos do procedimento de ajuste para uma simulação em 18 ambientes, para três valores de CR.

Observa-se que de modo geral o ajuste traz melhores resultados, uma vez que em somente três dos 54 casos houve aumento do valor de TC.

No entanto, não se pode obter uma conclusão acerca dos efeitos do ajuste em relação ao CR, uma vez que em alguns ambientes obtêm-se melhores resultados com um valor baixo de CR, enquanto que o oposto ocorre em outros casos.

Da mesma forma o parâmetro N não aponta para uma correlação entre o seu valor e os resultados obtidos com e sem ajuste.

Tabela 6 – Ganhos de Desempenho Decorrentes do Ajuste para *No-Shows* e *Walk-Ins*

Fonte: Cayirli et al. (2012)

Ambiente	T	Cv	Pn	Pw	N	μ'	σ'	TC (CR=1)			TC (CR=5)			TC (CR=50)		
								Sem Ajuste	Com Ajuste	Diferença (%)	Sem Ajuste	Com Ajuste	Diferença (%)	Sem Ajuste	Com Ajuste	Diferença (%)
1	10	0,4	0,05	0	7	1,40	0,5441	0,8990	0,8866	1,38	1,9030	1,8844	0,98	11,0835	11,0692	0,13
2	10	0,4	0,05	0,2	17	0,60	0,5138	1,0527	1,0421	1,01	2,4853	2,3848	4,04	13,5219	12,9605	4,15
3	10	0,4	0,05	0,4	10	1,00	0,9165	1,3582	1,3433	1,10	2,9787	2,8477	4,40	16,9297	16,5591	2,19
4	10	0,4	0,2	0	13	0,80	0,7694	1,2918	1,2800	0,92	2,9193	2,7888	4,47	16,0155	15,4705	3,40
5	10	0,4	0,2	0,2	10	1,00	0,6633	0,9775	0,9736	0,40	2,1523	2,1194	1,53	12,5967	12,1544	3,51
6	10	0,4	0,2	0,4	8	1,20	0,4561	0,6431	0,6370	0,94	1,4386	1,4265	0,84	8,7347	8,7212	0,15
7	10	0,4	0,4	0	7	1,40	1,0658	1,3911	1,4041	(0,93)	3,0541	3,0486	0,18	18,2184	18,2011	0,10
8	10	0,4	0,4	0,2	17	0,60	0,7899	1,5423	1,5414	0,06	3,3614	3,3600	0,04	18,6658	18,6644	0,01
9	10	0,4	0,4	0,4	10	1,00	1,2000	1,7664	1,7592	0,41	3,7748	3,7461	0,76	22,2775	22,2340	0,20
10	10	0,6	0,05	0	13	0,80	1,0354	1,7408	1,7359	0,28	3,7107	3,6901	0,55	20,8711	10,8188	0,25
11	10	0,6	0,05	0,2	10	1,00	1,0198	1,4874	1,4987	(0,76)	3,2275	3,2182	0,29	18,7256	18,6883	0,20
12	10	0,6	0,05	0,4	8	1,20	0,9633	1,2585	1,2426	1,27	2,7647	2,7787	(0,50)	16,6421	16,6292	0,08
13	20	0,6	0,2	0	14	1,40	0,5441	1,0796	1,0640	1,45	1,9984	1,9525	2,30	8,9379	8,7941	1,61
14	20	0,6	0,2	0,2	33	0,60	0,5138	1,0817	1,0766	0,47	2,4243	2,3467	3,20	10,8542	10,3031	5,08
15	20	0,6	0,2	0,4	20	1,00	0,9165	1,5905	1,5765	0,88	3,1230	3,0030	3,84	13,8682	13,2479	4,47
16	20	0,6	0,4	0	25	0,80	0,7694	1,3756	1,3659	0,71	2,8858	2,7735	3,89	12,9031	12,2517	5,05
17	20	0,6	0,4	0,2	20	1,00	0,6633	1,1027	1,1011	0,15	2,2698	2,2322	1,65	10,3209	9,8796	4,28
18	20	0,6	0,4	0,4	17	1,20	0,4561	0,8916	0,8846	0,78	1,7187	1,6953	1,36	7,3337	7,2183	1,57

4.4.6.2 Comparativos com Regras Tradicionais

Tendo em vista o melhor desempenho geral do modelo com ajustes, o estudo faz então dois comparativos com as regras tradicionais, também com ajustes, sendo o experimento restrito a faixas de valores para N, Cv, Pn e Pw, conforme apontado na Tabela 7. O comparativo delimita, ainda, os valores 2, 5 e 10 para o CR, gerando assim 162 situações diferentes. Segundo os autores, estes valores de CR são adequados para aplicação prática, uma vez que para valores de CR acima de um determinado valor, usualmente $CR > 10$, ganhos marginais em custos de ociosidade e horas-extras implicam grande aumento no custo de espera dos pacientes.

A regra OFFSET é excluída dos comparativos por ser interessante somente em casos em que o CR é inferior a 1, situação considerada rara pelos autores, uma vez que implica custos relacionados ao paciente como superiores aos atribuídos ao servidor.

No primeiro comparativo, a regra Dome é denominada Dome1 e tem os valores de k calculados com base nas três inclinações de retas (valores de CR) delimitadas para a simulação em cada ambiente. A precisão do valor de k é fixada em duas casas decimais. Como exemplo, na Figura 7 o valor de k com duas casas decimais que mais se aproxima de um valor de CR de 10 seria 1,28, o que resulta em um valor de CR de 10,38.

Tabela 7 – Parâmetros para Comparativo entre Regras Dome e Tradicionais

Fonte: Cayirli et al. (2012)

Fator	Valores
Cv	0,4, 0,6, 0,8
Pn (%)	5, 20, 40
Pw (C%)	0, 20, 40
N	10, 20

Os resultados do experimento podem ser conferidos na Tabela 8.

Tabela 8 – Ganhos de Desempenho da Regra Dome 1 em Relação à Melhor Regra Tradicional

Fonte: Cayirli et al. (2012)

Ambiente	Cv	Pn	Pw	N=10			N=20		
				TC (CR=2)	TC (CR=5)	TC (CR=10)	TC (CR=2)	TC (CR=5)	TC (CR=10)
1	0,4	0,05	0	1,11	6,76	4,04	8,24	4,42	2,21
2	0,4	0,05	0,2	2,40	4,48	0,58	7,97	4,35	2BEG (0,17)
3	0,4	0,05	0,4	3,67	2,02	2BEG (0,1)	9,76	2,61	2BEG (0,88)
4	0,4	0,2	0	1,77	1,36	2BEG (0,82)	6,26	2,13	2BEG (0,9)
5	0,4	0,2	0,2	2,14	0,26	0,11	6,64	1,23	2BEG (0,37)
6	0,4	0,2	0,4	3,08	2BEG (0,09)	0,37	8,14	1,03	2BEG (0,35)
7	0,4	0,4	0	1,58	0,60	3,40	4,69	0,38	2,89
8	0,4	0,4	0,2	1,67	0,70	3,61	5,03	0,61	2,54
9	0,4	0,4	0,4	1,52	0,84	3,91	5,74	0,56	2,02
10	0,6	0,05	0	1,21	3,97	0,31	7,04	4,08	2BEG (0,16)
11	0,6	0,05	0,2	2,48	1,66	2BEG (0,08)	8,19	2,46	2BEG (0,4)
12	0,6	0,05	0,4	3,53	0,73	0,05	9,92	1,94	2BEG (0,52)
13	0,6	0,2	0	2,09	0,21	0,34	5,32	0,96	0,16
14	0,6	0,2	0,2	2,62	0,00	1,15	6,49	0,82	0,42
15	0,6	0,2	0,4	3,23	2BEG (0,08)	1,39	8,10	0,92	0,38
16	0,6	0,4	0	0,99	1,51	5,01	3,68	0,93	4,29
17	0,6	0,4	0,2	1,28	1,41	4,80	4,40	0,95	3,49
18	0,6	0,4	0,4	1,35	1,36	4,87	5,23	0,81	2,77
19	0,8	0,05	0	0,95	0,84	0,01	6,05	1,82	2BEG (0,06)
20	0,8	0,05	0,2	2,33	0,39	0,38	7,85	1,57	0,05
21	0,8	0,05	0,4	3,34	0,13	0,63	9,61	1,55	2BEG (0,03)
22	0,8	0,2	0	2,38	0,04	1,54	4,77	0,63	1,20
23	0,8	0,2	0,2	2,86	0,17	2,11	6,03	0,78	1,22
24	0,8	0,2	0,4	3,05	0,16	2,26	7,70	0,95	1,13
25	0,8	0,4	0	0,66	2,33	6,25	3,02	1,54	5,49
26	0,8	0,4	0,2	1,06	2,07	5,76	3,86	1,36	4,38
27	0,8	0,4	0,4	1,28	1,85	5,65	4,66	1,04	3,54
Mín				0,66	(0,09)	(0,82)	3,02	0,38	(0,90)
Máx				3,67	6,76	6,25	9,92	4,42	5,49
Média				2,06	1,32	2,13	6,46	1,57	1,27

Dos 162 casos analisados, em apenas 11 a regra Dome1 é superada por uma regra tradicional, no caso a regra 2BEG, sendo que as diferenças são inferiores a 1%.

O segundo comparativo considera os mesmos cenários, porém a regra Dome é denominada Dome2 e tem a constante k obtida por meio da Equação 10. Os resultados deste comparativo podem ser observados na Tabela 9.

Tabela 9 – Ganhos de Desempenho da Regra Dome 2 em Relação à Melhor Regra Tradicional

Fonte: Cayirli et al. (2012)

Ambiente	Cv	Pn	Pw	N=10			N=20		
				TC (CR=2)	TC (CR=5)	TC (CR=10)	TC (CR=2)	TC (CR=5)	TC (CR=10)
1	0,4	0,05	0	IBFI (13,59)	IBFI (4,74)	2BEG (2,29)	2,95	IBFI (5,62)	2BEG (4,91)
2	0,4	0,05	0,2	IBFI (0,97)	4,22	0,37	7,77	4,15	2BEG (0,19)
3	0,4	0,05	0,4	IBFI (0,96)	1,69	2BEG (0,89)	9,71	2,61	2BEG (0,9)
4	0,4	0,2	0	IBFI (0,62)	1,20	2BEG (0,95)	6,25	2,06	2BEG (0,88)
5	0,4	0,2	0,2	2,12	2BEG (0,14)	2BEG (0,01)	5,91	0,01	2BEG (1,2)
6	0,4	0,2	0,4	2,86	2BEG (0,13)	0,34	7,29	0,11	2BEG (0,63)
7	0,4	0,4	0	1,42	0,60	3,36	4,49	2BEG (0,08)	2,71
8	0,4	0,4	0,2	1,46	0,29	3,52	3,59	2BEG (1,53)	1,70
9	0,4	0,4	0,4	1,42	0,84	3,78	4,81	2BEG (0,6)	1,84
10	0,6	0,05	0	IBFI (1,44)	3,41	2BEG (0,09)	6,01	1,35	2BEG (1,66)
11	0,6	0,05	0,2	1,80	1,66	2BEG (0,12)	8,04	2,37	2BEG (0,45)
12	0,6	0,05	0,4	2,10	0,67	2BEG (0,4)	9,88	1,90	2BEG (0,64)
13	0,6	0,2	0	2,07	0,17	0,34	5,33	0,96	0,16
14	0,6	0,2	0,2	2,57	2BEG (0,23)	1,09	6,17	0,39	0,22
15	0,6	0,2	0,4	3,00	2BEG (0,09)	1,21	7,53	0,62	0,39
16	0,6	0,4	0	0,98	1,50	5,00	3,64	0,70	4,25
17	0,6	0,4	0,2	1,26	1,35	4,80	3,99	0,13	3,38
18	0,6	0,4	0,4	0,91	1,26	4,37	4,99	0,47	2,77
19	0,8	0,05	0	0,94	0,71	2BEG (0,03)	5,81	1,19	2BEG (0,51)
20	0,8	0,05	0,2	2,30	0,26	0,38	7,67	1,50	2BEG (0,05)
21	0,8	0,05	0,4	2,46	0,09	0,30	9,51	1,46	2BEG (0,29)
22	0,8	0,2	0	2,06	2BEG (0,47)	1,37	4,78	0,63	1,21
23	0,8	0,2	0,2	2,80	2BEG (0,07)	2,09	6,00	0,66	1,21
24	0,8	0,2	0,4	2,77	0,16	2,06	7,55	0,93	1,03
25	0,8	0,4	0	0,64	2,23	6,25	3,03	1,42	5,49
26	0,8	0,4	0,2	1,03	2,05	5,71	3,86	1,15	4,38
27	0,8	0,4	0,4	0,29	1,48	4,73	4,64	1,03	3,23
Mín				(13,59)	(4,74)	(2,29)	2,95	(5,62)	(4,91)
Máx				3,00	4,22	6,25	9,88	4,15	5,49
Média				0,92	0,74	1,71	5,97	0,74	0,80
Mín*				(1,44)	(0,47)	(0,95)	3,03	(1,53)	(1,66)
Máx*				3,00	4,22	6,25	9,88	4,15	5,49

Observa-se que a conveniência de se obter um valor de k de modo mais prático traz, por outro lado, alguma perda no desempenho do modelo, dada a acuracidade da regressão que gera a equação para o seu cálculo.

