

## Avaliação da acurácia da informação em sites de saúde: métodos para construção de indicadores baseados em evidência

**Rodolfo Paolucci**

Doutor; Fundação de Apoio à Escola Técnica, Rio de Janeiro, RJ, Brasil;  
rpprjbr@gmail.com; ORCID: <https://orcid.org/0000-0003-3986-1118>

**André de Faria Pereira Neto**

Doutor; Escola Nacional de Saúde Pública Sergio Arouca, Rio de Janeiro, RJ, Brasil;  
andrepereiraneito@gmail.com; ORCID: <https://orcid.org/0000-0003-3631-8857>

**Paulo Nadanovsky**

Doutor; Universidade do Estado do Rio de Janeiro, Rio de Janeiro, RJ, Brasil;  
nadanovsky@ims.uerj.br; ORCID: <https://orcid.org/0000-0003-3345-9873>

**Resumo:** O problema da qualidade da informação de saúde na Internet tem sido apontado nas últimas duas décadas. Para lidar com esse problema, a avaliação de sites realizada por profissionais, pesquisadores e instituições tem se configurado como um novo campo de pesquisa e ação. Uma das principais questões nesse campo está relacionada aos critérios de avaliação. Acurácia é o critério mais utilizado, mas identificamos um problema em sua definição e aplicação: admitir consenso entre especialistas como parâmetro de qualidade. Entretanto, o consenso não garante um resultado atual e correto. Neste artigo, apresentamos e discutimos métodos orientados pela medicina baseada em evidência para construção de indicadores de acurácia. O conjunto de métodos desenvolvido compreende sete etapas: estratégia de busca, seleção da fonte de informação baseada em evidência, coleta e seleção de tópicos, desenvolvimento da primeira versão de indicadores, análises dos grupos, dos tópicos e dos indicadores. Nossos métodos garantem o desenvolvimento de indicadores de acordo com as melhores e atuais evidências científicas. Eles resolveram o problema do critério acurácia para o campo de avaliação da qualidade da informação em sites de saúde e viabilizaram a proposição de uma nova definição. Os resultados desta pesquisa preenchem lacunas no conhecimento do campo e orientam pesquisas e ações futuras, pois são replicáveis para qualquer assunto de saúde.

**Palavras-chave:** Medicina baseada em evidência; Internet e acesso à informação; Critérios; Indicadores de acurácia; Informação de Saúde; Sites de Saúde

## 1 Introdução

O acesso à informação de saúde de qualidade na Internet pode ter consequências positivas tanto para os gestores de sistemas de saúde como para os cidadãos. Para os gestores de saúde, o acesso à informação de qualidade reduz os custos dos sistemas de saúde, pois diminui a busca e o uso dos serviços oferecidos (PARK et al., 2015; SPOELMAN et al., 2016). Para os cidadãos, por exemplo, há quase duas décadas atrás, os benefícios obtidos pela maioria de pacientes com câncer que acessavam informação online foram relatados (EYSENBACH, 2003). O autor aborda o impacto da informação obtida em sites de saúde. Quando essa informação é de qualidade, os pacientes podem obter conhecimento, ficar menos confusos, ter expectativas realísticas e, assim, compartilhar das decisões, aderir aos tratamentos e autocuidar de forma eficaz (EYSENBACH, 2003). Informação de qualidade propicia o desenvolvimento de habilidades que conferem ao cidadão maior poder de decisão sobre sua saúde e seu autocuidado (ALBARRAK et al., 2016; LAGAN; SINCLAIR; KERNOHAN, 2011).

Contudo, a qualidade da informação de saúde é um problema a ser investigado na Internet. Na primeira revisão sistemática sobre o tema, a maioria dos estudos analisados (70%) concluiu que a qualidade da informação é um problema (EYSENBACH et al., 2002). Resultado semelhante foi encontrado em outra revisão publicada 13 anos depois. Zhang, Sun e Xie (2015) identificaram que 55,2% dos estudos avaliaram a informação de saúde na Internet como negativa e 37% como de qualidade variada. Segundo Lagan, Sinclair e Kernohan (2011), quando informação errada é acessada, pode provocar sentimentos de apavoramento, ansiedade e paranoia. Para Powell et al. (2011), informação incorreta, incompreensível ou desatualizada pode desencadear tomadas de decisão prejudiciais à saúde. Portanto, o acesso à informação de baixa qualidade pode ter consequências negativas para a saúde dos cidadãos. Então, como lidar com esse problema?

A principal proposta para lidar com o problema da qualidade da informação de saúde na Internet compreende as iniciativas institucionais públicas ou privadas de avaliação. Elas existem há mais de vinte anos e as

primeiras experiências são predominantemente europeias e norte americanas (MORENO et al., 2010). Elas podem realizar avaliações rigorosas envolvendo profissionais e usuários no julgamento da qualidade da informação de saúde na Internet. Algumas dessas iniciativas conferem um selo de qualidade aos sites que atenderem aos seus critérios de qualidade (PEREIRA NETO; PAOLUCCI, 2019). O selo é uma marca que comunica confiabilidade aos usuários dos sites e representa uma informação chancelada institucionalmente. Ele é uma solução e um serviço oferecido à maioria das pessoas que não tem condições de julgar a qualidade da informação de saúde na Internet, problema evidenciado por Silva-Jerez e Furnival (2018).

