

Experiências internacionais da aplicação de sistemas de apoio à decisão clínica em gastroenterologia

International outcomes from attempts to implement a clinical decision support system in gastroenterology

Experiencias internacionales de la aplicación de sistemas de apoyo a la decisión clínica en gastroenterología

Josceli Maria Tenório¹, Anderson Diniz Hummel¹, Vera Lucia Sdepanian², Ivan Torres Pisa³,
Heimar de Fátima Marin⁴

RESUMO

Descritores: Sistemas de apoio a decisões clínicas, inteligência artificial, gastroenterologia

Objetivo: Descrever as experiências recentes com a aplicação de sistemas de apoio à decisão clínica em gastroenterologia, de forma a estabelecer o nível de desenvolvimento, testes e vantagens conferidas à prática médica com a introdução desses softwares. **Métodos:** Foi realizada busca nas bases de dados PubMed, LILACS e ISI Web of Knowledge, utilizando termos relacionados à sistemas de apoio à decisão e à gastroenterologia, incluindo artigos originais publicados no período entre 2005 e 2010. Foram recuperadas 104 publicações, na busca inicial e, após a aplicação dos critérios de inclusão e exclusão, foram eleitos nove estudos para leitura do texto completo. **Resultados:** Os sistemas de apoio à decisão clínica apresentam grande multiplicidade de problemas clínicos e investigação de doenças. Em 89% dos casos, são descritos modelos experimentais para o desenvolvimento de sistemas de apoio à decisão clínica. A descrição dos resultados obtidos por técnicas de inteligência artificial em 78% das publicações. Em dois dos estudos foram realizadas comparações com o médico e em apenas uma publicação um estudo controlado foi descrito, mostrando evidências de melhorias na prática médica. **Conclusão:** Os estudos mostram potenciais benefícios dos sistemas de apoio à decisão clínica à prática médica, porém, estudos controlados em ambiente real devem ser realizados para comprovar esta perspectiva.

ABSTRACT

Keywords: Decision support systems, clinical; artificial intelligence; gastroenterology

Objective: This study aimed at describing the recent experience acquired with the implementation and use of clinical decision support system in gastroenterology in order to determine the level of development, tests used and advantages that such a system can offer to the medical practice. **Methods:** A search in the PubMed, LILACS and ISI Web of Knowledge databases for studies in decision-making support systems in gastroenterology including original papers produced from 2005 to 2010 was performed. A total of 104 scientific papers were retrieved initially. These were analyzed using inclusion and exclusion criteria, thus yielding nine studies for further analysis. **Results:** The clinical decision support system analyzed in the present study showed a great variety of clinical problems regarding the investigation of a disease and the determination of a diagnosis. Eighty-nine per cent of the studies showed experimental models for clinical decision support system development. Seventy-eight per cent of the studies described the outcomes obtained with artificial intelligence technique. Two studies compared the clinical decision support system performance with that of a doctor, and only one research work described a controlled study evidencing improvements in the medical practice. **Conclusion:** The studies analyzed showed evidence of potential benefits that clinical decision support system can bring to the clinical practice. However, further controlled studies performed in medical day-to-day conditions and environment should be performed in order to provide more clear evidence of the usefulness of clinical decision support system in the medical practice.

RESUMEN

Descriptores: Sistemas de apoyo a decisiones clínicas; gastroenterología, y establecer el nivel de desarrollo, pruebas y beneficios concedidos a la práctica médica con la introducción de dicho software. **Métodos:** Se realizó una búsqueda en bases de datos PubMed, LILACS e ISI Web of Knowledge el uso de términos relacionados con los sistemas de apoyo a toma de decisiones y gastroenterología, incluyendo artículos originales publicados entre 2005 y 2010. Se recuperaron 104 publicaciones en la búsqueda inicial y después de la aplicación de criterios de inclusión y exclusión, nueve estudios fueron elegidos para leer el texto completo. **Resultados:** Hay una gran variedad de sistemas de apoyo a la toma de decisiones clínicas sobre los más diversos problemas clínicos y de investigación de las enfermedades. En el 89% de los casos se describen los modelos experimentales para el desarrollo de los sistemas de apoyo a la toma de decisiones clínicas. En el 78% de las publicaciones hay descripción de los resultados obtenidos por técnicas de inteligencia artificial. En dos estudios, las se hicieron comparaciones del sistema con el médico y solamente una única publicación, un estudio controlado describe la evidencia de las mejoras en la práctica médica. **Conclusión:** Estos estudios muestran los beneficios potenciales de los sistemas de apoyo a la toma de decisión clínica a la práctica médica, sin embargo, estudios controlados en un entorno real debe llevarse a cabo para confirmar este punto de vista