Entretanto, excetuando-se o ambiente 1, em que a regra Dome2 tem desempenho expressivamente inferior, os seus resultados são melhores na maioria dos casos, apresentando números inferiores entre 1% e 1,66% em somente 4 das 156 situações analisadas.

Em havendo dificuldade por parte do usuário para o cálculo do parâmetro k (regra Dome1), os autores recomendam a aplicação das regras tradicionais para o ambiente 1.

Por fim, o estudo conclui que as regras propostas, em especial a Dome2, cumprem os objetivos especificados de oferecer flexibilidade e facilidade de implantação aliadas a um bom desempenho em custos.

4.5 DEFINIÇÃO DE CENÁRIOS PARA APLICAÇÃO DO MODELO SELECIONADO

Os fatores ambientais encontrados no caso, conforme tratamento dos dados coletados foram resumidos na Tabela 10

Tabela 10 – Fatores Ambientais do Caso Estudado

Fator	Valores
Cv	Não Disponíveis
Pn (%)	10 a 37
Pw (%)	Não se aplica
N	11 a 33

Confrontando-se estes dados com os parâmetros do experimento de Cayirli et al. (2012), constantes na Tabela 7, observa-se que a probabilidade de *no-shows* do caso se enquadra na faixa do modelo selecionado. A faixa superior do parâmetro N, no entanto, excede a da pesquisa de Cayirli et al. (2012).

Para o coeficiente de variação do tempo de serviço (Cv) não há informações disponíveis para o estudo, enquanto que o parâmetro Pw não se aplica ao caso, dada a ausência de *walk-ins*.

Quanto ao parâmetro CR, o experimento considera os valores de 2, 5 e 10, enquanto que para o caso estudado não há valores disponíveis.

Para definição dos parâmetros deste estudo, foram adotados os seguintes critérios e decisões:

- Coeficiente de variação do tempo de serviço (C_v): em virtude do portfólio diversificado de serviços da instituição estudada, decidiu-se por manter a amplitude de valores proposta no modelo selecionado. Vale ressaltar que seus autores utilizaram como critério os valores mais comumente encontrados na literatura
- Probabilidade de *No-Shows* (P_n): foi mantida a faixa de valores propostos por Cayirli et al. (2012), mais ampla que a do caso estudado, uma vez que ações para mitigação do *no-show* podem levar à redução de suas taxas. Além disso, uma segmentação mais refinada das taxas, como por período do dia, tende a ampliar esta faixa
- Tamanho da sessão (N): como o caso estudado abrange uma faixa expressivamente mais ampla de valores, tornou-se necessária a introdução de um novo limite superior. Como este número depende de P_n , conforme a Equação 2, e os valores de P_n podem ser refinados, conforme já mencionado, optou-se por adotar os valores 10, 20 e 40 para este parâmetro
- Relação de Custos entre Servidor e Cliente (CR): tendo em vista a significativa amplitude de preços de exames, conforme apontado na Tabela 4, optou-se por manter os valores de 2, 5 e 10, propostos por Cayirli et al. (2012)

A Tabela 11 resume os parâmetros definidos para o experimento.

Quanto aos custos de horas-extras (C_o), dada a indisponibilidade de dados, manteve-se o critério adotado no modelo proposto, o qual fixa seu valor como 1,5 vezes o custo de ociosidade do servidor (C_d).

4.6 APLICAÇÃO DO MODELO SELECIONADO AO CASO

Para os comparativos realizados com o modelo selecionado foram consideradas as regras IBFI, 2BEG e Dome 1. Um comparativo com a regra Dome 2 exigiria a revisão da Equação 10, mediante uma nova análise de regressão, excluindo a variável P_w (probabilidade de *walk-ins*) e estendendo a faixa de N para 40. Entretanto, esta

abordagem tem como meta propor um método ágil e simplificado para a construção de agendas, proporcionando conveniência em detrimento de desempenho, o que não está dentro dos objetivos da presente pesquisa.

O MS-Excel 2013 foi a ferramenta empregada para o desenvolvimento das simulações constantes neste tópico, cujos detalhes estão contemplados no Anexo 1.

Tabela 11 – Parâmetros para o Experimento

Parâmetro	Valores
Cv	0,4 - 0,6 - 0,8
Pn (%)	5 - 20 - 40
Pw	0
N	10 - 20 - 40
CR	2 - 5 - 10

4.6.1 Formulação

Os horários de agendamento dos clientes em uma sessão seguem as Equações 6 (sem ajustes para *no-shows* e *walk-ins*) e 8 (com ajustes) para a regra Dome 1. Para as regras IBFI e 2BEG são mantidas as Equações 11 e 12, respectivamente.

Para o cálculo dos inícios e térmios de cada serviço, o modelo selecionado baseia-se na formulação descrita no estudo de Ho e Lau (1992), conforme apontado nas Equações 14 e 15.

Considera-se $b_1 = A_1$. Para $i > 1$ ($i = 1, 2, \dots, N$), tem-se:

$$b_i = \text{Máx}(A_i, e_{i-1}) \quad (14)$$

$$e_i = b_i + t_i \quad (15)$$

Sendo:

A_i = horário agendado para o cliente i

t_i = tempo de serviço para o cliente i

b_i = horário de início do serviço para o cliente i

e_i = horário de término do serviço para o cliente i

Para os tempos de espera dos clientes e de ociosidade do servidor, também foram aplicadas as fórmulas de Ho e Lau (1992). As Equações 16 e 17 representam esta formulação:

$$P_i = \text{Máx}(0, b_i - A_i) \quad (16)$$

$$M_i = \text{Máx}(0, A_i - e_{i-1}) \quad (17)$$

Onde:

P_i = tempo de espera do cliente i

M_i = tempo de ociosidade do servidor durante o tempo agendado para atendimento do cliente $i-1$

Um tempo ocioso adicional para o servidor deve ser considerado caso o término de atendimento do último cliente ocorra antes do término programado da sessão. Este cálculo é apresentado na Equação 18.

$$E = \text{Máx}(0, S - e_N) \quad (17)$$

Onde:

E = tempo de ociosidade do servidor após término do serviço do cliente N

S = horário programado para término da sessão

O valor de S é obtido com base no número de clientes agendados e o tempo médio de serviço, conforme Equação 18.

$$S = T \cdot \mu \quad (18)$$

Sendo que o valor de T é obtido mediante a Equação 3. Dada a inexistência de *walk-ins* no caso estudado, o valor de Pw pode ser considerado como 0, tanto nesta equação quanto nas demais que contenham este parâmetro.

Como o modelo adotado considera tempos médios por paciente atendido, os indicadores WAIT, IDLE e OVER são calculados conforme as Equações 19 a 21.

$$WAIT = \frac{\sum_{i=1}^N P_i}{N'} \quad (19)$$

$$IDLE = \frac{\sum_{i=1}^N M_i + E}{N'} \quad (20)$$

$$OVER = \frac{\max(0; e_N - S)}{N'} \quad (21)$$

Onde N' representa o número de clientes atendidos em uma sessão, calculado pela Equação 22.

$$N' = \sum C_i \quad (i \text{ de } 1 \text{ a } N) \quad (22)$$

Sendo que Ci corresponde a um vetor binário, o qual assume valor 0, para o caso de *no-show*, ou 1 quando do comparecimento do cliente para o exame agendado.

O cálculo do tempo de serviço ti é obtido por meio da Equação 23.

$$t_i = s_i \cdot C_i \quad (23)$$

Onde si representa o tempo simulado para o atendimento do cliente i.

O cálculo do custo total (TC) teve por base a Equação 9. Para efeitos de simplificação dos cálculos e comparativos, o custo de espera dos clientes (C_p) foi fixado em 1, sem perda de generalidade. Deste modo, o custo de ociosidade do servidor (C_d) equivale a CR, uma vez que CR representa a razão C_d/C_p .

Com estas premissas, a Equação 7 gerou a Equação 24.

$$TC = WAIT + [(IDLE) + 1,5(OVER)]CR \quad (24)$$

4.6.2 Geração de dados para simulação

Dois tipos de dados foram gerados aleatoriamente para realização da pesquisa: tempos de serviço simulados (s_i) e comparecimento dos clientes (C_i).

Para os tempos de serviço simulados (s_i) foi mantida a distribuição lognormal, utilizada no modelo de Cayirli et al. (2012), e empregada a técnica de transformação inversa, utilizando o aplicativo MS-Excel 2013. Também foi mantida a média para o tempo de serviço (μ) = 1, sem perda de generalidade.

Com estas premissas foram simulados nove conjuntos de 10.000 sessões, para as diferentes combinações de C_v e N.

Para o comparecimento dos clientes agendados (C_i) foi empregado o recurso de geração de dados aleatórios do MS-Excel 2013, considerando distribuição binomial, conforme identificado no tratamento dos dados coletados. Uma vez que esta variável busca representar o comparecimento dos clientes, sua probabilidade foi parametrizada como $1-P_n$.

Foram gerados, então, vetores binários de tamanho N, para os três valores delineados para este parâmetro, os quais combinados com os três valores definidos para P_n , resultaram em 9 tipos de vetores. Para cada tipo de vetor foram gerados 10.000 conjuntos de vetores.

Combinando-se estes conjuntos de dados para os valores delineados para C_v , P_n e N, foram formados 27 cenários com 10.000 sessões cada.

Estes 27 cenários são compostos de 18 situações já contempladas no trabalho de Cayirli et al. (2012), e nove cenários novos, com $N=40$, valor introduzido para adequação ao caso estudado.

Desta maneira, construiu-se um modelo com aplicação das Equações 6 ou 8 e as Equações 16 a 25 aos cenários para obtenção do valor de custo total (TC) para cada uma das sessões simuladas. Como TC de cada cenário, que é o indicador para comparativo de desempenho entre regras de agendamento, foi considerada a média aritmética dos custos totais de suas 10.000 sessões.

4.6.3 Obtenção da Constante k para a Regra Dome 1

No modelo selecionado o parâmetro k , aplicado às Equações 6 ou 8, define em que grau se configura o formato de domo, ou seja, o quão acentuados são a redução dos intervalos entre serviços no início e final da sessão e o aumento dos intervalos em sua parte intermediária.

Cayirli et al. (2012) plotam uma série de valores para o parâmetro k , formando uma fronteira da eficiência de custos para cada cenário. Para determinar o valor de k mais aproximado do CR desejado, utilizam a técnica de interpolação de *splines*. Esta técnica de aproximação consiste em se dividir o intervalo de interesse em vários subintervalos e interpolar, da forma mais suave possível, nestes subintervalos com polinômios de grau pequeno. Neste caso foram empregados *splines* cúbicos, correspondentes a polinômios de 2º grau.

No presente estudo, os valores de k foram obtidos por meio da ferramenta Solver (da Frontline Solvers), em ambiente MS-Excel 2013.

Foi empregado o método GRG-Não Linear para minimização do custo total (TC), tendo como variável de decisão o parâmetro k e como função objetivo a Equação 24. Não foram inseridas restrições neste modelo.

Esta abordagem permitiu a obtenção de valores mais precisos em relação ao modelo de Cayirli et al. (2012), o qual considerou apenas 2 casas decimais para os valores de k .

Vale ressaltar que os indicadores de todos os cenários do presente trabalho foram padronizados com o uso desta técnica.

4.6.4 Ajustes para *No-Shows*

No modelo proposto, um comparativo inicial é realizado com o objetivo de se verificar o desempenho do cenário com e sem ajustes para *no-shows* na regra Dome 1.

Em razão da acuracidade de 1%, para um intervalo de confiança de 95% atribuída à geração de 10.000 sessões para os conjuntos de dados simulados, não foram repetidos os comparativos realizados para os 18 cenários já analisados no modelo selecionado. Naquela análise pôde-se observar que os ajustes realizados proporcionaram desempenho geral superior em relação às simulações sem ajuste, optando-se pelo ajuste para o comparativo entre a regra Dome 1 e as regras tradicionais.

No entanto, para os novos cenários introduzidos com a extensão de N para 40, esta verificação tornou-se necessária.

Deste modo, foram computados os valores de custo total (TC) para estes nove cenários sem ajuste, utilizando a Equação 6, e com ajuste para *no-shows*, utilizando a Equação 8 com os parâmetros das Equações 3 e 4. Considerando-se os três valores delineados para CR, foram comparados dois conjuntos com 27 indicadores cada (nove cenários com três valores de CR).

A Tabela 12 aponta os percentuais de ganho dos cenários com ajuste em relação aos cenários sem ajuste.

Tabela 12 – Ganhos de Desempenho Decorrentes do Ajuste para *No-Shows* e *Walk-Ins* para N=40

Cv	Pn	N=40		
		CR=2	CR=5	CR=10
0,4	0,05	-0,11%	0,06%	0,23%
0,4	0,2	-0,09%	0,52%	0,95%
0,4	0,4	0,61%	1,21%	1,58%
0,6	0,05	-0,03%	0,02%	0,08%
0,6	0,2	-0,03%	0,19%	0,34%
0,6	0,4	0,10%	0,29%	0,42%
0,8	0,05	-0,01%	0,00%	0,03%
0,8	0,2	0,00%	0,03%	0,06%
0,8	0,4	-0,01%	-0,03%	-0,04%
Mínimos		-0,11%	-0,03%	-0,04%
Máximos		0,61%	1,21%	1,58%
Médias		0,05%	0,26%	0,41%

Embora em alguns casos tenha-se observado uma ligeira queda de desempenho, em geral houve melhora nos custos totais. Deste modo, prosseguiu-se com os estudos comparativos considerando a versão com ajustes para todos os cenários e regras, mesmo critério adotado no modelo selecionado.

