

## **MINERACAO DE DADOS TEMPORAL NA GESTÃO DE EXAMES LABORATORIAIS EM HOSPITAL PÚBLICO**

Leila Cristina Dufлот (Especialista em Gestão Pública de Tecnologia da Informação)  
Deborah Ribeiro Carvalho (Doutora em Computação, PUCPR) E-mail: ribeiro.carvalho@pucpr.br  
Elaine Rossi Ribeiro (Doutora em Medicina)  
Rosyellen Rabelo Szvarça (Mestranda em Tecnologia em Saúde)

**Resumo:** Os custos operacionais de serviços hospitalares representam cada vez mais um desafio para os gestores. O objetivo é identificar padrões e reincidências nas solicitações de exames laboratoriais. Foi aplicada mineração de dados temporal sobre um conjunto de exames laboratoriais de 2015, considerando duas janelas temporais, 0 a 2 dias e 3 a 10 dias. Sobre o conjunto de 192.151 exames, seis representam 52% e dezoito correspondem a 85% do total de exames. Considerando as recorrências entre as demandas, 12% aconteceram entre o primeiro e o segundo dia de internação e 61% aconteceram até o décimo dia de internação. A reincidência é maior entre 0 a 2 dias. Os padrões descobertos contribuíram na identificação da aderência do corpo clínico às rotinas institucionais, na possibilidade de criação de novas rotinas que orientem a solicitação de exames e na avaliação e reestruturação da política de custeio e/ou contratação dos exames mais requisitados.

**Palavras-chave:** Mineração de Dados Temporal; Gestão em Saúde; Testes Laboratoriais.

## **IDENTIFICATION OF CARDIOVASCULAR DISEASES BENEFICIARIES FOR MANAGEMENT PROGRAMS**

**Abstract:** Hospital services operating costs have become increasingly a challenge for health managers. The goal was to identify patterns and recidivism in laboratory test requests. Temporal data mining was applied on laboratory tests database, carried out in 2015. It was processed two temporal intervals: 0-2 days and 3-10 days. Considering the set of 192,151 tests, six represent 52% and eighteen correspond to 85% of all tests. Almost 12% of exam's occurrences took place between the first and second day of hospitalization and 61% were carried out by the tenth day of hospitalization. It was evidenced that the recurrence is greater between 0 to 2 days. The patterns discovered contributed in identifying the adherence of clinical staff to institutional routines, the possibility of creating new routines to guide the request for laboratory tests as well as for the evaluation and restructuring of the funding policy and/or hiring of the most required tests.

**Keywords:** Temporal Data Mining; Health Management; Laboratory Tests.

### **1. INTRODUÇÃO**

A gestão de serviços de saúde é um grande desafio, ainda mais quando se trata de serviços hospitalares, considerando que estes sofrem grandes influências das inovações tecnológicas, adaptam-se e readaptam-se constantemente às transformações que a tecnologia provoca.

Castillo e Jacobo (2014) são categóricos ao afirmarem que a tecnologia veio para, de maneira essencial, modificar as relações profissionais. Os gestores hospitalares buscam alternativas tecnológicas para identificar, conhecer e controlar seus custos, melhorar sua eficiência e eliminar desperdícios.

A constante necessidade de atualização tecnológica, estrutural e de processos de trabalho, se de um lado trazem melhorias nos serviços prestados, por outro lado representam um aumento nos custos gerenciais. Por se tratar basicamente de uma prestação de serviços os gestores hospitalares precisam ir além do que conhecer a composição de seus custos e a formação dos preços.

Neste sentido cabe ao gestor um olhar analítico e atento a todas as áreas da instituição, focado nos processos de produção, no consumo, informações estas, preferencialmente, estruturadas num sistema apurado de custos, de forma a dar condições ao gestor de tomar decisões estratégicas que aperfeiçoem a utilização dos recursos sem afetar a qualidade dos serviços, principalmente ao se considerar que os recursos financeiros e materiais disponíveis são escassos e onerosos; que os exames laboratoriais representam uma despesa relevante a ser acompanhada e otimizada; que a receita gerada pelos exames laboratoriais realizados é insuficiente para cobrir os custos dispendidos para sua realização; poder-se-ia dizer que estes são os fatores basilares que justificam uma pesquisa sobre mineração de dados relacionados aos custos versus exames laboratoriais.