¹ Mestranda do Programa de Pós Graduação em Informática em Saúde - Departamento de Informática em Saúde da Universidade Federal de São Paulo, Brasil.

² Chefe da Disciplina de Gastroenterologia Pediátrica, Universidade Federal de São Paulo, Brasil.

³ Departamento de Informática em Saúde, Universidade Federal de São Paulo, Brasil.

⁴ Titular do Departamento de Enfermagem, Universidade Federal de São Paulo, Brasil.

INTRODUÇÃO

Um sistema de apoio à decisão clínica (SADC) é qualquer programa projetado para ajudar profissionais de saúde na tomada de decisão⁽¹⁾. Tais sistemas têm sido alvo de estudos durante as últimas décadas na área de informática em saúde. O aumento progressivo da quantidade de dados, informações e conhecimento necessários para a prática médica é o principal motivo para o desenvolvimento de sistemas de apoio à decisão clínica⁽²⁾. Segundo Denekamp⁽³⁾, a expectativa é que os SADCs diminuam a distância entre as evidências e a prática clínica, ao disponibilizar dados relevantes e conhecimento no ponto de cuidado.

Historicamente, um dos primeiros SADCs de sucesso foi desenvolvido para gastroenterologia no final da década de 60, especificamente, para o processo de diagnóstico e decisão da dor abdominal aguda não traumática: o Leeds Abdominal Pain System⁽⁴⁾. Dados clínicos e laboratoriais da primeira consulta foram analisados por meio da Teoria da Probabilidade de Bayes, sendo considerados dois pressupostos: a independência condicional e a exclusividade mútua de sete diagnósticos propostos, a saber, apendicite, diverticulite, úlcera perfurada, coleciste, obstrução do intestino delgado, pancreatite e dor abdominal inespecífica. A taxa de acerto obtida pelo SADC atingiu 91,8% dos 304 casos analisados, um valor muito superior à taxa de diagnósticos corretos apontados pelos médicos, que variou entre 65% e 80%. Duas questões fundamentais sobre SADCs foram levantadas no artigo original: os médicos realmente precisam de SADCs? Os SADCs podem oferecer qualquer vantagem mensurável em termos de precisão diagnóstica? A resposta só pode ser conferida em um ambiente clínico real, em um estudo controlado que possibilite verificar as métricas sem e com o SADC⁽⁴⁾.

Em busca de evidência para estes questionamentos, este artigo objetiva mostrar o estado da arte dos SADCs desenvolvidos no período entre 2005 e 2010, de forma a determinar o nível de desenvolvimento, testes e vantagens conferidas por estes estudos à prática médica.

MÉTODOS

Para fins de verificação do escopo dos SADCs desenvolvidos em gastroenterologia, no período entre 20 e 21 de janeiro de 2011 foram realizadas buscas nas bases de dados PubMed[®] (<http://www.ncbi.nlm.nih.gov/pubmed>), Literatura Latinoamericana y del Caribe en Ciencias de la Salud (LILACS) (<http://bases.bireme.br>) e ISI Web of Knowledge (<http://www.isiknowledge.com>) para fins de verificação da ocorrência de publicações que relatassem experiências nesse assunto. Foi selecionado um conjunto de descritores pesquisados no Medical Subject Headings (MeSH) (<http://www.ncbi.nlm.nih.gov/mesh>). De forma a ampliar o escopo da busca, alguns termos não encontrados no MeSH foram utilizados. Os descritores e termos selecionados foram agrupados da forma seguinte:

- Sistemas de apoio à decisão: *clinical decision support systems* [MeSH], *decision support systems, clinical* [MeSH], *decision support*,
- Termos médicos referentes a gastroenterologia: *gastroenterology* [MeSH], *abdominal pain* [MeSH], *reflux*, *gastritis* [MeSH], *bowel or intestines* [MeSH], *stomach* [MeSH],

gastrointestinal, liver [MeSH].