4.6.5 Comparativos entre Regras de Agendamento

Nesta fase foram repetidas as simulações para todos os cenários, com ajuste para *no-shows*, inclusive os 18 já pesquisados no modelo proposto, uma vez que se tem por objetivo um comparativo entre todos os cenários e regras.

Para os 27 cenários criados, que compõem 27 conjuntos de fatores ambientais, foram calculados os valores de custo total com os três valores de CR definidos para o experimento, gerando 81 valores para TC para cada regra de agendamento. Sendo o comparativo realizado entre três regras, estes 243 valores de TC foram confrontados.

A Tabela 13 apresenta os resultados, identificando o percentual de ganho da regra Dome 1 em relação à regra tradicional de melhor desempenho.

Tabela 13 – Ganhos de Desempenho da Regra Dome 1 em Relação à Melhor Regra Tradicional

Cv	Pn	N=10			N=20			N=40		
		CR=2	CR=5	CR=10	CR=2	CR=5	CR=10	CR=2	CR=5	CR=10
0,4	0,05	1,05%	6,75%	3,70%	7,30%	4,32%	2,32%	22,51%	5,95%	3,14%
0,4	0,2	1,56%	1,07%	-0,83%	5,51%	1,92%	-0,91%	18,12%	4,90%	0,06%
0,4	0,4	1,03%	0,86%	3,93%	3,84%	0,41%	3,35%	11,19%	1,59%	1,52%
0,6	0,05	1,30%	3,95%	0,28%	6,32%	4,16%	-0,12%	20,37%	5,59%	1,11%
0,6	0,2	2,24%	0,17%	0,47%	5,11%	0,87%	0,22%	16,33%	3,75%	0,47%
0,6	0,4	0,65%	1,52%	5,14%	2,96%	0,95%	4,63%	10,06%	1,71%	2,30%
0,8	0,05	0,85%	0,88%	-0,02%	4,97%	1,65%	-0,03%	18,55%	4,93%	0,81%
0,8	0,2	2,36%	0,07%	1,69%	4,47%	0,58%	1,41%	14,49%	3,01%	1,00%
0,8	0,4	0,47%	2,24%	6,18%	2,40%	1,50%	5,70%	8,73%	1,85%	3,09%
Mínimos		0,47%	0,07%	-0,83%	2,40%	0,41%	-0,91%	8,73%	1,59%	0,06%
Máximos		2,36%	6,75%	6,18%	7,30%	4,32%	5,70%	22,51%	5,95%	3,14%
Médias		1,28%	1,95%	2,28%	4,76%	1,82%	1,84%	15,59%	3,70%	1,50%

Para os tamanhos de sessão (N) 10 e 20, os valores obtidos ficaram bastante próximos dos apontados no estudo de Cayirli et al. (2012), com desempenho médio superior da regra Dome 1, conforme pode ser observado nos resultados em que Pw=0 na Tabela 8. Da mesma forma se repetiram os cinco casos de desempenho inferior, todos também em relação à regra 2BEG.

Para $N=40$, valor introduzido para adequação à realidade do caso estudado, percebeu-se um ganho em todos os cenários. Para os valores de $CR=5$ e $CR=10$, as diferenças ficaram em um patamar semelhante aos casos de $N=10$ e $N=20$. Para $CR=2$, contudo, os ganhos se apresentaram expressivamente superiores, variando entre 8,73% e 22,51%, com média de 15,59%.

Analisando-se todo o conjunto de cenários, percebe-se que para $CR=2$ os ganhos da regra Dome 1 em relação às regras tradicionais crescem consistentemente com o aumento de N , para todas as combinações de C_v e P_n . Assim como no estudo de Cayirli et al. (2012), contudo, esta análise não identifica claramente nenhuma outra tendência além desta.

Para efeitos de desempenho geral do modelo proposto, constata-se que dos 81 casos avaliados, apenas 5 apresentam desempenho inferior da regra Dome 1, sendo que as diferenças se situam em uma faixa entre 0,02% e 0,91%.

Para cada um dos 27 cenários, compostos por conjuntos de fatores ambientais (N , C_v e P_n), pode ser criada uma fronteira da eficiência com os diferentes valores de k da regra Dome 1.

Em seguida são analisados mais detalhadamente os cenários que tiveram o pior e o melhor desempenho da regra Dome 1 em relação às regras tradicionais.

O Gráfico 4 se refere ao cenário com $N=10$, $C_v=0,8$ e $P_n=0,05$, no qual a regra Dome 1 teve o menor incremento de resultados, com ganho, na média dos três valores de CR , de 0,57% no custo total (TC).

Os três valores de k formam a fronteira da eficiência da regra Dome 1 para a faixa do parâmetro CR delineada para o experimento, sendo que para cada valor de CR pode ser traçada uma reta de isocusto tangenciando esta fronteira.

Para as regras 2BEG e IBFI não há uma fronteira, uma vez cada uma gera um único ponto no gráfico, independentemente do valor de CR .

Pode-se notar que o ponto gerado pela regra 2BEG está fora da fronteira da eficiência, o que significa que para algum faixa de valores de CR esta regra supera a Dome 1. Traçando-se uma reta de isocusto com inclinação correspondente a $CR=2$, passando pelo ponto gerado pela regra 2BEG, esta ficaria mais distante da origem do que a reta

de isocusto gerada pela regra Dome 1, confirmando o melhor desempenho desta última. O mesmo pode ser constatado com $CR=5$.

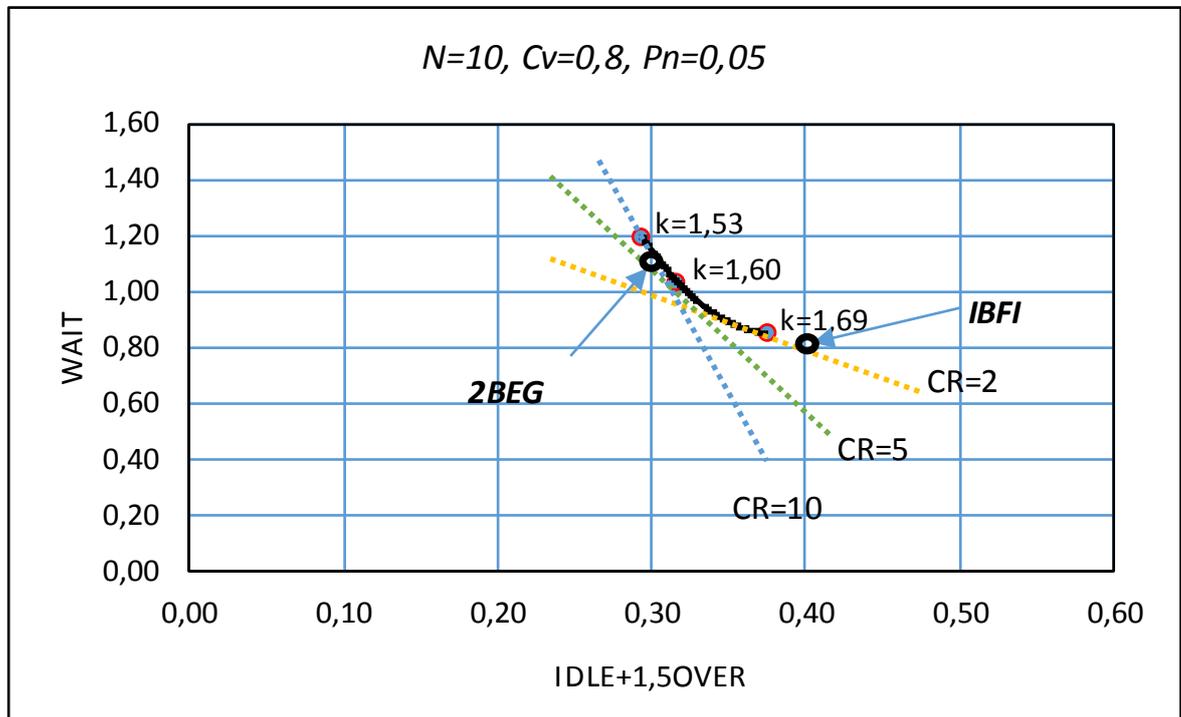


Gráfico 4 - Fronteira da Eficiência para $N=10$, $Cv=0,8$ e $Pw=0,05$

Já para $CR=10$, uma reta de isocusto passando pelo ponto gerado pela regra 2BEG ficaria mais próxima da origem do que a reta de isocusto da regra Dome 1, apontando seu melhor desempenho.

Esta análise condiz com os resultados da Tabela 13 relativos a este cenário.

Aplicando-se a mesma análise entre as regras Dome 1 e IBFI, percebe-se que retas de isocusto passando pelo ponto gerado por esta última ficariam mais distantes da origem do que as retas geradas pela regra Dome.

É importante destacar que, devido à delimitação da faixa do parâmetro CR aos valores entre 2 e 10, não se pode avaliar se o ponto correspondente à regra IBFI está dentro ou fora da fronteira de eficiência. O perfil da curva gerada leva a supor que para um valor de CR mais baixo a regra IBFI supere a Dome 1, mas somente a extensão da

faixa de valores de CR e, conseqüentemente, da fronteira da eficiência, poderia permitir uma conclusão.

Aplicando-se a mesma análise no cenário de melhor desempenho da regra Dome 1, com base no Gráfico 5, constata-se que a regra 2BEG tem pior desempenho que a Dome 1, uma vez que retas de isocusto que passem pelo ponto que a caracteriza no gráfico estarão mais afastadas da origem do que as retas geradas pela regra Dome 1. Uma extensão das simulações para uma faixa além da estipulada para CR poderia apontar se a regra 2BEG tem desempenho superior ou inferior à Dome 1 para valores mais altos de CR.

Quanto à regra IBFI, fazendo-se a mesma projeção das retas de isocusto, constata-se da mesma forma que seu desempenho é inferior à Dome 1. Neste caso, o ponto que a caracteriza está dentro da fronteira da eficiência, permitindo afirmar que seu desempenho é inferior para qualquer valor de CR, mesmo que fora da faixa definida para este estudo. Novamente o gráfico reflete os números da Tabela 13.

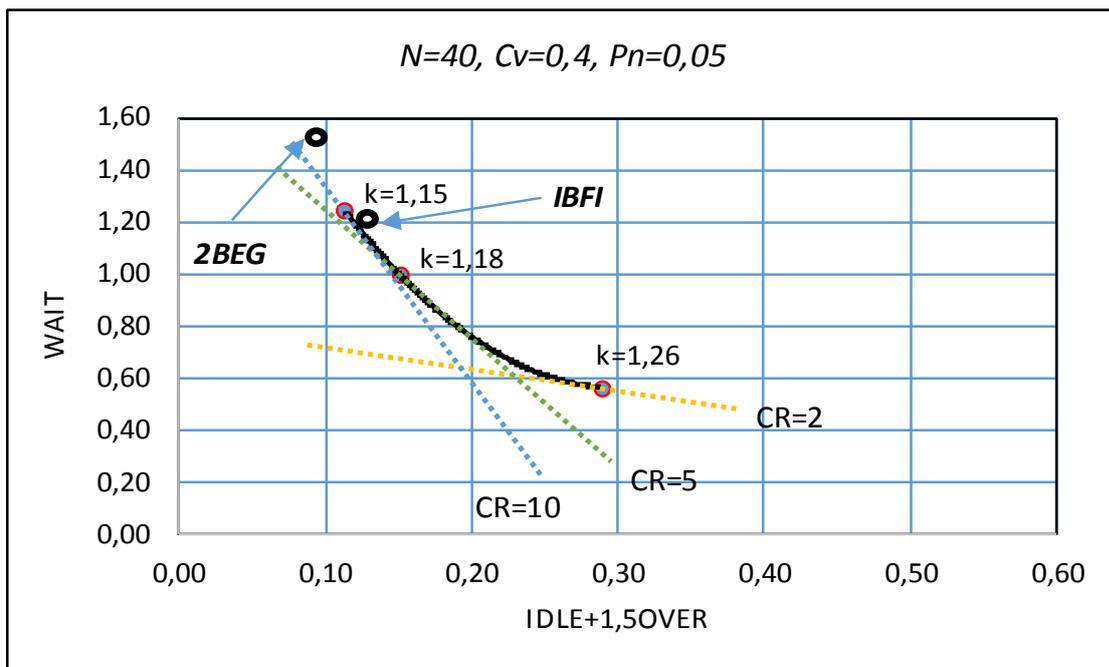


Gráfico 5 - Fronteira da Eficiência para $N=40$, $Cv=0,4$ e $Pw=0,05$

Foi realizado, ainda, um comparativo medindo as diferenças de desempenho da regra 2BEG em relação à IBFI, conforme resume a Tabela 14.

Percebe-se o predomínio de melhor desempenho da regra IBFI para CR=2, enquanto que na totalidade dos cenários com CR=10 a regra 2BEG proporciona resultados melhores. Pode-se destacar a amplitude das diferenças, partindo de -33,71%, para N=10, Cv=0,4, Pn=0,05 e CR=2, com valor máximo de 18,99% para N=10, Cv=0,4, Pn=0,4 e CR=10.