Ao iniciar-se este estudo, alguns descritores foram utilizados nas buscas de informações relevantes nas principais bases de dados científicas, tais como: mineração de dados; gestão em saúde; testes laboratoriais. A relevância do assunto é de tamanha extensão, que somente o número de artigos nacionais ultrapassa a casa dos milhares, somente nos últimos cinco anos. Os temas são os mais variados, desde o uso da mineração de dados em oncologia, radiologia, risco cardiovascular, febre hemorrágica, entre outros. No entanto, poucos são aqueles que relacionam mineração de dados com custos hospitalares de forma direta, com foco econômico-financeiro.

Baltzan (2016) explicita “mineração de dados (*data mining*) é o processo de análise de dados para extrair informações que não são fornecidas apenas pelos dados brutos.” Afirma ainda que “a mineração também pode começar em um nível de informação sumária (granularidade grossa) e progredir em níveis crescentes de detalhes (*drill down*) ou o contrário (*drill up*). Para executar a mineração de dados, os usuários precisam de ferramentas adequadas. As ferramentas de mineração de dados utilizam diversas técnicas para encontrar padrões e relações em grandes volumes de informações e, a partir deles, inferem regras que preveem um comportamento futuro e guiam a tomada de decisão”.

Com o objetivo de demonstrar como a aplicação de técnicas de mineração de dados permite ancorar o apoio à decisão, Gomes et al (2014) apresentam um estudo no Serviço de Atendimento de Urgências, destacando as demandas sucessivas ao serviço. Chaves, Silvério e Barroso (2015) relacionam a mineração de dados à incidência de HIV/Aids em determinadas microrregiões mineiras, o que resultou na qualificação do atendimento á saúde daquela localização.

Segundo Conceição e Lima (2013), a mineração de dados pode ser realizada de três diferentes formas, e isso ocorre em função do nível de conhecimento que se tem do problema estudado. Quando se desconhece o comportamento do fenômeno, pode-se simplesmente descobrir relações “novas” escondidas, e que a olho nu não poderiam ser localizadas com facilidade, a partir da mineração de dados. Esse método recebe o nome de descoberta não supervisionada de relações.

Entretanto, quando se possui algum conhecimento sobre o campo de atuação da empresa ou alguma ideia sobre que relação nova se está buscando, pode-se definir uma hipótese e verificar sua confirmação por meio da mineração de dados. E finalmente quando se tem um nível maior de conhecimento da área e da relação que se deseja estudar, procede-se com a metodologia de modelagem matemática de dados.

Neste sentido vislumbra-se uma oportunidade de utilização da ferramenta de mineração de dados no conjunto de informações clínicas e gerenciais disponíveis na instituição, mais especificamente no conjunto de dados relacionados às requisições de exames laboratoriais.

Assim sendo, este estudo apresenta a seguinte questão norteadora: A mineração de dados na prescrição de exames laboratoriais pode, em sua eficácia, otimizar a gestão de custos operacionais?

Para responder a esta questão, este artigo objetiva identificar padrões nas solicitações de exames laboratoriais em um hospital público. Traz como contribuição evidenciar que o uso da mineração de dados pode auxiliar na gestão de custos operacionais relacionados a indicação de exames laboratoriais.

## **2. MATERIAIS E MÉTODOS**

Trata-se de uma pesquisa quantitativa, do tipo exploratória-descritiva.

Para a realização da pesquisa foi utilizado um banco de dados de um Hospital Municipal, de porte secundário, com atendimento prioritário ao idoso, na cidade de Curitiba, que expediu carta de ciência autorizando estudo em questão, para utilização de informações não relativas à identificação de pacientes, utilizando banco de dados secundários, nos termos da Resolução n. 466/2012.