Uma das principais questões discutidas nesse campo está relacionada com os critérios utilizados na avaliação de sites de saúde (PAOLUCCI; PEREIRA NETO, 2021). Os critérios de avaliação predominantes na literatura especializada estão presentes em três revisões sistemáticas.

Eysenbach et al. (2002) realizaram a primeira revisão sistemática. Este trabalho reuniu os métodos e os processos adotados na avaliações da qualidade da informação de saúde na Internet. Esta revisão analisou 79 estudos publicados até o ano de 2001. Os autores identificaram 86 critérios de qualidade distintos. Dentre eles, muitos foram empregados segundo as mesmas regras, mas foram relatados com diferentes nomes. Para normalizar tais nomenclaturas e apresentar um quadro metodológico, eles sintetizaram essas 86 denominações em cinco critérios de qualidade (EYSENBACH et al., 2002). Esta é uma das principais contribuições desse estudo pioneiro nesse campo. Os cinco critérios são técnico, *design*, abrangência, legibilidade e acurácia.

A segunda revisão sistemática foi publicada treze anos depois (ZHANG; SUN; XIE, 2015). A amostra dessa revisão incluiu estudos publicados exclusivamente em inglês entre os anos de 2002 e 2013. Ela foi composta por 165 estudos distintos. Os autores identificaram 11 critérios de avaliação que foram agrupados nas categorias “*content*” e “*design*”. A categoria “*content*” possui critérios contemplados por aqueles sintetizados na primeira revisão sistemática, pois tratam das mesmas regras dos critérios técnico, acurácia,

abrangência e legibilidade. O mesmo pode ser afirmado em relação ao critério *design* encontrado nas duas revisões.

A terceira revisão sistemática também foi finalizada em 2015 (PAOLUCCI; PEREIRA NETO, 2021). Um dos objetivos desse estudo foi atualizar os dados encontrados na primeira revisão sobre o tema (EYSENBACH et al., 2002). Para selecionar sua amostra, o estudo empregou uma estratégia de busca semelhante à da primeira revisão de 2002. O período coberto foi do ano de 2001 até meados de 2014. Diferente da segunda revisão publicada no mesmo ano (ZHANG; SUN; XIE, 2015), não houve restrição em relação aos idiomas dos estudos que compuseram a amostra analisada por Paolucci e Pereira Neto (2021). A amostra contemplou 279 estudos. Paolucci e Pereira Neto (2021) verificou que os critérios sintetizados na primeira revisão sistemática continuam válidos. Ele constatou que os cinco critérios foram utilizados pela maioria dos estudos na seguinte proporção: acurácia (64%), legibilidade (38%), abrangência (27%), técnico (24%) e *design* (17%) (PAOLUCCI; PEREIRA NETO, 2021). De acordo com esse resultado, o critério acurácia continua sendo o critério mais utilizado por estudos que se propõem a avaliar a qualidade da informação de saúde na internet.

Contudo, identificamos um problema na definição das regras que constituem o critério acurácia. Na primeira revisão sistemática, acurácia foi definida como o “[...] grau de concordância da informação oferecida com a melhor evidência ou com a prática médica geralmente aceita” (EYSENBACH et al., 2002, p. 2695). O grau de concordância da informação com a prática médica é realizado, geralmente, através do consenso de especialistas. Este processo não garante que o resultado desse consenso esteja atual e correto. Pelo contrário, os profissionais envolvidos podem estar desatualizados em relação às melhores evidências científicas. Eles podem, ainda, apresentar opiniões e condutas que estejam em desuso.

A preocupação com práticas amparadas no consenso de especialistas não é nova. No fim da década de 1990, houve uma crescente oferta de cursos de formação continuada para médicos na tentativa de mantê-los atualizados com a melhor evidência científica (NADANOVSKY, 1999). Essa oferta é justificável,

porque geralmente o profissional recém-formado pode ter tido acesso ao conhecimento através de livros que frequentemente estão desatualizados. No entanto, segundo Nadanovsky (1999), a formação continuada não resolveu o problema da atualização dos médicos. Ele afirma que o consenso de especialistas é baseado no conhecimento dos profissionais envolvidos, indivíduos inevitavelmente influenciados por fatores pessoais (crenças, posições políticas, sua saúde física e mental, estado emocional, etc.) e por condições diversas de trabalho (infraestrutura, renda, motivação, etc.).

O consenso de especialistas é defendido em determinadas situações. Betting et al. (2003, p. 1068), apesar de afirmarem que as “decisões terapêuticas devem ser baseadas em evidências apresentadas pela literatura”, defendem que frente “[...] à persistência de questões clínicas não respondidas ou que permanecem controversas [...] a opinião de especialistas torna-se de grande valor” (BETTING et al., 2003, p. 1046). Contudo, as evidências, além de reafirmarem a “inabilidade do médico em se manter em dia com os importantes avanços”, demonstra que há “[...] variações extraordinárias no comportamento clínico do médico e nas taxas de intervenções benéficas, inúteis e prejudiciais” (NADANOVSKY, 1999, p. 22). Por estas razões, o fomento da medicina baseada em evidência (MBE) tem sido intensificado. Para Atallah (2018, p. 43), “apoiar as práticas em medicina com base nas melhores evidências oriundas de rigorosas metodologias que lhes dão validade comparativa é um processo civilizatório”, pois, ao observarmos a história da humanidade, tais práticas permitiram descobrir curas para doenças letais e salvaram milhares de vidas.