Tabela 14 – Comparativo de Desempenho da Regra 2BEG em Relação à Regra IBFI

Cv	Pn	N=10			N=20			N=40		
		CR=2	CR=5	CR=10	CR=2	CR=5	CR=10	CR=2	CR=5	CR=10
0,4	0,05	-33,71%	-5,17%	10,72%	-24,55%	-6,48%	6,93%	-16,01%	-6,74%	2,32%
0,4	0,2	-9,54%	8,11%	17,23%	-9,25%	3,36%	12,04%	-7,22%	-0,20%	6,20%
0,4	0,4	5,45%	14,72%	18,99%	1,07%	9,07%	13,89%	-1,04%	4,01%	8,12%
0,6	0,05	-15,87%	3,44%	14,21%	-11,76%	0,67%	9,91%	-7,98%	-1,64%	4,56%
0,6	0,2	-4,06%	9,53%	16,58%	-4,72%	4,98%	11,67%	-4,12%	1,24%	6,15%
0,6	0,4	5,61%	13,39%	17,00%	1,75%	8,48%	12,55%	-0,32%	3,91%	7,36%
0,8	0,05	-8,30%	6,07%	14,11%	-6,35%	2,92%	9,80%	-4,49%	0,29%	4,95%
0,8	0,2	-1,20%	9,69%	15,35%	-2,43%	5,31%	10,64%	-2,38%	1,92%	5,86%
0,8	0,4	5,39%	12,07%	15,16%	1,90%	7,62%	11,06%	0,03%	3,64%	6,58%
Mínimos		-33,71%	-5,17%	10,72%	-24,55%	-6,48%	6,93%	-16,01%	-6,74%	2,32%
Máximos		5,61%	14,72%	18,99%	1,90%	9,07%	13,89%	0,03%	4,01%	8,12%
Médias		-6,25%	7,98%	15,48%	-6,04%	3,99%	10,94%	-4,84%	0,72%	5,79%

Estes resultados estão alinhados com as afirmações de Cayirli et al. (2012) quanto às regras tradicionais não apresentarem bom desempenho em todas as situações aplicadas, ao mesmo tempo em que reforça o caráter universal da regra Dome 1.

Os números gerados nesta pesquisa permitem também uma análise da distribuição dos tempos de espera dos clientes. O fato de o indicador de desempenho do modelo proposto (TC) ser composto tanto deste tempo como do tempo de ociosidade do servidor não permite comparar este parâmetro isoladamente. Entretanto, sua variabilidade pode ser comparada entre as regras e cenários, mediante análise dos seus coeficientes de variação, conforme proposto na metodologia deste estudo.

O maior desafio de um sistema de agendamento é lidar com as incertezas, que no caso estudado, têm com fontes o tempo de serviço e o *no-show*. Assim sendo, foram selecionados os cenários de maior e menor variabilidade para o parâmetro N=40,

introduzido no estudo. O tempo de espera do primeiro cliente foi excluído em todas as medições de coeficientes de variação, uma vez que é sempre igual a zero.

Para o cenário de menor variabilidade, correspondente a $Cv=0,4$ e $Pn=0,005$, obtiveram-se os resultados da Figura 10. Observa-se que os tempos de espera do cliente apresentam um comportamento crescente ao longo da sessão em todas as regras. Para a regra Dome 1, os coeficientes de variação deste parâmetro aumentam em função do CR. Para o caso de $CR=2$, a variabilidade obtida com a regra Dome 1 é significativamente menor que nos demais valores de CR, assim como em relação às regras tradicionais. Ainda para o caso de $CR=2$, o comportamento crescente dos tempos de espera é menos acentuado na regra Dome 1.

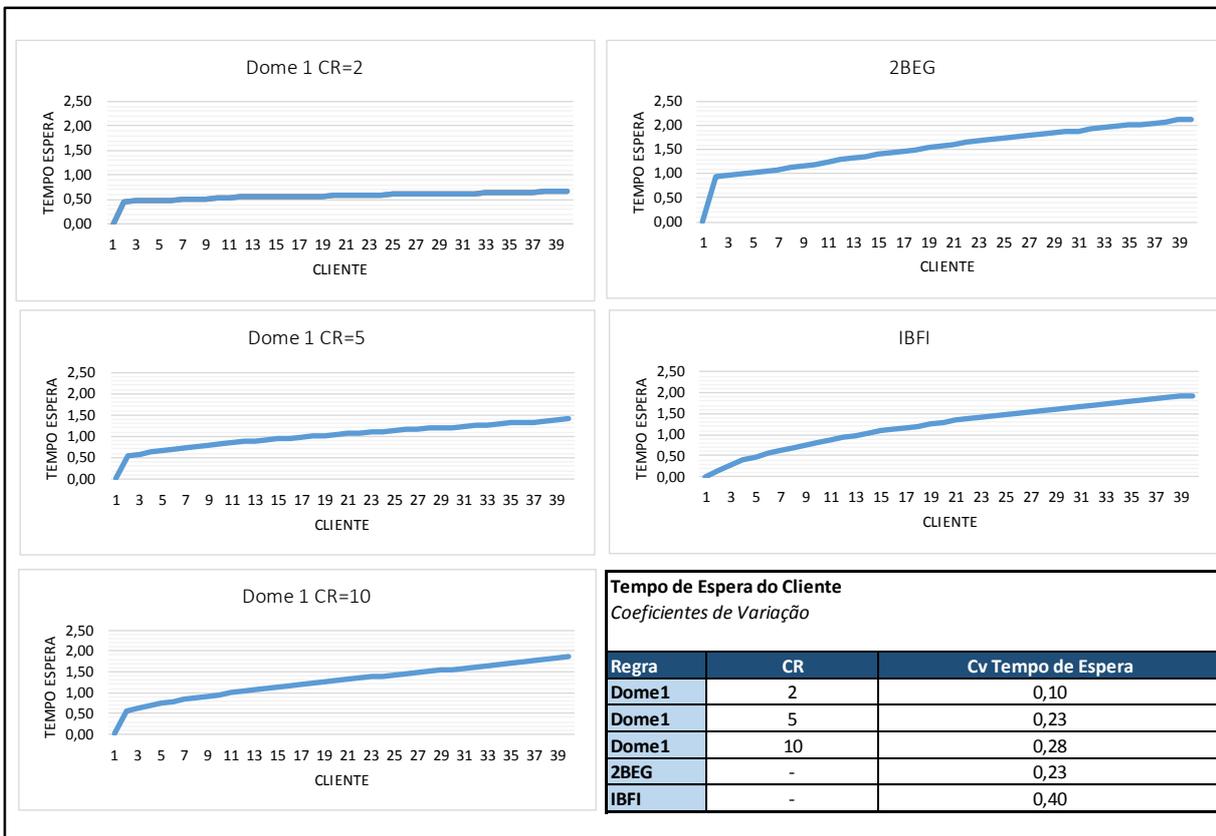


Figura 10 - Variabilidade dos Tempos de Espera para $N=40$, $Cv=0,4$ e $Pn=0,05$

Quanto ao cenário de maior variabilidade, correspondente a $Cv=0,8$ e $Pn=0,4$ (Figura 11), o crescimento dos tempos de espera ao longo da sessão se repetem em todas as situações, assim como o aumento do Cv em relação ao aumento de CR na regra

Dome 1. A redução do Cv obtida com esta regra para CR=2, contudo, não é tão acentuada quanto no cenário anterior.

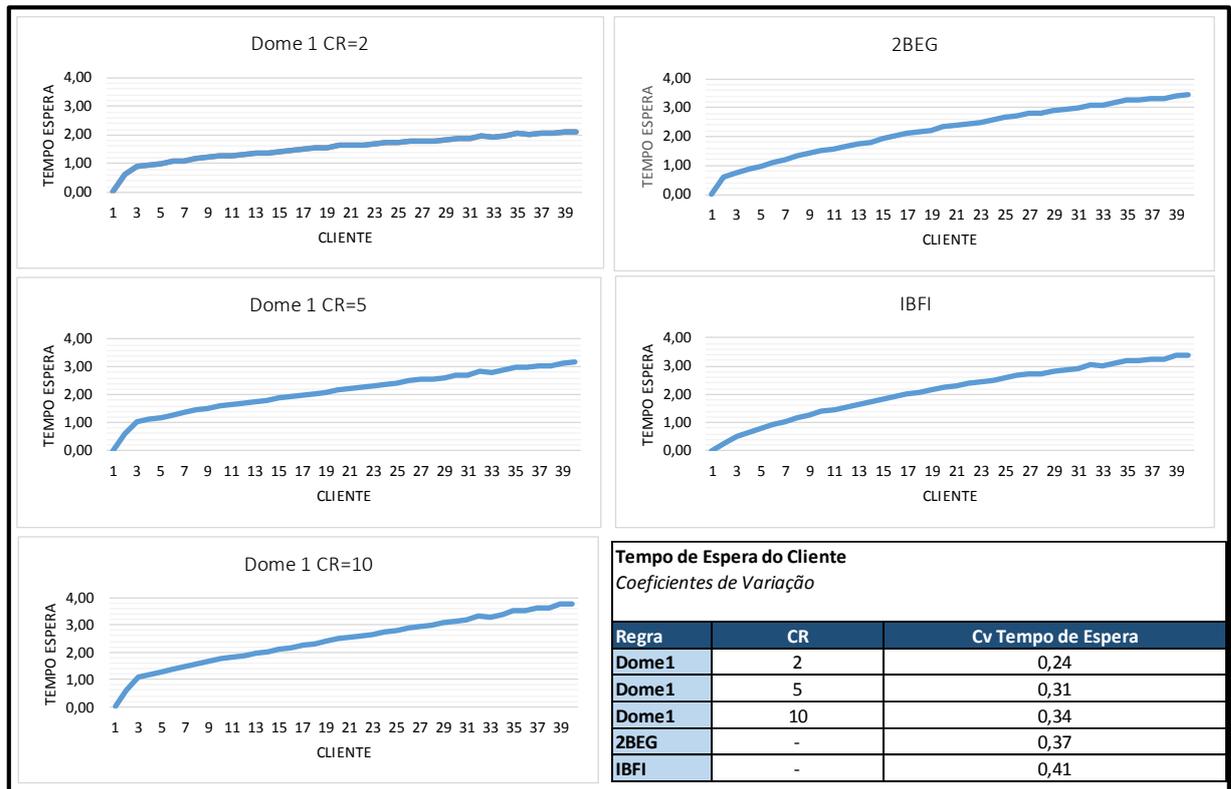


Figura 11 - Variabilidade dos Tempos de Espera para N=40, Cv=0,8 e Pn=0,4

4.7 APLICAÇÃO DO MODELO PROPOSTO PELO AUTOR

Uma última análise foi realizada com emprego do modelo proposto pelo autor deste trabalho. Empregando a ferramenta Solver da Frontline Solvers, em ambiente MS-Excel 2013, foi adotado o método GRG (Gradiente Reduzido Generalizado) -Não Linear para minimização do custo total (TC). Segundo Lasdon et al. (1974), este método apresenta bom desempenho na solução de problemas de otimização não-lineares. A submissão do modelo ao método Simplex, desenvolvido para otimizações lineares, foi realizada para confirmação da não-linearidade do problema.

Neste modelo, em vez de se aplicar uma regra, os horários de agendamento para os N clientes da sessão foram configurados com sendo as variáveis de decisão. Esta modelagem está resumida no Quadro 13.

Quadro 13 – Modelagem para Aplicação da Ferramenta Solver

Função objetivo	<i>Mín (TC)</i>
Variáveis de Decisão	A_i ($i=1$ a N)
Restrições	$A_i \geq 0$ ($i=1$ a N)
	$A_i \leq A_{i+1}$ ($i=1$ a N)

Este comparativo foi limitado aos cenários simulados com $C_v=0,8$, considerados pelos autores do modelo selecionado como sendo os mais desafiadores em termos de desempenho.

Os resultados da análise, comparados aos obtidos com a regra de melhor desempenho, podem ser observados na Tabela 15.

Tabela 15 – Ganhos de Desempenho da Regra Dome 1 com Aplicação da Ferramenta Solver

Cv	Pn	N=10			N=20			N=40		
		CR=2	CR=5	CR=10	CR=2	CR=5	CR=10	CR=2	CR=5	CR=10
0,8	0,05	2,12%	1,06%	0,69%	1,27%	0,52%	0,93%	1,08%	0,86%	0,84%
0,8	0,2	0,86%	0,16%	0,41%	0,54%	0,24%	0,95%	0,84%	0,43%	0,89%
0,8	0,4	0,12%	0,33%	0,61%	0,10%	0,63%	1,36%	0,87%	0,60%	1,64%
Mínimos		0,12%	0,16%	0,41%	0,10%	0,24%	0,93%	0,84%	0,43%	0,84%
Máximos		2,12%	1,06%	0,69%	1,27%	0,63%	1,36%	1,08%	0,86%	1,64%
Médias		1,03%	0,52%	0,57%	0,64%	0,46%	1,08%	0,93%	0,63%	1,12%

Os ganhos em custo total (TC) resultantes desta abordagem variaram entre 0,1% e 2,12%. Pode-se afirmar que, ao mesmo tempo em que é possível atingir um desempenho superior com este método, em todos os cenários analisados, a regra Dome 1 proporciona números bastante próximos, aliados a uma menor complexidade.