Como estratégia de pesquisa, uma expressão foi composta, utilizando o operador OR para termos de mesmo grupo e AND para a junção dos grupos.

Para a base de dados PubMed[®] a expressão utilizada foi ("*gastroenterology*" OR "*abdominal pain*" OR "*reflux*" OR "*gastritis*" OR "*bowel*" OR "*intestines*" OR "*stomach*" OR "*gastrointestinal*" OR "*liver*") AND ("*clinical decision support systems*" OR "*decision support systems, clinical*").

Para a base de dados LILACS, foi utilizada a expressão "*clinical decision support systems*" OR "*decision support*".

Para a busca de SADCs na base ISI, a expressão composta foi *Topic=((gastroenterology* OR "*abdominal pain*" OR *reflux* OR *gastritis* OR *bowel* OR *intestines* OR *stomach* OR *gastrointestinal* OR *liver*) AND ("*clinical decision support systems*" OR "*decision support systems, clinical*" OR "*decision support*").

Um questionário para avaliação das publicações foi esquematizado, com o objetivo de facilitar a caracterização e comparação dos estudos. O Quadro 1 mostra as questões propostas.

Todos os artigos foram lidos e analisados. Na fase seguinte, o questionário para avaliação foi preenchido para cada publicação selecionada.

Quadro 1 - Questionário para avaliação das publicações

#	Questão
1	Qual o problema clínico em questão?
2	O SADC é aplicado à investigação de que doença?
3	Que ferramentas de inteligência artificial são integradas?
4	Qual o tamanho das amostras utilizadas?
5	Quais os valores das métricas avaliadas?
6	Há comparação entre as métricas do SADC e os especialistas?
7	Descreve experiências com pacientes ou trata-se de um protótipo?
8	Descreve uma avaliação dos usuários?
9	Há evidências de melhoria em algum aspecto da prática médica a partir da introdução do SADC?
10	São relatadas críticas ou problemas em relação ao SADC?

RESULTADOS

Foram incluídos artigos originais, com o resumo disponível, publicados entre 2005 e 2010. Foram excluídas todas as publicações cujos títulos não indicassem a descrição de SADC em gastroenterologia. Após leitura dos resumos, foram excluídos os artigos que não apresentaram uma descrição de um SADC computadorizado aplicado a doenças do aparelho digestivo.

A base de dados LILACS não retornou nenhuma ocorrência. A Tabela 1 mostra a quantidade de publicações obtidas a partir das buscas realizadas nas bases de dados e após a aplicação dos critérios de inclusão.

Tabela 1 – Quantidade de publicações selecionadas em cada fase

Base de dados	Busca	Resumo	Texto completo
PubMed	33	25	5
ISI	71	19	7
Total	104	44	12

Tabela 2 - Dados coletados das publicações selecionadas

Questão	1	2	3	4	5	6	7	8	9	10
Autor	Problema clínico	Doença investigada	Técnicas de IA	Amostra	Métrica (taxa de acerto)	Comparação com médicos?	Tipo de modelo (experimental/produção)	Avaliação dos usuários?	Descrição de Melhorias?	Críticas ou problemas?
Lin, 2009 ⁹	Diagnóstico	Hepática	CART CBR	510 casos clínicos	CART: 92,94% CBR: 91,09%	Não	Experimental	Não	Não	Não
Farion et al., 2009 ⁶	Diagnóstico	Dor abdominal	-	-	-	Não	Experimental	Não	Não	Não
Das et al., 2008 ⁷	Conduta clínica	Hemorragia gastrointestinal aguda	RNA (vários modelos)	194 - Tr 193 - Te 200 - VE	81% - Te 61% - VE	Não	Experimental	Não	Não	Não
Lahner et al., 2008 ⁸	Diagnóstico	Distúrbios da tireóide em pacientes com gastrite atrófica	RNA	253 casos clínicos	75,8%	Não	Experimental	Não	Não	Não
Chu et al., 2008 ⁹	Conduta clínica	Hemorragia gastrointestinal aguda	SVM, RNA, KNN, LDA, SC, RF, RL, Boosting	189 casos clínicos	80%	Não	Experimental	Não	Não	Não
Aruna et al., 2007 ⁽¹⁰⁾	Diagnóstico	Distúrbios gastrointestinais	RNA (BP e RBF)	1125 casos clínicos	-	Não	Experimental	Não	Não	Não
Yang et al., 2007 ⁽¹¹⁾	Eficácia do medicamento	Hepate C	SVM e KNN	112 casos clínicos	Eficaz: 85% Ineficaz: 83%	Não	Experimental	Não	Não	Não
Bemer et al., 2006 ⁽¹²⁾	Segurança na prescrição de medicamento	Hemorragia gastrointestinal	-	68 médicos	-	Sim	Experimental	Não	Sim	Não
Sadeghi et al., 2006 ⁽¹³⁾	Conduta clínica	Dor abdominal não traumática	Rede bayesiana	90 casos clínicos	56%	Sim	Produção	Não	Não	Sim