Diante desse cenário, manter a prática médica geralmente aceita na definição do critério acurácia é um problema para o campo de avaliação da qualidade da informação de saúde na Internet. Os indicadores de acurácia tem sido construídos através do consenso de especialistas, manuais técnico científicos, diretrizes médicas, livros didáticos ou literatura.

Neste sentido, o objetivo deste artigo é apresentar e discutir os métodos orientados pela MBE que desenvolvemos para a construção de indicadores de acurácia. Como objetivo secundário, este artigo pretende atualizar a definição do

critério acurácia para o campo de avaliação da qualidade de informação de saúde na Internet.

## 2 Referencial metodológico

David Sackett é considerado o pai da MBE, pois, no final da década de 1960 na Universidade de McMaster, ele e colegas de departamento passaram a ensinar disciplinas clínicas com uma nova proposta (SMITH; RENNIE, 2014). A proposta nomeada de *critical appraisal* incorporava epidemiologia e estatística e utilizava uma abordagem voltada para os problemas dos pacientes. Em 1996, Sackett et al. (1996) relatam que a MBE evoluiu desde sua fundação e permanecia como um tópico de discussão importante para uma ampla audiência composta por profissionais de saúde, gestores, investidores e público em geral. Eles esclarecem o quê a MBE é e o quê não é. Eles apresentam a seguinte definição:

Medicina baseada em evidência é o uso consciente, explícito e criterioso da melhor evidência atual na tomada de decisão sobre o cuidado de pacientes individuais. A prática da medicina baseada em evidência significa integrar o conhecimento clínico individual com a melhor evidência clínica externa disponível proveniente de pesquisa sistemática. (SACKETT et al., 1996, p. 71, tradução nossa)

Eles explicitam o que estão chamando de conhecimento clínico individual e de melhor evidência clínica externa disponível. O primeiro é definido por eles como o conhecimento que os clínicos adquirem com a experiência da prática clínica. Eles afirmam que esse conhecimento resulta em diagnósticos mais assertivos e em uma postura mais sensível aos direitos e às escolhas dos pacientes. O segundo é definido por esses autores como a evidência oriunda de pesquisa clínica considerada relevante como as pesquisas das ciências básicas e pesquisas centradas no paciente sobre testes diagnóstico, prognósticos e intervenções de tratamento e de prevenção. Por exemplo, no caso da tuberculose, existe evidência proveniente de pesquisa clínica realizada no ano de 2005 afirmando que tosse por duas ou mais semanas é indicativo da doença e essa medida deve ser utilizada para realização de um diagnóstico precoce (DYNAMED, 2019a). No Brasil, o Ministério da Saúde ainda informa que a

tosse por três ou mais semanas é sintoma da tuberculose (BRASIL, 2021). Para Sackett et al. (1996, p. 72, tradução nossa) a melhor evidência clínica externa disponível “invalida testes diagnóstico e tratamentos previamente aceitos e os substitui por novos mais poderosos, mais precisos, mais eficazes e mais seguros”.

É importante ressaltar que o binômio experiência clínica e melhor evidência externa são ambos essenciais na MBE. O trecho a seguir resume a interdependência desse binômio para a prática da MBE:

Sem *expertise* clínica, a prática corre o risco de ser tiranizada pelas evidências, pois mesmo evidências externas excelentes podem ser inaplicáveis ou inadequadas para um paciente em particular. Sem as melhores evidências atuais, a prática corre o risco de ficar rapidamente desatualizada, em detrimento dos pacientes. (SACKETT et al., 1996, p. 72, tradução nossa).

Sackett et al. (1996) ressaltam o caráter integrativo contido na definição do que é a MBE, mas o papel do paciente na tomada de decisão parece restrito à manifestação de preferências e valores. A MBE tem foco no profissional de saúde e parece que o mantém como único responsável pelo conhecimento e por fornecer informação aos pacientes. Em nossa visão de mundo, consideramos que os pacientes podem exercer um papel mais ativo sobre sua saúde. Eles podem integrar a melhor e mais atual evidência científica com sua experiência pessoal, ou seja, com sua vivência na questão de saúde que o afeta. Desta forma, os pacientes têm maiores condições de discutir as tomadas de decisão com os profissionais de saúde.

## 2.1 Evidência pré avaliada

Segundo Dicenso, Bayley e Haynes (2009), no passado, para obtermos a melhor evidência, precisávamos aprender técnicas para efetuar buscas na literatura especializada e desenvolver habilidades de avaliação crítica para identificar e compreender os estudos clínicos de qualidade. Contudo, os autores explicam que foram criados muitos recursos para facilitar a obtenção das evidências. Eles caracterizam tais recursos como *preappraised* por fornecerem *preappraised evidence* que traduzimos como ‘evidência pré avaliada’. Esses recursos passam

por “[...] um processo de filtragem para incluir apenas os estudos de maior qualidade e são atualizados regularmente para que as evidências que acessamos por meio desses recursos sejam atuais” (DICENSO; BAYLEY; HAYNES, 2009, p. 99, tradução nossa).