Por fim, um comparativo de desempenho do modelo proposto com o atual sistema de agendamento praticado na instituição do caso estudado não foi possível, uma vez que os dados históricos de tempos de serviço não estavam disponíveis.

Concluída esta análise, responde-se à quarta pergunta de pesquisa, mediante os resultados obtidos na Tabela 13:

- Qual o desempenho do modelo de sistema de agendamento de maior potencial, quando aplicado especificamente em um prestador de serviços de medicina diagnóstica com ampla variedade de serviços oferecidos?

Para os tamanhos de sessão (N) 10 e 20, os valores obtidos ficaram bastante próximos dos apontados no estudo de Cayirli et al. (2012), com desempenho médio superior da regra Dome 1 em relação às regras consideradas tradicionais.

Para N=40, valor introduzido para adequação à realidade do caso estudado, percebeu-se um ganho em todos os cenários. Para os valores de CR=5 e CR=10, as diferenças ficaram em um patamar semelhante aos casos de N=10 e N=20. Para CR=2, contudo, os ganhos se apresentaram expressivamente superiores, variando entre 8,73% e 22,51%, com média de 15,59%.

Para efeitos de desempenho geral do modelo proposto, constata-se que dos 81 casos avaliados, apenas 5 apresentam desempenho inferior da regra Dome 1, sendo que as diferenças se situam em uma faixa entre 0,02% e 0,91%.

5 CONCLUSÕES

Motivada pela crescente pressão por ganhos de eficiência e redução de custos na área de saúde, esta pesquisa teve como foco sistemas de agendamento para serviços ambulatoriais, mais precisamente os serviços de medicina diagnóstica.

Buscando reduzir a lacuna entre a teoria e a prática acerca do tema, o trabalho foi conduzido na forma de simulação com modelagem quantitativa, tendo como unidade de análise um laboratório de medicina diagnóstica de grande porte, com extenso portfólio de serviços agendados.

O trabalho foi norteado pela seguinte pergunta de pesquisa:

Qual modelo de sistema de agendamento é o mais adequado para um prestador de serviços de medicina diagnóstica com ampla variedade de serviços oferecidos?

A seguir é apresentado o desdobramento desta pergunta em quatro questões de pesquisa, com suas respectivas respostas.

Quais são os sistemas de agendamento existentes na literatura e direcionados para aplicação em serviços ambulatoriais?

Em revisão da literatura a partir da década de 1950, quando foram desenvolvidos os primeiros estudos sobre sistemas de agendamento para serviços ambulatoriais, foram identificados três fatores que caracterizam um sistema de agendamento: regras de agendamento, classificação de pacientes e ajustes para *no-shows* e *walk-ins*

Conforme apontado no Quadro 14, das 24 combinações básicas destes fatores, as quais formam 24 sistemas de agendamento, 18 foram encontradas na teoria. Seis combinações de fatores, portanto, formam uma lacuna a ser explorada.

Quadro 14 – Sistemas de Agendamento encontrados na literatura

Intervalo entre Serviços	Bloco	Ajustes No-Shows/Walk-Ins	Classificação Pacientes	Encontrado na Literatura?
Fixo	Individual	Sim	Sim	Sim
Fixo	Individual	Sim	Não	Sim
Fixo	Individual	Não	Sim	Sim
Fixo	Individual	Não	Não	Sim
Fixo	Múltiplo	Sim	Sim	Sim
Fixo	Múltiplo	Sim	Não	Sim
Fixo	Múltiplo	Não	Sim	Sim
Fixo	Múltiplo	Não	Não	Sim
Fixo	Variável	Sim	Sim	Não
Fixo	Variável	Sim	Não	Sim
Fixo	Variável	Não	Sim	Sim
Fixo	Variável	Não	Não	Sim
Variável	Individual	Sim	Sim	Sim
Variável	Individual	Sim	Não	Sim
Variável	Individual	Não	Sim	Sim
Variável	Individual	Não	Não	Sim
Variável	Múltiplo	Sim	Sim	Não
Variável	Múltiplo	Sim	Não	Não
Variável	Múltiplo	Não	Sim	Não
Variável	Múltiplo	Não	Não	Sim
Variável	Variável	Sim	Sim	Não
Variável	Variável	Sim	Não	Sim
Variável	Variável	Não	Sim	Não
Variável	Variável	Não	Não	Sim

Quais modelos de sistemas de agendamento encontrados na literatura apresentam maior potencial para aplicação em prestadores de serviços de medicina diagnóstica com ampla diversidade de serviços agendados?

Os 51 estudos sobre modelos de sistemas de agendamento contemplados na revisão teórica foram submetidos a uma seleção, buscando compatibilidade com os objetivos do estudo, ainda sem utilizar os dados do caso estudado como critério. Foram ainda consideradas questões referentes a aplicação prática e o desempenho dos modelos para esta análise. Este processo está ilustrado na Figura 2.

Como resultado, seis modelos (Quadro 15) foram considerados como de potencial para aplicação em prestadores de serviços de medicina diagnóstica com amplo portfólio de serviços agendados.

Quadro 15 – Potencias modelos da literatura para aplicação em medicina diagnóstica com ampla diversidade de serviços agendados

Modelo
Berg et al. (2014)
Cayirli et al. (2006)
Cayirli et al. (2012); Cayirli e Yang (2014)
Ho e Lau (1992, 1999); Ho et al. (1995)
Vanden Bosch e Dietz (2000)
Yang et al. (1998)

Quais indicadores de desempenho podem ser utilizados para avaliar a adequação desses modelos?

Ainda com base na literatura sobre sistemas de agendamento para serviços ambulatoriais, uma ampla gama de indicadores foi identificada e classificada em cinco grupos principais, conforme apresentado no Quadro 16.

Para o presente trabalho foi utilizado um indicador de custo total, composto pelos custos associados aos tempos de espera do paciente e ociosidade e horas-extras do servidor, conforme abordagem do modelo selecionado para aplicação no caso estudado.

Qual o desempenho do modelo de sistema de agendamento de maior potencial, quando aplicado especificamente em um prestador de serviços de medicina diagnóstica com ampla variedade de serviços oferecidos?

Entre as seis opções elencadas na resposta à segunda questão de pesquisa (Quadro 13), o modelo de Cayirli et al. (2012) foi identificado como de maior potencial para aplicação no caso estudado. Além de propor flexibilidade quanto à cobertura de diferentes cenários e complexidade relativamente baixa de implantação, este sistema de agendamento apresentou melhor desempenho quanto aos indicadores de custos adotados para este estudo. Foi constatado, ainda, que o modelo selecionado acumula grande parte dos conceitos da literatura revisada.

Quadro 16 – Indicadores de desempenho identificados na literatura (Cayirli e Veral, 2003)

Indicadores Baseados em Custos
Custo médio total calculado com combinações de:
Tempo de espera de pacientes
Tempo de fluxo de pacientes
Tempo ocioso do provedor
Horas-extras do provedor
Indicadores Baseados em Tempo
Média, máximo e distribuição de frequência do tempo de espera de pacientes
Média, variância e distribuição de frequência do tempo ocioso do provedor
Média, máximo e desvio padrão de horas-extras do provedor
Média e distribuição de frequência do tempo de fluxo de pacientes
Percentual de pacientes atendidos em até 30 min após o horário agendado
Indicadores de Congestionamento
Média e distribuição de frequência do número de pacientes na fila
Média e distribuição de frequência do número de pacientes no sistema
Indicadores de Equidade
Tempo médio de espera de pacientes conforme o estágio da sessão
Variância do tempo de espera de pacientes
Variância do tamanho da fila
Outros Indicadores
Produtividade do provedor
Ocupação média do provedor
Tempo entre a solicitação do serviço e o agendamento
Percentual de pacientes atendidos em urgência
Probabilidade de os pacientes serem agendados nos horários requisitados
Eficiência da sessão

Este modelo consiste em um sistema com intervalos variáveis entre agendamentos, os quais seguem um comportamento de “domo”, ou seja, são menores no início da sessão e aumentam progressivamente até a sua fase intermediária, voltando a valores menores em direção ao seu final.

O sistema configura, ainda, uma regra universal para agendamento e, por meio de parametrização, se propõe a cobrir diferentes proporções entre custos de pacientes e servidores.

Utilizando modelagem quantitativa, os cenários do caso estudado foram reproduzidos com abordagem de simulação, tendo como indicador o custo ponderado entre a ociosidade do servidor e a espera dos clientes. Os resultados, resumidos na Tabela 16, apontaram para ganhos expressivos de desempenho em um amplo espectro de

cenários, quando comparados a regras tradicionais de agendamento, igualmente submetidas às simulações.

Tabela 16 – Ganhos de Desempenho da Regra Dome 1 em Relação à Melhor Regra Tradicional

Cv	Pn	N=10			N=20			N=40		
		CR=2	CR=5	CR=10	CR=2	CR=5	CR=10	CR=2	CR=5	CR=10
0,4	0,05	1,05%	6,75%	3,70%	7,30%	4,32%	2,32%	22,51%	5,95%	3,14%
0,4	0,2	1,56%	1,07%	-0,83%	5,51%	1,92%	-0,91%	18,12%	4,90%	0,06%
0,4	0,4	1,03%	0,86%	3,93%	3,84%	0,41%	3,35%	11,19%	1,59%	1,52%
0,6	0,05	1,30%	3,95%	0,28%	6,32%	4,16%	-0,12%	20,37%	5,59%	1,11%
0,6	0,2	2,24%	0,17%	0,47%	5,11%	0,87%	0,22%	16,33%	3,75%	0,47%
0,6	0,4	0,65%	1,52%	5,14%	2,96%	0,95%	4,63%	10,06%	1,71%	2,30%
0,8	0,05	0,85%	0,88%	-0,02%	4,97%	1,65%	-0,03%	18,55%	4,93%	0,81%
0,8	0,2	2,36%	0,07%	1,69%	4,47%	0,58%	1,41%	14,49%	3,01%	1,00%
0,8	0,4	0,47%	2,24%	6,18%	2,40%	1,50%	5,70%	8,73%	1,85%	3,09%
Mínimos		0,47%	0,07%	-0,83%	2,40%	0,41%	-0,91%	8,73%	1,59%	0,06%
Máximos		2,36%	6,75%	6,18%	7,30%	4,32%	5,70%	22,51%	5,95%	3,14%
Médias		1,28%	1,95%	2,28%	4,76%	1,82%	1,84%	15,59%	3,70%	1,50%

Limitações do Estudo

Embora tenha procurado cobrir uma ampla diversidade de cenários, o emprego de simulação em um caso único como método de pesquisa do trabalho traz limitações quanto à generalização dos resultados obtidos.

Fatores não existentes no caso estudado, como atendimento de *walk-ins* e arranjos com multiservidores, são exemplos que seguramente estão presentes em outros contextos reais.

A indisponibilidade de dados relativos a tempos e custos de prestador de serviços e pacientes, bem como tempos de serviço e sua variabilidade, levou à imposição de premissas, as quais foram assumidas como realistas.

Proposições para Contribuições Futuras

Além de novos cenários, o acréscimo de variáveis ao modelo poderia representar outros fatores presentes na realidade de serviços ambulatoriais, além dos contemplados neste estudo. Entre elas, a pontualidade e o tempo de serviço como

função da fila de pacientes figuram como candidatos abordados na literatura, além dos fatores mencionados nas limitações deste estudo.

A aplicação de um estudo de caso em laboratórios acoplados a um ambiente hospitalar possivelmente permitiria explorar cenários mais complexos em termos de ambientes e variáveis.

Comparativos adicionais também configuram alternativas interessantes de avaliação entre sistemas de agendamento. Além das regras consideradas tradicionais neste estudo, outras combinações de intervalos fixos e bloco inicial podem ser simulados e comparados ao modelo proposto.

Em particular, as regras IBFI e 2BEG, referências de regras tradicionais para comparativos do presente trabalho, poderiam ser simuladas com diferentes valores de média de tempo de serviço, e não somente a média ou a média ajustada. Deste modo, o conceito de fronteira da eficiência poderia ser estendido a estas regras, enriquecendo os comparativos.

Devido à ausência de *walk-ins* no caso estudado, um comparativo também poderia ser desenvolvido com o modelo de intervalos variáveis de Yang et al. (1998), uma vez que o mesmo não contempla este fator. Da mesma forma que os demais modelos comparados, este sistema poderia ser submetido aos ajustes de média e variância de tempos de serviço.

Mesmo que complexos do ponto de vista prático, modelos de otimização robustos poderiam ser desenvolvidos com o propósito de *benchmark*, estabelecendo um “teto” para avaliação de sistemas de agendamento.

Quanto a indicadores, sugere-se adotar o tempo de espera de pacientes como restrição, fixando-se níveis de serviço, e considerar os custos de ociosidade e horas extras do servidor como métrica de desempenho. A distribuição dos tempos de espera, contemplada neste trabalho, complementa esta proposta.

Por fim, a relação entre os coeficientes de variação dos tempos de espera e os custos totais poderia ser melhor explorada a partir dos resultados deste estudo.

ANEXO 1 - SIMULAÇÕES

Este anexo tem por objetivo detalhar as simulações de cenários elaboradas com uso da ferramenta MS-Excel 2013.