Foram selecionados dados referentes ao ano de 2015, os quais foram compostos por exames laboratoriais realizados em pacientes internados, sua respectiva data de solicitação e o tempo das internações. As variáveis selecionadas foram: número de atendimento; data de internação; data de alta hospitalar; tempo de internação; exame solicitado; data da solicitação do exame.

O critério adotado para a seleção das variáveis cumpriu os objetivos deste estudo no sentido de identificar a frequência de realização de exames, sua recorrência e faixa temporal de recorrência.

A análise dos dados foi implementada em duas etapas: estatísticas gerenciais e mineração de dados. Para a estatística gerencial foi analisada a série histórica de internações e exames laboratoriais realizados no ano de 2015, o tempo médio das internações e, além disso, buscaram-se correlacionar os gastos dispendidos com os exames realizados e quanto estes valores corresponderam do valor total recebido do Sistema Único de Saúde.

Para a mineração de dados foi adotada a tarefa de descoberta de regras de associação a partir do algoritmo APRIORI (BORGELT, 2014) e para a identificação das janelas temporais o algoritmo ASSOCTEMP (SOKOLOSKI, CARVALHO, DALLAGASSA, 2014). A escolha pela tarefa de associação deve-se ao interesse em identificar a reincidência de solicitação de exames dentro de um período de tempo determinado.

Por exemplo, a regra R1, descoberta a partir do algoritmo APRIORI (BORGELT, 2014), permite evidenciar que 75,5% realizou dosagem de sódio, dos quais 95,7% também realizou Dosagem de Creatinina. A R2 evidencia que 77,2% realizou Dosagem de Potássio, dos quais 96,8% também realizou Dosagem de Sódio.

R1: DOSAGEM DE SODIO → DOSAGEM DE CREATININA (76.5%, 95.7%)

R2: DOSAGEM DE POTASSIO → DOSAGEM DE SODIO (77.2%, 96.8%)

Porém, como o algoritmo APRIORI (BORGELT, 2014) foi proposto para associar elementos, com objetivo mais focado na gestão de negócios, a partir das associações descobertas não é possível inferir a sequência na qual os elementos ocorreram, por exemplo,

quantas Dosagens de Sódio antecederam a Dosagem de Creatinina. Bem como não é possível evidenciar a repetição do mesmo elemento, no caso desta pesquisa, solicitação de exames.

Para tratar esta limitação da proposta original da tarefa de descoberta de regras de associação, o algoritmo ASSOCTEMP (SOKOLOSKI, CARVALHO, DALLAGASSA, 2014) pós-processa as regras de associação descobertas indetificando sequência e a janela temporal entre os elementos associados.

Para a seleção dos exames com vistas ao processo de mineração, foi adotada a frequência relativa de demanda por exames e seu respectivo custo total. Como linha de corte foram selecionados os exames que representassem as maiores frequências até atingir em torno de 80% dos gastos, os quais são especificados à seguir: Potássio, Creatinina, Sódio, Ureia, Hemograma Completo, Proteína C Reativa, Gasometria, Cultura de Bactérias, Antibiograma, Lactato, Protrombina (Tap), Tromboplastina Parcial Ativada (Ttp Ativada), Transaminase Glutâmico-Oxalacética (Tgo), Transaminase Glutâmico-Pirúvica (Tgp), Bilirrubina total e frações, Urina, Glicose e Hemocultura.

Como janela temporal foi adotada entre 0 a 2 dias e de 3 a 10 dias. Essa escolha se deve ao fato de que nos primeiros dois dias de internação é realizada a investigação diagnóstica para subsidiar o plano terapêutico do paciente e a partir do terceiro dia de internação busca-se evidenciar e/ou confirmar outras hipóteses diagnósticas, ademais, observou-se o tempo médio das internações.

### **3. RESULTADOS E DISCUSSÃO**

Entre as estatísticas gerenciais, destaca-se que o Hospital em questão, realizou 192.151 exames laboratoriais, distribuídos em 2.918 internações, uma média de 66 exames por internação e as internações tiveram um tempo médio de 11 dias, em 2015. Do ponto de vista financeiro, o total de internações gerou uma receita de R\$5.965.365. Os exames laboratoriais geraram um gasto de R\$887.326,13, representado 15% da receita.