Para Burwell (2019), os processos de filtragem dos recursos de ‘evidência pré avaliada’ são explícitos e os parâmetros para julgar a qualidade dos estudos são baseados no mérito científico e na facilidade com que as evidências podem ser aplicadas nas decisões clínicas. Caracterizar tais evidências como pré avaliadas ressalta seu potencial para a prática da MBE. O autor aponta três motivos que justificam seu potencial. O primeiro diz respeito ao fato dos profissionais nem sempre saberem realizar avaliação crítica de estudos. Caso o profissional conheça critérios para julgar a qualidade dos estudos, o segundo motivo é não ter esses critérios disponíveis quando for preciso. O terceiro motivo é aplicar tais critérios erroneamente. Nesses casos, as ‘evidências pré avaliadas’ são a melhor opção para os profissionais de saúde, devendo realizar buscas mais extensas na literatura especializada somente se não encontrar o que precisa previamente avaliado (BURWELL, 2019).

Há dez anos atrás, já existiam recursos que ofereciam ‘evidência pré avaliada’ para diversos problemas clínicos (DICENSO; BAYLEY; HAYNES, 2009). Alguns deles disponibilizam sumários das evidências. Para Chapa et al. (2013), sumários de recursos de ‘evidência pré avaliada’ fazem com que os profissionais de saúde acessem rapidamente informação relevante sobre condições e intervenções destinadas tanto para pacientes específicos como para perfis populacionais. Burwell (2019) apresenta um guia que auxilia observarmos se um recurso é realmente baseado em evidência: *Guides for judging whether an online clinical text is evidence based and current*. O guia é composto por 14 itens que podemos observar para avaliar textos clínicos online. A proposta do autor consiste na aplicação desses itens na avaliação de textos clínicos online, indicando se cada um está ausente ou presente. Para ele, um texto que tenha menos de cinco itens presentes não pode ser considerado como um texto baseado em evidência (BURWELL, 2019).

Burwell (2019) apresenta o recurso de *meta-search* que auxilia o esforço despendido para encontrar sites que oferecem ‘evidência pré avaliada’ em formato de textos clínicos online. A *meta-search* é um tipo de busca amplamente utilizada em serviços na Internet para reunir em uma única plataforma os resultados de buscas em diferentes sites. Esse tipo de busca se tornou uma forma de lidar com o grande e crescente volume de informação disperso que está disponível na Internet, inclusive informação científica e tecnológica.

Existem inúmeros sites dedicados à produção e à disponibilização de ‘evidência pré avaliada’. Esse volume tende a crescer devido ao fomento da MBE e de sua expansão para áreas do conhecimento além da medicina. Com esse crescimento, podem ser encontrados sites com resumos críticos de estudos clínicos randomizados, de revisões sistemáticas, de *guidelines* ou, ainda, sites que consolidam sumários das ‘evidências pré avaliadas’ de todos esses tipos de materiais na forma de textos clínicos online.

## **2.2 Pirâmide 5.0 de atenção à saúde baseada em evidência**

Para ensinar a prática da MBE, modelos têm sido desenvolvidos e aperfeiçoados. Esses modelos orientam os profissionais de saúde na busca e na seleção de informação baseada em evidência que possa ser integrada com suas experiências e *expertises* no contexto da prática clínica. Alper e Haynes (2016) apresentam um resumo desses modelos.

O primeiro modelo foi proposto por Haynes (2001). Ele propôs um modelo de estrutura hierárquica na forma de pirâmide chamado de “4S”, primeira letra dos nomes de cada nível. Da base ao topo da pirâmide, esse modelo possui quatro níveis de organização das evidências encontradas nas pesquisas: *Studies*, *Syntheses*, *Synopses* e *Systems*. Cinco anos após essa publicação, o mesmo autor publicou um editorial atualizando o modelo (HAYNES, 2006). Nessa atualização, ele adicionou o nível “*Summaries*” na pirâmide entre sistemas e sinopses. De acordo com o autor, a proposta desse nível sumários é integrar as evidências provenientes das três camadas inferiores.

Ele explica que os três níveis inferiores representam evidências sobre apenas um aspecto relacionado a um problema ou condição de saúde como um medicamento. Além disso, sua configuração exige que os profissionais de saúde integrem os achados e, no caso do primeiro nível estudos, façam avaliações críticas por conta própria. O novo nível sumários permite uma visão abrangente de um problema ou condição de saúde.

Nos anos seguintes, os serviços de informação baseada em evidência continuaram a evoluir, além do aumento do volume de publicações tanto de estudos originais, revisões sistemáticas e até revisões de revisões. Outros modelos foram propostos para atender às evoluções com novos níveis para a pirâmide como sinopses para sínteses por exemplo. Diante de um cenário permeado por “aliterações crescentes suplantando definições (por exemplo, sinopses de sínteses de recomendações sistematicamente derivadas) e complexidade ultrapassando a utilidade de tais modelos”, Alper e Haynes (2016) desenvolveram uma nova versão simplificada da pirâmide.