Para a construção destes cenários, duas variáveis tiveram valores gerados aleatoriamente: tempo de serviço simulado (s_i) e comparecimento (C_i).

Para a geração de valores de s_i foi empregada a técnica de transformação inversa, conforme ilustrado na Figura 12, em um exemplo com 10 clientes por sessão ($N=10$) e coeficiente de variação (Cv) de 0,6.

	A	B	K
1	Simulação de Tempos de Serviço		
2	Transformação Inversa		
3			
4	Parâmetro	Valor	
5	Média (μ)	1	
6	Coefficiente Variação (Cv)	0,6	
7	Desvio Padrão (σ)	=B6*B5	
8	Variância (σ^2)	=B7^2	
9	Scaled Mean	=LN(B5^2/RAIZ(B5^2+B8))	
10	Scaled Standard Deviation	=RAIZ(LN((B5^2+B8)/B5))	
11			
12	Cliente (i)	1	10
13	Tempo de Serviço Simulado (s_i)	=INV.LOGNORMAL(ALEATÓRIO();\$B\$9;\$B\$10)	=INV.LOGNORMAL(ALEATÓRIO();\$B\$9;\$B\$10)
14		=INV.LOGNORMAL(ALEATÓRIO();\$B\$9;\$B\$10)	=INV.LOGNORMAL(ALEATÓRIO();\$B\$9;\$B\$10)
10011		=INV.LOGNORMAL(ALEATÓRIO();\$B\$9;\$B\$10)	=INV.LOGNORMAL(ALEATÓRIO();\$B\$9;\$B\$10)
10012		=INV.LOGNORMAL(ALEATÓRIO();\$B\$9;\$B\$10)	=INV.LOGNORMAL(ALEATÓRIO();\$B\$9;\$B\$10)

10.000 sessões

10 clientes por sessão (N=10)

Figura 12 - Geração de Dados Aleatórios para Tempos de Serviço Simulados

Os dados para C_i , por sua vez, foram criados com a função Geração de Número Aleatório do MS-Excel, conforme ilustrado na Figura 13, em um exemplo com 10 clientes por sessão ($N=10$) e *no-show* de 5% ($P_n=0,05$).

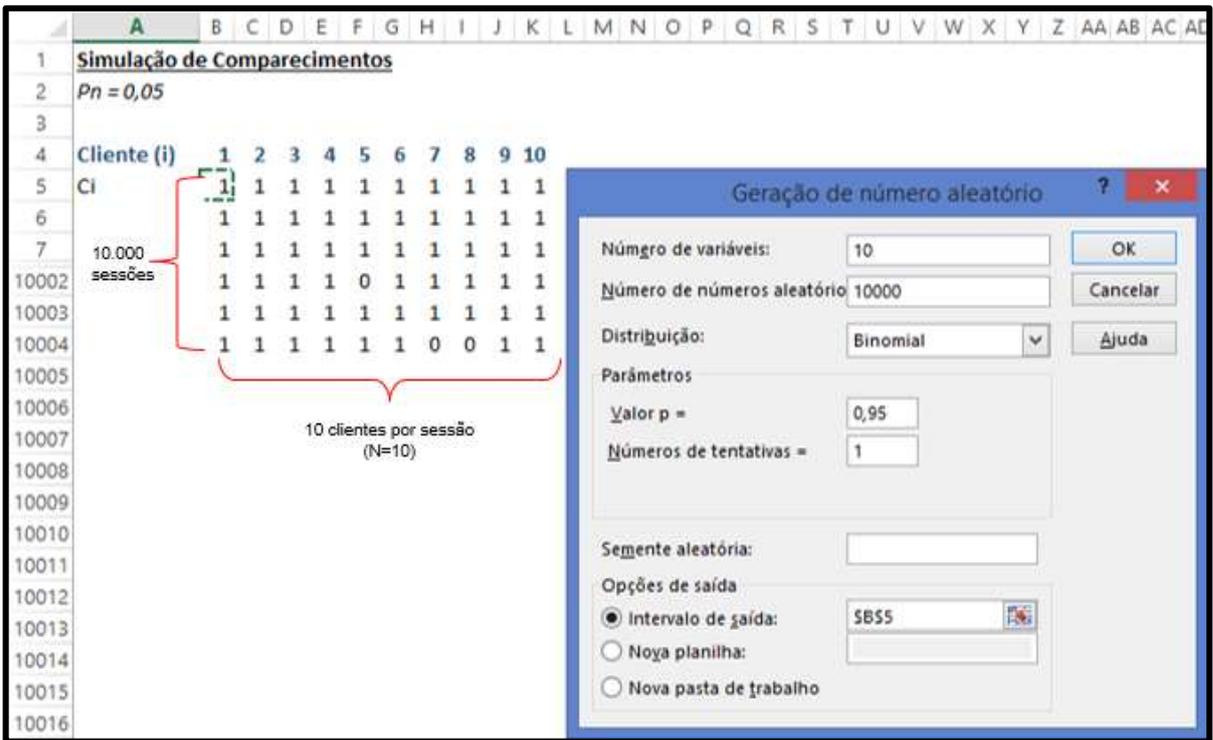


Figura 13 - Geração de Dados Aleatórios para Comparcimentos

Para cada cenário foi criada uma planilha com os dados aleatórios, horários de agendamento, fatores ambientais e indicadores, aplicando as formulações descritas na seção de resultados desta pesquisa, conforme ilustrado na Figura 14.

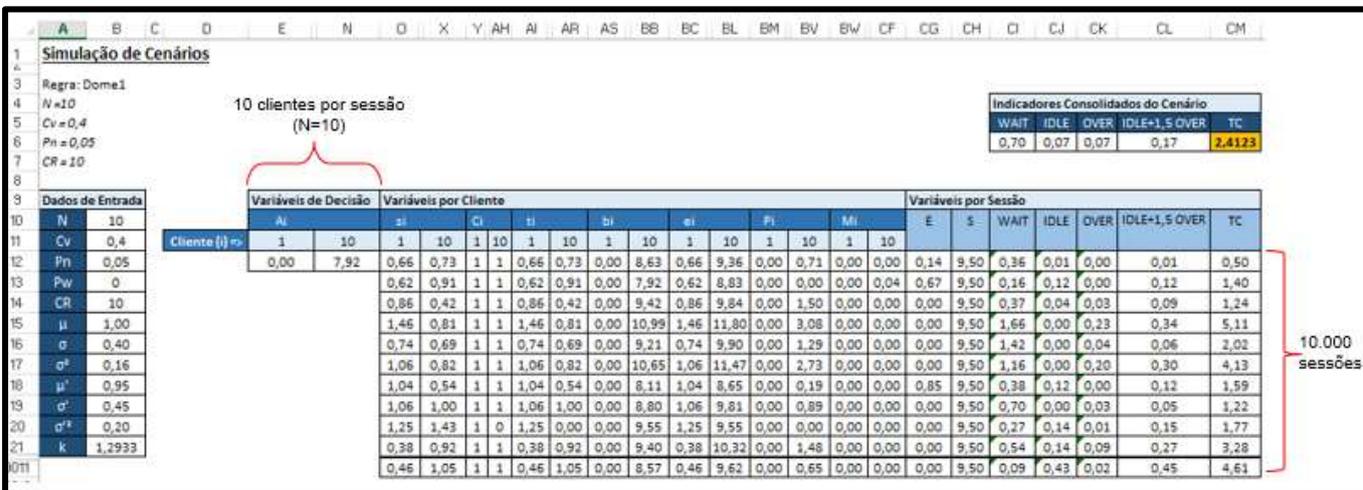


Figura 14 - Simulação de Cenário

As fórmulas aplicadas estão detalhadas nos exemplos das Figuras 15 a 21.

	A	B
1	Simulação de Cenários	
2	Regra: Dome1	
3	N=10	
4	Cv=0,4	
5	Pn=0,05	
6	CR=10	
7		
8		
9	Dados de Entrada	
10	N	10
11	Cv	0,4
12	Pn	0,05
13	Pw	0
14	CR	10
15	μ	1
16	σ	=B11*B15
17	σ^2	=B16^2
18	μ'	=(1-B12+B13)*B15
19	σ'	=RAIZ(B20)
20	σ'^2	=(1-B12-B13)*(B17+(B12-B13)^2*(B15^2))+B12*((1-B12+B13)^2*(B15^2)+B13*(2*B17+((1+B12-B13)^2)*(B15^2)))
21	k	1,29327292897194

Figura 15 - Formulação para os Dados de Entrada

	A	B	C	D	E	N
1	Simulação de Cenários					
2	Regra: Dome1					
3	N=10					
4	Cv=0,4					
5	Pn=0,05					
6	CR=10					
7						
8						
9	Dados de Entrada					
10	N	10				
11	Cv	0,4				
12	Pn	0,05				
13	Pw	0				
14	CR	10				
15	μ	1				
16	σ	=B11*B15				
17	σ^2	=B16^2				
18	μ'	=(1-B12+B13)*B15				
19	σ'	=RAIZ(B20)				
20	σ'^2	=(1-B12-B13)*(B17+(B12-B13)^2*(B15^2))+B12*((1-B12+B13)^2*(B15^2)+B13*(2*B17+((1+B12-B13)^2)*(B15^2)))				
21	k	1,29327292897194				
			10 clientes por sessão (N=10)			
			Variáveis de Decisão			
			Aj			
	Cliente (i) =>	1	10			
			=MÁXIMO(0;S821*(E511-1)*S818-S819*RAIZ(E511)*((S810+E511)/(S810-1)))			
			=MÁXIMO(0;S821*(N511-1)*S818-S819*RAIZ(N511)*((S810+N511)/(S810-1)))			

Figura 16 - Formulação para os Horários de Agendamento (Regra Dome 1)

	A	B	C	D	E	F	N
1	Simulação de Cenários						
2							
3	Regra: IBFI						
4	$N=10$						
5	$Cv=0,4$						
6	$Pn=0,05$						
7	$CR=10$						
8							
9	Dados de Entrada						
10	N	10					
11	Cv	0,4					
12	Pn	0,05					
13	Pw	0					
14	CR	10					
15	μ	1					
16	σ	=B11*					
17	σ^2	=B16^2					
18	μ'	=(1-B1					
19	σ'	=RAIZ(
20	σ'^2	=(1-B1					
21	k	1,2931					

10 clientes por sessão (N=10)			
Variáveis de Decisão			
Ai			
Cliente (j) => 1	2	10	
0	=E12+\$B\$18	=M12+\$B\$18	

Figura 17 - Formulação para os Horários de Agendamento (Regra IBFI)

	A	B	C	D	E	F	G	N
1	Simulação de Cenários							
2								
3	Regra: 2BEG							
4	$N=10$							
5	$Cv=0,4$							
6	$Pn=0,05$							
7	$CR=10$							
8								
9	Dados de Entrada							
10	N	10						
11	Cv	0,4						
12	Pn	0,05						
13	Pw	0						
14	CR	10						
15	μ	1						
16	σ	=B11*						
17	σ^2	=B16^2						
18	μ'	=(1-B1						
19	σ'	=RAIZ(
20	σ'^2	=(1-B1						

10 clientes por sessão (N=10)			
Variáveis de Decisão			
Ai			
Cliente (i) => 1	2	3	10
0	0	=F12+\$B\$18	=M12+\$B\$18

Figura 18 - Formulação para os Horários de Agendamento (Regra 2BEG)

10 clientes por sessão (N=10)

	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI
8																
9	Variáveis por Cliente															
10	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI	AI
11	1	10	1	10	1	10	1	10	1	10	1	10	1	10	1	10
12	0,4	0,7	1	1	=K12*Y12	=K12*AH12	0	=MÁXIMO(N512;BK12)	=A512+AF12	=BB12+AR12	=MÁXIMO(A512-ES12)*Y12	=MÁXIMO(BB12-N512)*AH12	0	=MÁXIMO(N512-BK12)		
13	0,5	0,8	1	1	=K13*Y13	=K13*AH13	0	=MÁXIMO(N513;BK13)	=A513+AF13	=BB13+AR13	=MÁXIMO(A513-ES13)*Y13	=MÁXIMO(BB13-N513)*AH13	0	=MÁXIMO(N513-BK13)		
14	0,6	0,4	1	1	=K14*Y14	=K14*AH14	0	=MÁXIMO(N514;BK14)	=A514+AF14	=BB14+AR14	=MÁXIMO(A514-ES14)*Y14	=MÁXIMO(BB14-N514)*AH14	0	=MÁXIMO(N514-BK14)		
15	1,4	0,8	1	1	=K15*Y15	=K15*AH15	0	=MÁXIMO(N515;BK15)	=A515+AF15	=BB15+AR15	=MÁXIMO(A515-ES15)*Y15	=MÁXIMO(BB15-N515)*AH15	0	=MÁXIMO(N515-BK15)		
16	0,7	0,6	1	1	=K16*Y16	=K16*AH16	0	=MÁXIMO(N516;BK16)	=A516+AF16	=BB16+AR16	=MÁXIMO(A516-ES16)*Y16	=MÁXIMO(BB16-N516)*AH16	0	=MÁXIMO(N516-BK16)		
17	1,0	0,8	1	1	=K17*Y17	=K17*AH17	0	=MÁXIMO(N517;BK17)	=A517+AF17	=BB17+AR17	=MÁXIMO(A517-ES17)*Y17	=MÁXIMO(BB17-N517)*AH17	0	=MÁXIMO(N517-BK17)		
18	1,0	0,5	1	1	=K18*Y18	=K18*AH18	0	=MÁXIMO(N518;BK18)	=A518+AF18	=BB18+AR18	=MÁXIMO(A518-ES18)*Y18	=MÁXIMO(BB18-N518)*AH18	0	=MÁXIMO(N518-BK18)		
19	1,0	1,0	1	1	=K19*Y19	=K19*AH19	0	=MÁXIMO(N519;BK19)	=A519+AF19	=BB19+AR19	=MÁXIMO(A519-ES19)*Y19	=MÁXIMO(BB19-N519)*AH19	0	=MÁXIMO(N519-BK19)		
20	1,2	1,4	1	1	=K20*Y20	=K20*AH20	0	=MÁXIMO(N520;BK20)	=A520+AF20	=BB20+AR20	=MÁXIMO(A520-ES20)*Y20	=MÁXIMO(BB20-N520)*AH20	0	=MÁXIMO(N520-BK20)		
21	0,3	0,8	1	1	=K21*Y21	=K21*AH21	0	=MÁXIMO(N521;BK21)	=A521+AF21	=BB21+AR21	=MÁXIMO(A521-ES21)*Y21	=MÁXIMO(BB21-N521)*AH21	0	=MÁXIMO(N521-BK21)		
10011	0,4	1,0	1	1	=K10011*Y10011	=K10011*AH10011	0	=MÁXIMO(N512;BK10011)	=A510011+AF10011	=BB10011+AR10011	=MÁXIMO(A510011-ES10011)*Y10011	=MÁXIMO(BB10011-N512)*AH10011	0	=MÁXIMO(N512-BK10011)		