A partir da Tabela 1, é possível observar a relevância dos seis primeiros tipos de exames mais requisitados, os quais representaram 52% do total. Já considerando os dezoito exames mais requisitados, estes representaram 85% do total. Porém os seis exames mais requisitados representaram 35% do total dos gastos dos exames, e os dezoito exames, representaram 86% do total. Pode-se afirmar que os dezoito exames mais requisitados representam 85% dos exames realizados e 86% dos gastos dispendidos.

Considerando os 18 exames mais requisitados, a respectiva frequência de solicitação e o tempo de internação, apresentam-se as seguintes situações: Em média, cada paciente realizou 14 dos 18 tipos; a menor frequência de realização de exames foi de um, ocorrendo com seis pacientes, na sua maioria com internação de um 1 dia. Exceção ocorreu com um paciente internado por sete dias; 12% dos pacientes realizaram os 18 tipos, não apresentando padrão no tempo de internação dado que variou de 1 a 121 dias; 49% dos pacientes realizaram de 14 a 18 tipos, também sem apresentar padrão em relação ao tempo de internação; 12% do total dos exames foram realizados entre o primeiro e o segundo dia de internação; 61% do total dos exames foram realizados entre 1 e 10 dias de internação; 30% do total dos exames foram realizados 11 e 30 dias de internação; 9% do total dos exames foram realizados entre 31 e 230 dias de internação.

**Tabela 1** - Frequência absoluta e relativa de demanda, valor unitário, total e percentual, por exames realizados em 2015.

Código	Exames realizados em 2015	Quantidade	%	Valor Un.	Valor Total	%
202010600	Dosagem De Potassio	17.940	9,34%	1,85	33.189,00	3,74%
202010317	Dosagem De Creatinina	17.915	9,32%	1,85	33.142,75	3,74%
202010635	Dosagem De Sodio	17.688	9,21%	1,85	32.722,80	3,69%
202010694	Dosagem De Ureia	17.561	9,14%	1,85	32.487,85	3,66%
202020380	Hemograma Completo	16.999	8,85%	4,11	69.865,89	7,87%
202030083	Determinacao Quantitativa De Proteina C Reativa	11.880	6,18%	9,25	109.890,00	12,38%
202010732	Gasometria (Ph Pco2 Po2 Bicarbonato As2 (Exceto Base )	9.636	5,01%	15,65	150.803,40	17,00%
202080080	Cultura De Bacterias P/ Identificacao	9.564	4,98%	5,62	53.749,68	6,06%
202080021	Antibiograma C/ Concentracao Inibitoria Minima	9.227	4,80%	13,33	122.995,91	13,86%
202010538	Dosagem De Lactato	6.477	3,37%	3,68	23.835,36	2,69%
202020142	Determinacao De Tempo E Atividade Da Protrombina (Tap)	4.976	2,59%	2,73	13.584,48	1,53%
202020134	Determinacao De Tempo De Tromboplastina Parcial Ativada (Ttp Ativada)	4.416	2,30%	5,77	25.480,32	2,87%
202010643	Dosagem De Transaminase Glutamico-Oxalacetica (Tgo)	3.459	1,80%	2,01	6.952,59	0,78%
202010651	Dosagem De Transaminase Glutamico-Piruvica (Tgp)	3.448	1,79%	2,01	6.930,48	0,78%
202010201	Dosagem De Bilirrubina Total E Fracoes	3.419	1,78%	2,01	6.872,19	0,77%
202050017	Analise De Caracteres Fisicos, Elementos E Sedimento Da Urina	3.385	1,76%	3,70	12.524,50	1,41%
202010473	Dosagem De Glicose	3.130	1,63%	1,85	5.790,50	0,65%
202080153	Hemocultura	1.915	1,00%	11,49	22.003,35	2,48%
Outros (207 exames diversos que representam menos de 1%)		29.116	15,15%		124.505,05	14,03%
Total		192.151	100%		887.326,13	100%

Fonte: Hiza (2015)

A tabela 2 evidencia a frequência de repetições pós-processadas pelo algoritmo ASSOCTEMP (SOKOLOSKI, CARVALHO, DALLAGASSA, 2014) ocorridas nos exames sob a óptica de dois determinados períodos, podendo-se afirmar que 18,70% das repetições de dosagem de potássio ocorreram entre 0 a 2 dias e 41,60% entre 3 a 10 dias, da mesma forma que 19% das repetições de dosagem de creatinina ocorreram entre 0 a 2 dias e 41,60% entre 3 a 10 dias, e, assim sucessivamente.