Este modelo pode ser considerado uma versão amadurecida após 15 anos de evolução. Alper e Haynes (2016) explicam as ideias que enriqueceram o modelo de pirâmide. A primeira é os três primeiros níveis apresentarem uma visão filtrada e pré avaliada das evidências de estudos originais, revisões sistemáticas e *guidelines*. Além dessa visão, esses níveis também podem apresentar sinopses construídas para facilitar o acesso às evidências. A segunda ideia é cada nível se desenvolver a partir dos níveis abaixo, fazendo com que a informação encontrada em níveis acima melhor oriente as tomadas de decisão clínica. A terceira ideia é a configuração do quarto nível da pirâmide que inclui sumários sintetizados integrando os três níveis inferiores com objetivo de ser referência clínica.

O quarto nível da pirâmide pode ser aquele pelo qual todo profissional de saúde deve iniciar uma busca, pois ele apresenta um sumário da informação baseada em evidência proveniente de qualquer um dos tipos de estudo dos níveis inferiores (estudo original, revisão sistemática ou *guideline*). Para Alper e Haynes (2016), o desenvolvimento de uma ferramenta de busca que organize os

resultados por níveis da pirâmide permite encontrar as melhores respostas para um problema ou condição de saúde.

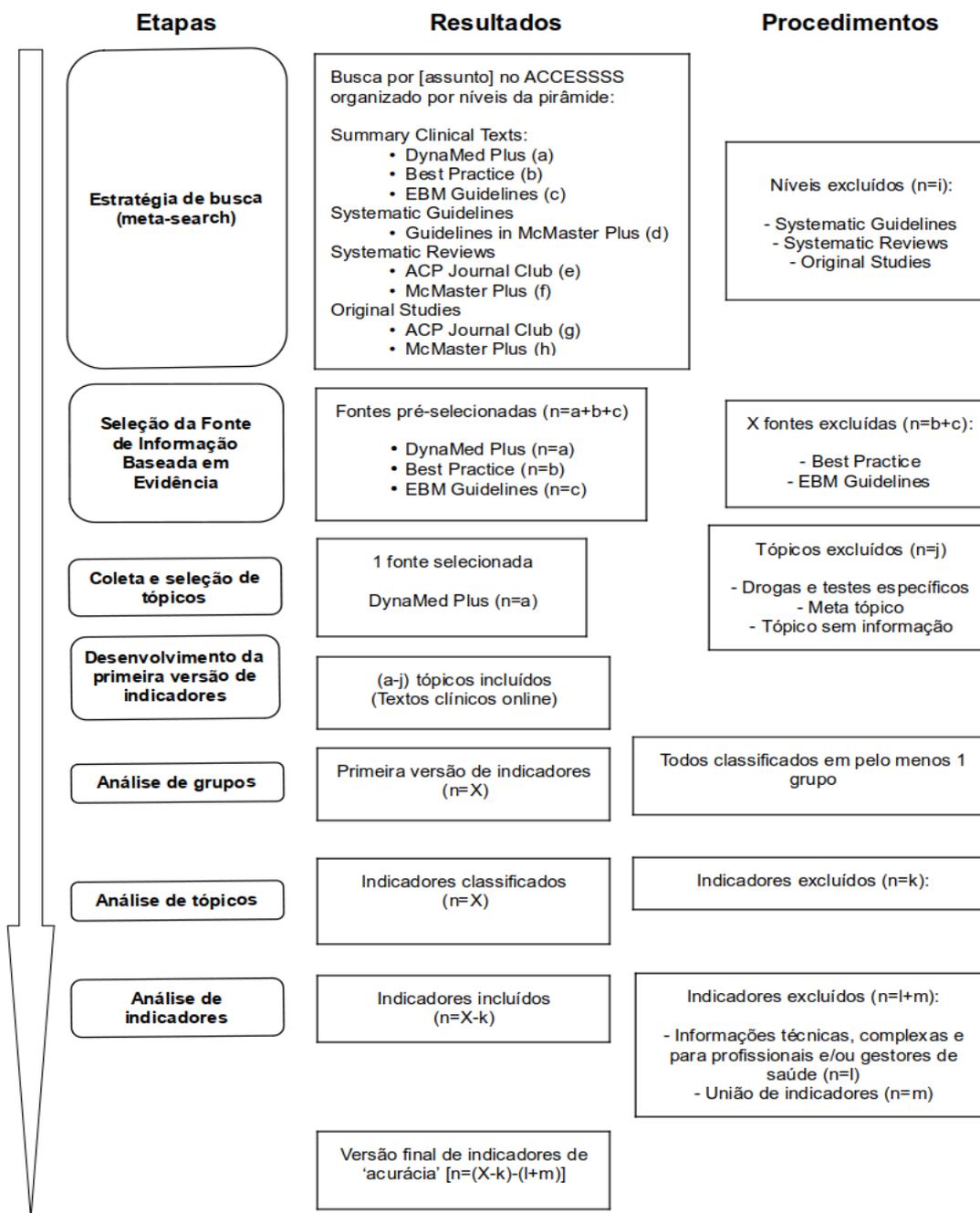
Decidimos nos apropriar desse modelo no desenvolvimento dos métodos para construção de indicadores de acurácia para avaliação da qualidade da informação de saúde na internet. De forma geral, Dicenso, Bayley e Haynes (2009) orientam como seguir e aplicar as ideias desse modelo para adquirir as melhores e atuais evidências. Mas nossa principal referência para obter a melhor e atual evidência e desenvolver os métodos para construção dos indicadores de acurácia é o capítulo de Burwell (2019).