KNN: *k-nearest neighbors* (k-vizinhos próximos); CART: *classification and regression tree*; CBR: *case-based reasoning*; RNA: *redes neurais artificiais*; SVM: *support vector machines*; LDA: *linear discriminant analysis*; SC: *shrunken centroid*; RF: *random forest*; RL: *regressão logística*; BP: *back-propagation*; RBF: *radio basis function*; Tr: *treinamento*; Te: *teste*; VE: *validação externa*

Três publicações foram comuns nas bases PubMed e ISI. Desta forma, nove publicações foram selecionadas para leitura do texto completo.

Na Tabela 2 foram relacionadas características essenciais dos SADCs descritos nos artigos, de acordo com o questionário para análise das publicações.

A maioria dos estudos selecionados descreveu modelos experimentais que cujos resultados das métricas mostraram potencial possibilidade de sucesso nas tarefas de diagnóstico clínico, conduta e eficácia de medicamento. A questão do diagnóstico clínico é o tema mais freqüente dos estudos, descrito em quatro publicações (44,4%), seguida por conduta clínica, descrita em três estudos (33,3%).

O distúrbio clínico mais freqüente foi hemorragia gastrointestinal, presente em três estudos^(7,9,12) (33,3%), com abordagens diferenciadas. No estudo de Das et al.⁽⁷⁾ na fase inicial um modelo experimental foi desenvolvido e validado para prever a necessidade de tratamento endoscópico. No estudo de Chu et al.⁽⁹⁾ um modelo experimental foi desenvolvido para a previsão da origem da hemorragia, necessidade de transfusão de sangue ou derivados, endoscopia urgente ou predisposição à hemorragia gastrointestinal aguda, de forma a apoiar a conduta clínica em uma situação de emergência. No estudo de Berner et al.⁽¹²⁾ um experimento controlado com médicos residentes mostrou que o SADC aumentou a taxa de prescrição segura de medicamento de 0,27 para 0,45, o que sugere que os médicos prescreveram tratamentos mais seguros quando assistidos por um SADC.

Em dois estudos^(6,13) os SADCs descritos foram projetados para utilização para auxílio de casos no setor de emergência. Farion et al.⁽⁶⁾ descreveram um SADC desenvolvido para triagem de pacientes por meio da análise da história clínica, exame físico e testes de laboratório, utilizando computadores de mão. Sadeghi et al.⁽¹³⁾ descreveram um SADC, baseado em rede bayesiana, cujo objetivo é realizar uma triagem automática de pacientes que apresentam dor abdominal não traumática, indicando a conduta a ser realizada. O sistema apresentou um alto valor de sensibilidade (90% versus 64%) e baixa especificidade (25% versus 48%) quando comparado às decisões tomadas por um especialista de emergência.

A aplicação de técnicas de inteligência artificial para o desenvolvimento de SADCs ocorreu em sete (78%) publicações, sendo que as redes neurais artificiais (RNA) foram aplicadas em quatro (57%) dessas. A concorrência entre as técnicas para obtenção de melhores métricas foi aplicada em quatro (57%) dos modelos experimentais.