**Tabela 2** - Frequência de realização de exames dentro de determinada janela temporal

Frequência de realização de exames		Janela temporal 0 a 2 dias	Janela temporal 3 a 10 dias
Código	Descrição do Exame	Percentual	Percentual
202010600	Dosagem De Potassio	18,70%	41,60%
202010317	Dosagem De Creatinina	19,00%	41,60%
202010635	Dosagem De Sodio	19,10%	41,90%
202010694	Dosagem De Ureia	19,20%	42,20%
202020380	Hemograma Completo	17,90%	40,90%
202030083	Determinacao Quantitativa De Proteina C Reativa	17,80%	40,30%
202010732	Gasometria (Ph Pco2 Po2 Bicarbonato As2 (Exceto Base )	20,80%	48,00%
202080080	Cultura De Bacterias P/ Identificacao	79,80%	0,40%
202080021	Antibiograma C/ Concentracao Inibitoria Minima	33,60%	30,70%
202010538	Dosagem De Lactato	21,40%	41,60%
202020142	Determinacao De Tempo E Atividade Da Protrombina (Tap)	26,60%	44,00%
202020134	Determinacao De Tempo De Tromboplastina Parcial Ativada (Ttp Ativada)	27,00%	42,70%
202010643	Dosagem De Transaminase Glutamico-Oxalacetica (Tgo)	28,30%	43,60%
202010651	Dosagem De Transaminase Glutamico-Piruvica (Tgp)	28,20%	43,60%
202010201	Dosagem De Bilirrubina Total E Fracoes	25,90%	43,80%
202010473	Dosagem De Glicose	27,20%	38,20%
202050017	Analise De Caracteres Fisicos, Elementos E Sedimento Da Urina	27,70%	40,70%
202080153	Hemocultura	19,90%	41,10%

Com relação ao índice de frequência pode-se afirmar que na janela temporal de 3 a 10 dias houve uma redução que variou de 34% a 60% em relação à frequência realizada na janela temporal de 0 a 2 dias. Exceção para dois casos, que apresentaram drástica redução, 74% e 99%.

Dentre os 18 exames selecionados destaca-se um exame onde ocorreu maior redução na frequência de realização, trata-se da cultura de bactéria para identificação, exame este que faz parte da rotina de admissão de pacientes admitidos por transferência. O resultado encontrado demonstra a aderência do corpo clínico à rotina institucional, uma vez que esta estipula que as culturas devem ocorrer até o segundo dia de internação.

#### **4. CONCLUSÕES**

O estudo revelou que embora a instituição tenha, no ano de 2015, realizado 207 tipos de exames laboratoriais e total de 192.151 exames, os seis exames mais solicitados correspondem a 52% e dezoito mais solicitados correspondem a 85% do total de exames realizados.

A aplicação do algoritmo de mineração de dados ASSOCTEMP (SOKOLOSKI, CARVALHO, DALLAGASSA, 2014) nestes dezoito tipos de exames mais solicitados, possibilitou a visualização da temporalidade e descoberta de padrões nas ocorrências de exames laboratoriais.

Um padrão evidenciado é que a reincidência de realização destes exames é maior na janela temporal de 0 a 2 dias, o que vem ao encontro da prática clínica, uma vez que na admissão dos pacientes é onde se concentra o maior número de solicitação de exames, a fim de estruturar o plano terapêutico do paciente.

Outro padrão evidenciado foi de que as reincidências de realização de exames ocorridas na janela temporal de 3 a 10 dias, não evidenciaram um percentual médio de redução na frequência de realização, uma vez que apresentou grande variabilidade, o que sugere um estudo mais aprofundado visando a criação de protocolos que orientem a solicitação destes exames.