### **3 Resultados e discussão: métodos para construção de indicadores**

Nossa principal referência para o desenvolvimento dos métodos orientados pela MBE para construção de indicadores de acurácia é o capítulo de Burwell (2019, tradução nossa) intitulado “Aquisição de evidências: Como encontrarmos as melhores e atuais evidências e sermos encontrados por elas”. Trata-se do capítulo 2 de uma das principais referências no campo da MBE, o livro “Medicina Baseada em Evidências: Como praticar e ensinar MBE”. A primeira edição deste livro data de 1997 e atualmente está em sua quinta edição.

O conjunto de métodos desenvolvido é composto por sete etapas: estratégia de busca, seleção da fonte de informação baseada em evidência, coleta e seleção de tópicos baseados em evidência, desenvolvimento da primeira versão de indicadores, análise dos grupos, análise dos tópicos e análise dos indicadores. A Figura 1 apresenta o diagrama de fluxo metodológico contendo as sete etapas, os possíveis resultados a serem obtidos e possibilidades de procedimentos aplicados em cada uma delas. Cada etapa é explicada detalhadamente nesta seção.

Figura 1 – Diagrama de fluxo metodológico



Fonte: Elaborado pelos autores

### 3.1 Estratégia de busca

Inicialmente, é necessário definir onde a busca será realizada, ou seja, qual ou quais serviços de informação baseada em evidência serão utilizados como fonte

de informação. O objetivo de trabalhos que pretendam desenvolver indicadores de acurácia para avaliar sites de saúde não deve ser reunir todas as evidências disponíveis e apresentar um estado da arte sobre determinado assunto. Neste campo de avaliação, o objetivo deve ser desenvolver indicadores de acurácia baseados em evidência para avaliar sites de saúde que contenham informação de interesse para seus públicos alvos. Portanto, ao invés de buscar em vários serviços, há uma opção mais adequada e viável que atende à esse objetivo, o uso de serviço de *meta-search*. É necessário definir onde realizar esse tipo de busca, a tarefa inicial da primeira etapa metodológica.

Burwell (2019) apresenta três serviços de *meta-search* baseados em evidência, mas destaca um deles: o ACCESSSS. Ele é mantido pela canadense *McMaster University*. De acordo com o site oficial da ferramenta (MCMASTER, 2019a), as evidências disponibilizadas são obtidas a partir de mais de 120 periódicos clínicos através de um processo rigoroso de avaliação crítica realizada por especialistas. Essa parceria criou o serviço *McMaster Premium Literature Service* (McMasterPLUS™). Assim, eles garantem que novas evidências sejam rapidamente processadas e incorporadas.

O ACCESSSS também se destaca por ser baseado no modelo da “Pirâmide 5.0” (ALPER; HAYNES, 2016; BURWELL, 2019). Além de realizar a *meta-search* em serviços de informação baseada em evidência, os resultados encontrados são apresentados de forma ordenada de acordo com a estrutura da “Pirâmide 5.0” (BURWELL, 2019). Em uma única página, encontramos uma visão geral sobre um assunto. Essa visão fornece subsídios para a tomada de decisão sobre as próximas etapas de construção de indicadores de acurácia. Portanto, o ACCESSSS é a ferramenta de *meta-search* que compõe nosso conjunto de métodos.

Nesse sentido, o método desta etapa é realizar a busca pelo assunto de saúde de interesse no ACCESSSS. O resultado da busca é apresentado conforme a coluna “Resultados” da Figura 1. No nível “*Original Studies*” (base da pirâmide), é apresentado o total encontrado em cada fonte de informação baseada em evidência sobre sinopses de estudos originais pré avaliados. No próximo nível “*Systematic Reviews*”, o total encontrado em cada fonte sobre

estudos de revisão sistemática. No terceiro “*Systematic Guidelines*”, o total encontrado em cada fonte sobre *guidelines*. No quarto nível “*Summary Clinical Texts*”, o ACCESSSS apresenta o total de tópicos encontrado em cada fonte de textos clínicos resumidos. Este nível apresenta resumos integrados das três primeiras camadas da pirâmide. Isto significa o seguinte: se a melhor evidência sobre um determinado aspecto do assunto de saúde estiver em um estudo clínico, em uma revisão sistemática ou em um *guideline*, ela provavelmente está disponível de forma resumida nesse quarto nível da pirâmide, independente do tipo de material pré avaliado (BURWELL, 2019).

Assim, prioritariamente, nosso método consiste em utilizar os resultados encontrados no quarto nível para selecionar a fonte de informação a ser consultada para construção dos indicadores. Por esta razão, devem ser aplicados os procedimentos dispostos nesta etapa na Figura 1.

### 3.2 Seleção da fonte de informação baseada em evidência

O total de tópicos resultante do quarto nível da busca realizada no ACCESSSS pode ser extenso, disperso em diferentes serviços e pouco circunscrito para determinado assunto de um site de saúde. De acordo com os termos escolhidos para realizar a busca, a *meta-search* pode apresentar mais de um serviço de textos clínicos resumidos, conforme a coluna “Resultados” desta etapa na Figura 1. Neste caso, será necessário selecionar o melhor serviço para cada caso. Esta etapa metodológica apresenta métodos para fazer essa seleção.