A utilização de técnicas complementares foi descrita por Lin⁽⁵⁾. Na fase I do estudo foi aplicada a técnica de árvores de decisão *classification and regression tree* (CART), para descobrir e regras, de forma a classificar pacientes saudáveis ou com doenças hepáticas. Na fase II, um modelo baseado em *case-based reasoning* (CBR), foi desenvolvido para identificar o tipo de doença presente, de acordo com o CID-10, dentro do grupo com doença hepática positiva.

Técnicas de seleção de variáveis foram aplicadas em dois modelos experimentais^(8,11) e foram eficientes para a melhoria das métricas.

O SADC descrito por Farion et al.⁽⁶⁾ utilizou ontologia como mecanismo de representação do conhecimento, uma forma que tem se mostrado efetiva na solução de problemas de apoio à decisão.

DISCUSSÃO

Apesar da restrita quantidade de publicações disponíveis identificadas como SADCs em gastroenterologia, este estudo mostrou-se válido para verificar o estado de desenvolvimento, testes e vantagens conferidas por estes estudos à prática médica.

A análise das publicações recentes mostrou que modelos experimentais em fase inicial de desenvolvimento são predominantes. Desta forma, não foi possível conferir as evidências de vantagens à prática médica dos SADCs.

Os SADCs em gastroenterologia avaliados foram aplicados a uma multiplicidade de problemas clínicos e investigação de doenças, o que mostra que a amplitude do escopo dos SADC no auxílio à prática médica, está muito além de apenas propiciar auxílio ao diagnóstico clínico.

O desenvolvimento de SADCs no ambiente de emergência sugere uma confiabilidade em relação à informação indicada pelo sistema em uma situação em que o tempo para a tomada de decisão é um fator crítico.

Técnicas de inteligência artificial foram aplicadas para análise de dados clínicos e previsão de classes em 78% dos estudos. As métricas mostraram excelentes resultados na identificação de diagnósticos, conduta clínica e eficiência na utilização de um medicamento.

Em 89% das publicações foi descrito o desenvolvimento de modelos experimentais baseados na aplicação de técnicas de inteligência artificial para suportar um SADC, apesar de serem identificados como SADC. Apenas Sadeghi et al.⁽¹³⁾ relataram um sistema em produção em que os dados clínicos dos pacientes eram inseridos por enfermeiras.

A realização de validação externa, em que o modelo experimental é testado com dados de fontes externas, é um item importante na confiabilidade das métricas apuradas. Porém, apenas o estudo descrito por Das et al.⁽⁷⁾ apresentou resultados referentes à validação externa do modelo experimental.

Um aspecto fundamental dos SADCs é a busca de evidência de melhorias da prática médica a partir da introdução do SADC, ou a verificação das suas influências, como reportado por Berner et al.⁽¹²⁾. Neste estudo, o grupo com intervenção do SADC mostrou uma sensível melhora em termos de segurança da prescrição de medicamentos. Nos outros estudos, encontram-se apenas expectativas, ao analisarem-se as potencialidades trazidas pelas métricas.

Um exemplo promissor de melhorias na prática médica é encontrado no estudo de Onisko⁽¹⁴⁾, que descreve um sistema de apoio à decisão para doenças hepáticas (HEPAR II) baseado em redes bayesianas. A taxa de acerto diagnóstico do HEPAR II foi de 70%, contra 33% dos especialistas. A interação dos médicos com o sistema resultou no dobro da taxa de acerto

diagnóstica, em relação à situação anterior, respectivamente, 65,8% contra 33,6%.

Infelizmente, não foram descritas avaliações qualitativas dos usuários (médicos ou enfermeiras) acerca do funcionamento e impacto do SADC na prática clínica, que são importantes itens para a adequação dos SADCs às necessidades dos usuários e melhoria do atendimento ao paciente.

Vale destacar que não foram encontrados na base PubMed, estudos que descrevem experiências de SADCs em gastroenterologia no Brasil, em nenhum período.

CONCLUSÃO

As experiências que descrevem a utilização dos SADCs em gastroenterologia mostraram que estes são potencialmente úteis em aplicações diversas e à prática

médica, segundo expectativas trazidas pelos modelos experimentais analisados, apesar do reduzido número de publicações sobre o assunto no período analisado.