Além dos padrões anteriormente citados, outro padrão foi encontrado no exame “cultura de bactéria para identificação”, que revelou uma baixíssima reincidência de ocorrência na janela temporal de 3 a 10 dias. Este fato comprova que o exame está sendo solicitado na admissão do paciente, ou seja, entre o primeiro e segundo dia de internação, observando a rotina institucional, que definiu este exame como obrigatório para todos os pacientes admitidos oriundos de outros serviços de saúde.

Os resultados da mineração de dados, contribuíram para avaliar a aderência do corpo clínico às rotinas institucionais estabelecidas, a possibilidade de criação de novas rotinas que orientem a solicitação de exames, bem como para a avaliação e reestruturação da política de custeio e/ou contratação dos exames mais requisitados.

Os dados revelados no estudo podem contribuir para a realização de outros estudos, para a implementação de protocolos clínicos, para o processo de educação permanente do corpo clínico, bem como, para a área administrativa no sentido de subsidiar políticas de compras e/ou contratação deste serviço.

No sentido de limitações e fragilidades deste estudo, o mesmo foi realizado abrangendo todas as áreas da instituição, porém a escolha por áreas críticas, como por exemplo, Unidade de Terapia Intensiva, poderia trazer resultados mais específicos.

Enfim, pode-se afirmar que o processo de mineração de dados demonstrou ser relevante para a gestão clínica e de primordial importância para complementar as informações geradas por sistemas de gestão informatizados. Sob o ponto de vista de gestão em saúde, os resultados sugerem novos meios de análise e tomada de decisão que venham apoiar medidas de monitoramento e padronização da prescrição e realização dos exames.

## **REFERÊNCIAS**

BALTZAN P. **Tecnologia orientada para gestão** [internet]. 6.ed. Porto Alegre: AMGH; 2016 [acesso em 2016 março 16]. Disponível em: [https://play.google.com/books/reader?id=pUCPCwAAQBAJ&printsec=frontcover&output=reader&hl=pt\\_BR&pg=GBS.PR3](https://play.google.com/books/reader?id=pUCPCwAAQBAJ&printsec=frontcover&output=reader&hl=pt_BR&pg=GBS.PR3).

BORGELT C. **Apriori** [internet]. 2014. [acesso em 2014 dez 15]. Disponível em: <http://www.borgelt.net/apriori.html>.

CASTILLO A.F.E, JACOBO B.O.M. Use of information technologies by the plastic surgeons of Uruguay. A national online survey. **Cir Plást Iberolatinoam**. 2014;40(4): 429-435

CHAVES A.R.C., SILVÉRIO T.A.B., BARROSO V.A. Pré-processamento para mineração de dados de pacientes com HIV. **J Health Inform**. V.7, n 1, p. 3-7, 2015

CONCEIÇÃO G.C., LIMA A.B. (2013). **Data mining como ferramenta de apoio a análise mercadológica** [internet]. Taquaratinga: Editora Anderson Barbosa de Lima; 2013.[acesso em 2016 março 23]. Disponível em: [https://play.google.com/books/reader?id=WnNJBQAAQBAJ&printsec=frontcover&output=reader&hl=pt\\_BR&pg=GBS.PA2](https://play.google.com/books/reader?id=WnNJBQAAQBAJ&printsec=frontcover&output=reader&hl=pt_BR&pg=GBS.PA2).

GOMES D.D.C., CARVALHO D.R., CUBAS M.R., SHMEIL M.A.H. Mineração de dados no serviço de atendimento de urgências. **J Health Inform**. V.6,n 4, p. 141-146; 2014

SOKOLOSKI W.F., CARVALHO D.R., DALLAGASSA M.R. Regra de Associação Temporal. In: **Anais do XIV Congresso Brasileiro de Informática em Saúde**. Santos: CBIS. 2014. Disponível em: [http://www.sbis.org.br/biblioteca\\_virtual/cbis/Anais\\_CBIS2014\\_Completo.pdf](http://www.sbis.org.br/biblioteca_virtual/cbis/Anais_CBIS2014_Completo.pdf)