A primeira possibilidade é verificar estudos que tenham avaliado os serviços encontrados. O estudo de Prorok et al. (2012) apresentou resultados da avaliação de serviços desse tipo em relação à pontualidade na atualização do conteúdo, à amplitude da cobertura e à qualidade da evidência oferecida. O estudo analisou dez serviços e elaborou um *ranking* de acordo com uma escala de 1 a 10 para cada um dos três critérios observados. Nessa escala, o número 1 é a melhor avaliação. O serviço denominado *DynaMed Plus* recebeu a melhor avaliação geral e ocupou a primeira posição do *ranking* desenvolvido por Prorok et al.(2012). O resultado da avaliação realizada por estudos como o de Prorok et

al.(2012) é o primeiro parâmetro que pode orientar a escolha da fonte de informação para construção de indicadores de acurácia.

Como o *DynaMed Plus* se destaca dentre os serviços de informação baseada em evidência encontrados no quarto nível da “Pirâmide 5.0” na busca realizada no ACCESSSS, investigamos essa fonte e como ela se estrutura. Para conhecer o *DynaMed Plus*, visitamos o site oficial. O serviço é conduzido por profissionais vinculados à universidades localizadas nos Estados Unidos (DYNAMED, 2019b). Ele conta com os seguintes parceiros: *American College of Physicians*, *American Medical Association*, *American Association of Nurse Practitioners*, *Guidelines International Network*, *McMaster University*, *Micromedex® Solutions*, *Wikipedia* e *Zynx Health* (DYNAMED, 2019c). Além disso, o site têm uma página sobre sua política de conflito de interesses (DYNAMED, 2019d).

Contudo, no âmbito da MBE, fomentamos a ideia de que:

[...] processos explícitos para encontrar, avaliar, incorporar e atualizar evidências relativas ao diagnóstico, curso, causa, prevenção, tratamento e reabilitação de problemas de saúde são mais importantes do que a posição acadêmica ocupada. (BURWELL, 2019, p. 61, tradução nossa)

Neste sentido, fizemos uma avaliação do *DynaMed Plus* em relação aos seus processos com tais finalidades que configura um segundo método para fornecer parâmetros para seleção de uma fonte. Utilizamos a ferramenta apresentada no referencial metodológico, o “*Guides for judging whether an online clinical text is evidence based and current*” (BURWELL, 2019). O resultado foi o cumprimento de 11 dos 14 critérios. Assim, podemos concluir que os tópicos do *DynaMed Plus* são baseados em evidência, pois seguem um padrão que atende a quase todos os critérios sugeridos por Burwell (2019).

Há, ainda, um terceiro método de seleção. Outro aspecto relevante que um serviço de informação baseado em evidência deve dispor é uma política de níveis de evidência para seus textos e os métodos empregados para que tais níveis sejam estabelecidos. No caso do *DynaMed Plus*, é disponibilizado um documento PDF online que descreve como a evidência pode ser classificada em

três níveis: *Level 1 (likely reliable) Evidence*, *Level 2 (mid-level) Evidence* e *Level 3 (lacking direct) Evidence*.

Portanto, consideramos que o *DynaMed Plus* é o melhor serviço atualmente e compõe nosso método. Ele deve ser a primeira opção de fonte de informação baseada em evidência para construção de indicadores de acurácia para avaliação da qualidade da informação em sites de saúde. Logo, as outras fontes podem ser excluídas conforme a coluna “Procedimentos” desta etapa na Figura 1.

### **3.3 Coleta e seleção de tópicos baseados em evidência**

A busca realizada no ACCESSSS resulta em tópicos encontrados no serviço do *DynaMed Plus* (Figura 1). De acordo com o assunto pesquisado, podem ser encontrados poucos ou muitos tópicos. Caso sejam encontrados poucos tópicos e eles atendam às necessidades informacionais para construção dos indicadores de acurácia sobre um determinado assunto de interesse, o trabalho poderá seguir para a coleta da informação baseada em evidência. Este é o cenário ideal. No entanto, as buscas podem resultar em um número elevado de tópicos que oferecem informação sobre o assunto desejado. Neste caso, é necessário fazer uma seleção dos tópicos adequados aos indicadores que pretendemos desenvolver.

Para lidar com um grande volume de tópicos, inclusive porque alguns são estritamente técnicos, elaboramos quatro critérios de exclusão que podem orientar na circunscrição de tópicos interessantes para geração de indicadores de avaliação de sites de saúde: (1) drogas específicas que parecem bula: tais informações entram em muitos detalhes que provavelmente não trarão benefícios se dispostas em sites destinados aos usuários finais; (2) teste diagnóstico específico: um tópico contendo informações técnicas sobre um teste como tipos de medidas utilizadas; (3) meta tópico: o *DynaMed Plus* apresenta tópicos contendo listas de *links* para outros; (4) não fornece informação: o *DynaMed Plus* apresenta tópicos sobre ferramentas como calculadoras de saúde.