A utilização de técnicas de inteligência artificial é a fase inicial para o desenvolvimento dos SADCs, desde que sejam realizados testes para validação interna e externa dos modelos experimentais. Os benefícios da utilização dessas técnicas em medicina foram sintetizados por Ramesh⁽¹⁵⁾ e Lisboa⁽¹⁶⁾.

A análise das publicações recuperadas neste estudo ainda não foi suficiente para responder as questões levantadas por deDombal⁽⁴⁾. Não há ainda evidência de que os médicos realmente precisem de SADCs. Apenas um estudo mostrou um indício de que há uma vantagem em termos de precisão, mas não diagnóstica. Desta forma, a necessidade de estudos controlados em um ambiente clínico real continua presente.

REFERÊNCIAS

- Musen MA, Shahar Y, Shortliffe EH. Clinical decision-support systems. In: Shortliffe EH, Cimino JJ. *Biomedical informatics: computer applications in health care and biomedicine*. 3a ed. New York: Springer; 2006. p. 698-736.
- Sigulem D, Anção MS, Ramos MP, Leão BF. *Sistemas de apoio à decisão em medicina. Atualização Terapêutica – Manual prático de diagnóstico e tratamento* [Internet]. 1998 [citado 2007 Set 10]. Disponível em http://www.virtual.epm.br/material/tis/curr-med/sad_html/sistema.htm.
- Denekamp Y. Clinical decision support systems for addressing information needs of physicians. *Isr Med Assoc J*. 2007;9(11):771-6.
- de Dombal FT, Leaper DJ, Staniland JR, McCann AP, Horrocks JC. Computer-aided diagnosis of acute abdominal pain. *Br Med J*. 1972; 2(5804):9-13.
- Lin RH. An intelligent model for liver disease diagnosis. *Artif Intell Med*. 2009;47(1):53-62.
- Farion K, Michalowski W, Wilk S, O'Sullivan D, Rubin S, Weiss D. Clinical decision support system for point of care use—ontology-driven design and software implementation. *Methods Inf Med*. 2009;48(4):381-90.
- Das A, Ben-Menachem T, Farooq FT, Cooper GS, Chak A, Sivak MV Jr, Wong RC. Artificial neural network as a predictive instrument in patients with acute nonvariceal upper gastrointestinal hemorrhage. *Gastroenterology*. 2008 Jan;134(1):65-74.
- Lahner E, Intraligi M, Buscema M, Centanni M, Vannella L, Grossi E, Annibale B. Artificial neural networks in the recognition of the presence of thyroid disease in patients with atrophic body gastritis. *World J Gastroenterol*. 2008;14(4):563-8.
- Chu A, Ahn H, Halwan B, Kalmin B, Artifon EL, Barkun A, Lagoudakis MG, Kumar A. A decision support system to facilitate management of patients with acute gastrointestinal bleeding. *Artif Intell Med*. 2008;42(3):247-59.
- Aruna P, Puviarasan N, Palaniappan, B. Diagnosis of gastrointestinal disorders using DIAGNET. *Expert Syst Appl*. 2007; 32(2):329-35.
- Yang J, Nugroho AS, Yamauchi K, Yoshioka K, Zheng J, Wang K, et al. Efficacy of interferon treatment for chronic hepatitis C predicted by feature subset selection and support vector machine. *J Med Syst*. 2007;31(2):117-23.
- Berner ES, Houston TK, Ray MN, Allison JJ, Heudebert GR, Chatham WW, et al. Improving ambulatory prescribing safety with a handheld decision support system: a randomized controlled trial. *J Am Med Inform Assoc*. 2006;13(2):171-9.
- Sadeghi S, Barzi A, Sadeghi N, King B. A Bayesian model for triage decision support. *Int J Med Inform*. 2006;75(5):403-11.
- Onisko, A. Probabilistic causal models in medicine: application to diagnosis of liver disorders [tese]. Pittsburgh: University of Pittsburgh; 2002.
- Ramesh AN, Kambhampati C, Monson JRT, Drew PJ. Artificial intelligence in medicine. *Ann R Coll Surg Engl*. 2004; 86(5):334–8.
- Lisboa PJ. A review of evidence of health benefit from artificial neural networks in medical intervention. *Neural Netw*. 2002;15(1):11-39.