10.000 sessões

Figura 19 - Formulação para as Variáveis por Cliente

	CI	CJ	CK	CL	CM
8					
9	Indicadores por Sessão				
10	CI	CJ	CK	CL	CM
11	1	1	1	1	1
12	=MÁXIMO(CI12-BL12)	=MÁXIMO(CJ12-BL12)*Y12	=MÁXIMO(CK12-BL12)/SOMAY12	=MÁXIMO(CL12-BL12)/SOMAY12	=MÁXIMO(CM12-BL12)/SOMAY12
13	=MÁXIMO(CI13-BL13)	=MÁXIMO(CJ13-BL13)*Y13	=MÁXIMO(CK13-BL13)/SOMAY13	=MÁXIMO(CL13-BL13)/SOMAY13	=MÁXIMO(CM13-BL13)/SOMAY13
14	=MÁXIMO(CI14-BL14)	=MÁXIMO(CJ14-BL14)*Y14	=MÁXIMO(CK14-BL14)/SOMAY14	=MÁXIMO(CL14-BL14)/SOMAY14	=MÁXIMO(CM14-BL14)/SOMAY14
15	=MÁXIMO(CI15-BL15)	=MÁXIMO(CJ15-BL15)*Y15	=MÁXIMO(CK15-BL15)/SOMAY15	=MÁXIMO(CL15-BL15)/SOMAY15	=MÁXIMO(CM15-BL15)/SOMAY15
16	=MÁXIMO(CI16-BL16)	=MÁXIMO(CJ16-BL16)*Y16	=MÁXIMO(CK16-BL16)/SOMAY16	=MÁXIMO(CL16-BL16)/SOMAY16	=MÁXIMO(CM16-BL16)/SOMAY16
17	=MÁXIMO(CI17-BL17)	=MÁXIMO(CJ17-BL17)*Y17	=MÁXIMO(CK17-BL17)/SOMAY17	=MÁXIMO(CL17-BL17)/SOMAY17	=MÁXIMO(CM17-BL17)/SOMAY17
18	=MÁXIMO(CI18-BL18)	=MÁXIMO(CJ18-BL18)*Y18	=MÁXIMO(CK18-BL18)/SOMAY18	=MÁXIMO(CL18-BL18)/SOMAY18	=MÁXIMO(CM18-BL18)/SOMAY18
19	=MÁXIMO(CI19-BL19)	=MÁXIMO(CJ19-BL19)*Y19	=MÁXIMO(CK19-BL19)/SOMAY19	=MÁXIMO(CL19-BL19)/SOMAY19	=MÁXIMO(CM19-BL19)/SOMAY19
20	=MÁXIMO(CI20-BL20)	=MÁXIMO(CJ20-BL20)*Y20	=MÁXIMO(CK20-BL20)/SOMAY20	=MÁXIMO(CL20-BL20)/SOMAY20	=MÁXIMO(CM20-BL20)/SOMAY20
21	=MÁXIMO(CI21-BL21)	=MÁXIMO(CJ21-BL21)*Y21	=MÁXIMO(CK21-BL21)/SOMAY21	=MÁXIMO(CL21-BL21)/SOMAY21	=MÁXIMO(CM21-BL21)/SOMAY21
10011	=MÁXIMO(CI10011-BL10011)	=MÁXIMO(CJ10011-BL10011)*Y10011	=MÁXIMO(CK10011-BL10011)/SOMAY10011	=MÁXIMO(CL10011-BL10011)/SOMAY10011	=MÁXIMO(CM10011-BL10011)/SOMAY10011

10.000 sessões

Figura 20 - Formulação para os Indicadores das Sessões

	CI	CJ	CK	CL	CM
1					
2					
3					
4	Indicadores Consolidados do Cenário				
5	WAIT	IDLE	OVER	IDLE+1,5 OVER	TC
6	=MÉDIA(CI12:CI10011)	=MÉDIA(CJ12:CJ10011)	=MÉDIA(CK12:CK10011)	=MÉDIA(CL12:CL10011)	=MÉDIA(CM12:CM10011)
7					

Figura 21 - Formulação para os Indicadores do Cenário

REFERÊNCIAS

- ANDERSON, Kelsey et al. An Overlapping Appointment Scheduling Model with Stochastic Service Time. In: **IIE Annual Conference. Proceedings**. Institute of Industrial Engineers-Publisher, 2013. p. 912.
- ASADUZZAMAN, Md; CHAUSSALET, Thierry J.; ROBERTSON, Nicola J. A loss network model with overflow for capacity planning of a neonatal unit. **Annals of Operations Research**, v. 178, n. 1, p. 67-76, 2010.
- BABES, Malika; SARMA, G. V. Out-patient queues at the Ibn-Rochd health centre. **Journal of the Operational Research Society**, p. 845-855, 1991.
- BARRON, William M. Failed appointments. Who misses them, why they are missed, and what can be done. **Primary care**, v. 7, n. 4, p. 563-574, 1980.
- BAILEY, Norman T.J. A study of queues and appointment systems in hospital out-patient departments, with special reference to waiting-times. **Journal of the Royal Statistical Society. Series B (Methodological)**, p. 185-199, 1952.
- BENNETT, Joanne C.; WORTHINGTON, D. J. An example of a good but partially successful OR engagement: Improving outpatient clinic operations. **Interfaces**, v. 28, n. 5, p. 56-69, 1998.
- BERG, Bjorn P. et al. Optimal booking and scheduling in outpatient procedure centers. **Computers & Operations Research**, v. 50, p. 24-37, 2014.
- BLANCO WHITE, M. J. AND M. C. PIKE, "Appointment Systems in Outpatients' Clinics and the Effect on Patients' Unpunctuality," **Medical Care** 2, 133–145, 1964
- BOWERS, John Andrew. Simulating waiting list management. **Health Care Management Science**, v. 14, n. 3, p. 292-298, 2011.
- BOWERS, John. Balancing operating theatre and bed capacity in a cardiothoracic centre. **Health Care Management Science**, p. 1-9, 2013.
- BRAHIMI, M.; WORTHINGTON, D. J. The finite capacity multi-server queue with inhomogeneous arrival rate and discrete service time distribution—and its application to continuous service time problems. **European Journal of Operational Research**, v. 50, n. 3, p. 310-324, 1991a.

BRAHIMI, M.; WORTHINGTON, D. J. Queueing models for out-patient appointment systems-a case study. **Journal of the Operational Research Society**, p. 733-746, 1991b.

BRYMAN, Alan et al. **Research methods and organization studies**. Routledge, 1989.

BUCHAN, I. C.; RICHARDSON, I. M. Time study of consultations in general practice. **Scottish Home and Health Department**, 1973.

CAYIRLI, Tugba; GUNES, Evrim Didem. Outpatient appointment scheduling in presence of seasonal walk-ins. **Journal of the Operational Research Society**, v. 65, n. 4, p. 512-531, 2013.

CAYIRLI, Tugba; VERAL, Emre. Outpatient scheduling in health care: a review of literature. **Production and Operations Management**, v. 12, n. 4, p. 519-549, 2003.

CAYIRLI, Tugba; VERAL, Emre; ROSEN, Harry. Designing appointment scheduling systems for ambulatory care services. **Health Care Management Science**, v. 9, n. 1, p. 47-58, 2006.

CAYIRLI, Tugba; YANG, Kum Khiong. A Universal Appointment Rule with Patient Classification for Service Times, No-Shows, and Walk-Ins. **Service Science**, v. 6, n. 4, p. 274-295, 2014.

CAYIRLI, Tugba; YANG, Kum Khiong; QUEK, Ser Aik. A Universal Appointment Rule in the Presence of No-Shows and Walk-Ins. **Production and Operations Management**, v. 21, n. 4, p. 682-697, 2012.

CHARNETSKI, J. R. Scheduling operating room surgical procedures with early and late completion penalty costs. **Journal of Operations Management**, v. 5, n. 1, p. 91-102, 1984.

COCHRAN, Jeffery K.; ROCHE, Kevin T. A multi-class queuing network analysis methodology for improving hospital emergency department performance. **Computers & Operations Research**, v. 36, n. 5, p. 1497-1512, 2009.

COX, T. F., J. F. BIRCHALL, AND H. WONG, "Optimising the Queuing System for an Ear, Nose and Throat Outpatient Clinic," **Journal of Applied Statistics**, 12, 113–126, 1985.

DE ANDRADE MARTINS, Gilberto; THEÓPHILO, Carlos Renato. **Metodologia da investigação científica para ciências sociais aplicadas**. Atlas, 2009.

DE VUYST, Stijn; BRUNEEL, Herwig; FIEMS, Dieter. Computationally efficient evaluation of appointment schedules in health care. **European Journal of Operational Research**, v. 237, n. 3, p. 1142-1154, 2014.

DENTON, Brian; GUPTA, Diwakar. A sequential bounding approach for optimal appointment scheduling. **IIE Transactions**, v. 35, n. 11, p. 1003-1016, 2003.

DEYO, Richard A.; INUI, Thomas S. Dropouts and broken appointments: a literature review and agenda for future research. **Medical Care**, p. 1146-1157, 1980.

ERDOGAN, S. Ayca; DENTON, Brian. Dynamic appointment scheduling of a stochastic server with uncertain demand. **INFORMS Journal on Computing**, v. 25, n. 1, p. 116-132, 2013.

FETTER, Robert B.; THOMPSON, John D. Patients' waiting time and doctors' idle time in the outpatient setting. **Health Services Research**, v. 1, n. 1, p. 66, 1966.

FRIES, Brant E.; MARATHE, Vijay P. Determination of optimal variable-sized multiple-block appointment systems. **Operations Research**, v. 29, n. 2, p. 324-345, 1981.

FRY, John. Appointments in general practice. **OR**, p. 233-237, 1964.

GARG, Lalit et al. Intelligent Patient Management and Resource Planning for Complex, Heterogeneous, and Stochastic Healthcare Systems. **Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on**, v. 42, n. 6, p. 1332-1345, 2012.

GREEN, Linda V.; SAVIN, Sergei; WANG, Ben. Managing patient service in a diagnostic medical facility. **Operations Research**, v. 54, n. 1, p. 11-25, 2006.

GUPTA, Diwakar et al. Capacity planning for cardiac catheterization: a case study. **Health Policy**, v. 82, n. 1, p. 1-11, 2007.

GUPTA, Diwakar; DENTON, Brian. Appointment scheduling in health care: Challenges and opportunities. **IIE Transactions**, v. 40, n. 9, p. 800-819, 2008.

HASSIN, Refael; MENDEL, Sharon. Scheduling arrivals to queues: A single-server model with no-shows. **Management Science**, v. 54, n. 3, p. 565-572, 2008.

HEANEY, D. J.; HOWIE, J. G.; PORTER, A. M. Factors influencing waiting times and consultation times in general practice. **British Journal of General Practice**, v. 41, n. 349, p. 315-319, 1991.

HO, Chrwan-jyh; LAU, Hon-Shiang; LI, Jing. Introducing variable-interval appointment scheduling rules in service systems. **International Journal of Operations & Production Management**, v. 15, n. 6, p. 59-68, 1995.

HO, Chrwan-Jyh; LAU, Hon-Shiang. Minimizing total cost in scheduling outpatient appointments. **Management Science**, v. 38, n. 12, p. 1750-1764, 1992.

HO, Chrwan-Jyh; LAU, Hon-Shiang. Evaluating the impact of operating conditions on the performance of appointment scheduling rules in service systems. **European Journal of Operational Research**, v. 112, n. 3, p. 542-553, 1999.

HOFMANN, Paul B.; ROCKART, John F. Implications of the no-show rate for scheduling OPD appointments. **Hospital Progress**, v. 50, n. 8, p. 35, 1969.

HONG, Tang Sai et al. Use of simulation to solve outpatient clinic problems: A review of the literature. **South African Journal of Industrial Engineering**, v. 24, n. 3, p. 27-47, 2013.

HUANG, Xiao-Ming. Patient attitude towards waiting in an outpatient clinic and its applications. **Health Services Management Research**, v. 7, n. 1, p. 2-8, 1994.

HUARNG, Fenghueih; LEE, Mong Hou. Using simulation in out-patient queues: a case study. **International Journal of Health Care Quality Assurance**, v. 9, n. 6, p. 21-25, 1996.

JACKSON, R. R. P. Design of an appointments system. **OR**, p. 219-224, 1964.

JANSSON, Birger. Choosing a good appointment system—A study of queues of the type (D, M, 1). **Operations Research**, v. 14, n. 2, p. 292-312, 1966.

JEANG, Angus; CHIANG, An-Jen. Economic and Quality Scheduling for Effective Utilization of Operating Rooms. **Journal of Medical Systems**, v. 36, n. 3, p. 1205-1222, 2012.

JOHNSON, Walter L.; ROSENFELD, Leonard S. Factors affecting waiting time in ambulatory care services. **Health Services Research**, v. 3, n. 4, p. 286, 1968.

KAANDORP, Guido C.; KOOLE, Ger. Optimal outpatient appointment scheduling. **Health Care Management Science**, v. 10, n. 3, p. 217-229, 2007.

KATZ, Jesse H. Simulation of outpatient appointment systems. **Communications of the ACM**, v. 12, n. 4, p. 215-222, 1969.

KC, Diwas Singh; TERWIESCH, Christian. An econometric analysis of patient flows in the cardiac intensive care unit. **Manufacturing & Service Operations Management**, v. 14, n. 1, p. 50-65, 2012.

KELLER, T. F.; LAUGHUNN, D. J. AN APPLICATION OF QUEUING THEORY TO A CONGESTION PROBLEM IN AN OUTPATIENT CLINIC*. **Decision Sciences**, v. 4, n. 3, p. 379-394, 1973.

KLASSEN, Kenneth J.; ROHLEDER, Thomas R. Scheduling outpatient appointments in a dynamic environment. **Journal of Operations Management**, v. 14, n. 2, p. 83-101, 1996.

LASDON, Leon S.; FOX, Richard L.; RATNER, Margery W. Nonlinear optimization using the generalized reduced gradient method. **Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle**, v. 8, n. 3, p. 73-103, 1974.

LAU, Hon-Shiang; LAU, Amy Hing-Ling. A fast procedure for computing the total system cost of an appointment schedule for medical and kindred facilities. **IIE Transactions**, v. 32, n. 9, p. 833-839, 2000.

LEHANEY, B.; CLARKE, S. A.; PAUL, R. J. A case of an intervention in an outpatients department. **Journal of the Operational Research Society**, p. 877-891, 1999.

LEUNG, Joseph YT (Ed.). **Handbook of scheduling: algorithms, models, and performance analysis**. CRC Press, 2004.

LIAO, CHTNG-JONG; PEGDEN, C. DENNIS; ROSENSHINE, MATTHEW. Planning timely arrivals to a stochastic production or service system. **IIE Transactions**, v. 25, n. 5, p. 63-73, 1993.

LINDLEY, David V. The theory of queues with a single server. In: **Mathematical Proceedings of the Cambridge Philosophical Society**. Cambridge University Press, 1952. p. 277-289.

LIU, L. LIU, Xiaoming. Block appointment systems for outpatient clinics with multiple doctors. **Journal of the Operational Research Society**, v. 49, n. 12, p. 1254-1259, 1998a.

LIU, Liming; LIU, Xiaoming. Dynamic and static job allocation for multi-server systems. **IIE Transactions**, v. 30, n. 9, p. 845-854, 1998b.

MA, Guoxuan; DEMEULEMEESTER, Erik. A multilevel integrative approach to hospital case mix and capacity planning. **Computers & Operations Research**, 2012.

MAHACHEK, Arnold R.; KNABE, Terry L. Computer simulation of patient flow in obstetrical/gynecology clinics. **Simulation**, v. 43, n. 2, p. 95-101, 1984.

MCDERMOTT, Christopher; STOCK, Gregory N. Hospital operations and length of stay performance. **International Journal of Operations & Production Management**, v. 27, n. 9, p. 1020-1042, 2007.

MERCER, A. A queueing problem in which the arrival times of the customers are scheduled. **Journal of the Royal Statistical Society. Series B (Methodological)**, p. 108-113, 1960.

MERCER, A. Queues with scheduled arrivals: A correction, simplification and extension. **Journal of the Royal Statistical Society. Series B (Methodological)**, p. 104-116, 1973.

MEZA, James P. Patient waiting times in a physician's office. **The American Journal of Managed Care**, v. 4, n. 5, p. 703-712, 1998.

MUTHURAMAN, Kumar; LAWLEY, Mark. A stochastic overbooking model for outpatient clinical scheduling with no-shows. **IIE Transactions**, v. 40, n. 9, p. 820-837, 2008.

O'KEEFE, Robert M. Investigating outpatient departments: implementable policies and qualitative approaches. **Journal of the Operational Research Society**, p. 705-712, 1985.

OLESKOVICZ, Marcelo et al. Técnica de overbooking no atendimento público ambulatorial em uma unidade do Sistema Único. **Cad. Saúde Pública**, v. 30, n. 5, p. 1009-1017, 2014.

OLIVEIRA, F. G. F. T.; FONSECA, L. M. B.; KOCH, Hilton Augusto. Responsabilidade civil do radiologista no diagnóstico do câncer de mama através do exame de mamografia. **Radiol Bras**, v. 44, p. 183-7, 2011.

PATRICK, Jonathan. Access to Long- Term Care: The True Cause of Hospital Congestion?. **Production and Operations Management**, v. 20, n. 3, p. 347-358, 2011.

PAUL, Ray J.; KULJIS, Jasna. A generic simulation package for organising outpatient clinics. In: **Proceedings of the 27th Conference on Winter Simulation**. IEEE Computer Society, 1995. p. 1043-1047.

PEDROSO, Marcelo Caldeira; MALIK, Ana Maria. Cadeia de valor da saúde: um modelo para o sistema de saúde brasileiro. **Ciênc. Saúde Coletiva**, v. 17, n. 10, p. 2757-2772, 2012.

PEGDEN, Claude Dennis; ROSENSHINE, Matthew. Scheduling arrivals to queues. **Computers & Operations Research**, v. 17, n. 4, p. 343-348, 1990.

PIERSKALLA, William P.; BRAILER, David J. Applications of operations research in health care delivery. **Handbooks in Operations Research and Management Science**, v. 6, p. 469-505, 1994.

RAMWADHDOEBE, Sabrina et al. A tutorial on discrete-event simulation for health policy design and decision making: Optimizing pediatric ultrasound screening for hip dysplasia as an illustration. **Health Policy**, v. 93, n. 2, p. 143-150, 2009.

RAU, Chi-Lun et al. Using discrete-event simulation in strategic capacity planning for an outpatient physical therapy service. **Health Care Management Science**, p. 1-14, 2013

RECHEL, Bernd et al. Hospital capacity planning: from measuring stocks to modelling flows. **Bulletin of the World Health Organization**, v. 88, n. 8, p. 632-636, 2010.

RISING, Edward J.; BARON, Robert; AVERILL, Barry. A systems analysis of a university-health-service outpatient clinic. **Operations Research**, v. 21, n. 5, p. 1030-1047, 1973.

ROBINSON, Lawrence W.; CHEN, Rachel R. Scheduling doctors' appointments: optimal and empirically-based heuristic policies. **IIE Transactions**, v. 35, n. 3, p. 295-307, 2003.

ROCKART, John F.; HOFMANN, Paul B. Physician and patient behavior under different scheduling systems in a hospital outpatient department. **Medical Care**, v. 7, n. 6, p. 463-470, 1969.

ROCHE, Kevin T.; RIVERA, Daniel E.; COCHRAN, Jeffery K. A control engineering framework for managing whole hospital occupancy. **Mathematical and Computer Modelling**, v. 55, n. 3, p. 1401-1417, 2012.

ROHLEDER, Thomas R.; KLASSEN, Kenneth J. Using client-variance information to improve dynamic appointment scheduling performance. **Omega**, v. 28, n. 3, p. 293-302, 2000.

ROMERO, H. L. et al. Admission and capacity planning for the implementation of one-stop-shop in skin cancer treatment using simulation-based optimization. **Health Care Management Science**, v. 16, n. 1, p. 75-86, 2013.

SCHAFER, W. B., "Keep Patients Waiting? Not in My Office," **Medical Economics**, 63, 10 (May 12), 137-141, 1986

SMITH-DANIELS, Vicki L.; SCHWEIKHART, Sharon B.; SMITH-DANIELS, Dwight E. Capacity Management in Health Care Services: Review and Future Research Directions*. **Decision Sciences**, v. 19, n. 4, p. 889-919, 1988.

SORIANO, Alfonso. Comparison of two scheduling systems. **Operations Research**, v. 14, n. 3, p. 388-397, 1966.

SWARTZMAN, Gordon. The patient arrival process in hospitals: statistical analysis. **Health Services Research**, v. 5, n. 4, p. 320, 1970.

SWISHER, James R. et al. Modeling and analyzing a physician clinic environment using discrete-event (visual) simulation. **Computers & Operations Research**, v. 28, n. 2, p. 105-125, 2001.

TAI, Guangfu; WILLIAMS, Peter. Quantitative analysis of investment allocation over various resources of health care systems by using views of product lines. **International Journal of Systems Science**, n. ahead-of-print, p. 1-11, 2012.

TAYLOR, Bruce. Patient use of a mixed appointment system in an urban practice. **British Medical Journal (Clinical Research Ed.)**, v. 289, n. 6454, p. 1277-1278, 1984.

- TSAI, Pei-Fang Jennifer; TENG, Guei-Yu. A stochastic appointment scheduling system on multiple resources with dynamic call-in sequence and patient no-shows for an outpatient clinic. **European Journal of Operational Research**, 2014.
- VAN DIJK, Nico M.; KORTBEEK, Nikky. Erlang loss bounds for OT-ICU systems. **Queueing Systems**, v. 63, n. 1-4, p. 253-280, 2009.
- VANDEN BOSCH, Peter M.; DIETZ, Dennis C. Minimizing expected waiting in a medical appointment system. **IIE Transactions**, v. 32, n. 9, p. 841-848, 2000.
- VANDEN BOSCH, Peter M.; DIETZ, Dennis C.; SIMEONI, John R. Scheduling customer arrivals to a stochastic service system. **Naval Research Logistics (NRL)**, v. 46, n. 5, p. 549-559, 1999.
- VILLEGAS, E. L. "Outpatient Appointment System Saves Time for Patients and Doctors," **Hospitals**, J. A. H. A., 41, 52-57, 1967
- VIRJI, Abbas. A study of patients attending without appointments in an urban general practice. **BMJ: British Medical Journal**, v. 301, n. 6742, p. 22, 1990.
- VISSERS, J. Selecting a suitable appointment system in an outpatient setting. **Medical Care**, p. 1207-1220, 1979.
- VISSERS, Jan MH; BERTRAND, J. W. M.; DE VRIES, G. A framework for production control in health care organizations. **Production Planning & Control**, v. 12, n. 6, p. 591-604, 2001.
- VISSERS, J.; WIJNGAARD, Jacob. The outpatient appointment system: Design of a simulation study. **European Journal of Operational Research**, v. 3, n. 6, p. 459-463, 1979.
- WALTER, S. D. A comparison of appointment schedules in a hospital radiology department. **British Journal of Preventive & Social Medicine**, v. 27, n. 3, p. 160-167, 1973.
- WANG, P. Patrick. Static and dynamic scheduling of customer arrivals to a single-server system. **Naval Research Logistics (NRL)**, v. 40, n. 3, p. 345-360, 1993.
- WANG, P. Patrick. Optimally scheduling N customer arrival times for a single-server system. **Computers & Operations Research**, v. 24, n. 8, p. 703-716, 1997.

WELCH, J. D. Appointment systems in hospital outpatient departments. **OR**, p. 224-232, 1964.

WELCH, J. D.; BAILEY, NormanT J. Appointment systems in hospital outpatient departments. **The Lancet**, v. 259, n. 6718, p. 1105-1108, 1952.

WEISS, Elliott N. Models for determining estimated start times and case orderings in hospital operating rooms. **IIE Transactions**, v. 22, n. 2, p. 143-150, 1990.

WESTMAN, Göuran et al. Waiting room time in the assessment of an appointment system in primary care. **Scandinavian Journal of Primary Health Care**, v. 5, n. 1, p. 35-40, 1987.

WILLIAMS, William J.; COVERT, R. P.; STEELE, J. D. Simulation modeling of a teaching hospital outpatient clinic. **Hospitals**, v. 41, n. 21, p. 71-5 passim, 1967.

YANG, Kum Khiong; LAU, Mun Ling; QUEK, Ser Aik. A new appointment rule for a single-server, multiple-customer service system. **Naval Research Logistics (NRL)**, v. 45, n. 3, p. 313-326, 1998.

YIN, Robert K. **Estudo de Caso-: Planejamento e Métodos**. Bookman editora, 2015.