Em nosso método, para aplicar esses quatro critérios de exclusão, os tópicos devem ser acessados e a primeira seção de cada página deve ser avaliada, na maioria dos casos, a seção “*Overview and Recommendations*”, comum nos tópicos do *DynaMed Plus*. Esse procedimento permite excluir tópicos das amostras das pesquisas conforme a coluna “Procedimentos” desta etapa na Figura 1. Com a inclusão de uma amostra mais circunscrita de tópicos, podemos avançar para a quarta etapa metodológica e construir a primeira versão de indicadores de acurácia para avaliação da qualidade de sites de saúde.

### **3.4 Desenvolvimento da primeira versão de indicadores**

Os tópicos encontrados são textos clínicos online com a finalidade de oferecer informação acessível das melhores evidências científicas disponíveis. Entendemos que tal informação é acessível por resumir a evidência, fazendo com que o usuário, profissional de saúde ou não, obtenha informação sobre o tratamento mais eficaz e opções de diagnósticos e formas de prevenção de doenças mais eficientes. Desta forma, o usuário não precisa possuir conhecimentos avançados para compreender métodos estatísticos como aqueles comumente encontrados em estudos de revisão sistemática por exemplo. Contudo, o objetivo desses textos clínicos online não é a construção de indicadores para avaliar sites de saúde.

Elaboramos um método para desenvolver indicadores de acurácia utilizando a informação baseada em evidência disponível nos tópicos do *DynaMed Plus*. Primeiro, é necessário compreender a estrutura dos tópicos como as informações estão dispostas nesse site. Esta fonte tem a vantagem de ser padronizada. A informação baseada em evidência encontrada em cada tópico no *DynaMed Plus* é organizada em seções. Geralmente, a primeira seção desse menu lateral esquerdo da página de um tópico apresenta a opção “*Overview and Recommendations*” que, quando expandida, mostra as três subseções “*Background*”, “*Evaluation*” e “*Management*”. Esta primeira seção contém um resumo das evidências. Trata-se de um texto clínico claro e direto. O usuário pode ter uma visão geral sobre os diferentes aspectos que envolvem o tópico

pesquisado. Neste sentido, cada seção e subseção contém informação contextualizada. De forma geral, na subseção “*Background*”, é apresentada informação como a definição do assunto do tópico estatísticas epidemiológicas, fatores de risco e formas de transmissão. Quando há casos específicos sobre alguma informação, é comum encontrar ponderações e recomendações sobre a necessidade de conhecer melhor a evidência para considerá-la adequadamente em cada caso. Em “*Evaluation*”, as evidências resumidas abordam informação relacionada aos sintomas e às formas de diagnóstico existentes. Nesta subseção, também são encontradas ponderações sobre os diferentes sintomas e possibilidades de diagnóstico. A última subseção “*Management*” contém informação sobre os melhores tratamentos comprovados, assim como efeitos colaterais dos mesmos.

De acordo com este padrão de informação em cada tópico, o procedimento inicial de nosso método é analisar a possibilidade de seus conteúdos serem utilizados na composição de indicadores de acurácia, considerando que serão aplicados na avaliação de sites destinados aos usuários interessados em determinado tema de saúde. Na maioria dos casos, há informação relevante nessa primeira seção. Dependendo do tema de saúde, é necessário recorrer às outras seções como “*Diagnosis*”, “*Treatment*” e “*Prevention and Screening*”. Essas apresentam informação detalhada sobre a evidência. É possível encontrar dados estatísticos, opções para diferentes populações e os estudos que foram utilizados para sustentar os textos clínicos resumidos.

Com a compreensão da estrutura informacional de seções e subseções dos tópicos, o próximo passo é desenvolver os indicadores. O próximo método consiste no seguinte. Para cada tópico incluído (coluna “*Resultados*” da Figura 1), primeiro, realizamos a leitura das seções iniciais. Após a leitura, para cada indicador encontrado, deve ser feita a redação de uma versão inicial, o registro da evidência na qual é baseado e a referência do *DynaMed Plus*.

Além disso, outro método deve ser previamente aplicado nesta etapa, é importante classificar os indicadores em grupos. Inicialmente, podem ser especificados alguns grupos. Outros podem ser gerados através da classificação

resultante da leitura e abstração do contexto da evidência. Cada grupo representa um assunto dentro do tema de saúde avaliado. Por exemplo, no caso da tuberculose apresentado anteriormente, um indicador pode ser redigido assim: Os sintomas sugestivos da tuberculose pulmonar são febre, fadiga, perda de peso, sudorese noturna, hemoptise ou tosse por duas ou mais semanas (DYNAMED, 2019e). Este indicador pode ser classificado em um grupo de indicadores denominado “Sintomas”.

A primeira versão de indicadores pode ser considerada um resultado bruto que precisa ser revisado, corrigido, adaptado, ou seja, passar por um processo de melhorias. A finalidade é chegar a um conjunto de indicadores viável para ser aplicado na avaliação de sites de saúde destinados aos usuários. Neste sentido, as próximas etapas consistem em processos de análise.

### **3.5 Análise dos grupos**

Nesta primeira etapa de análise, o primeiro procedimento do nosso método é verificar se todos os indicadores da primeira versão (Figura 1) estão em pelo menos um dos grupos criados. Caso algum indicador não esteja no âmbito de um grupo, devemos refletir e abstrair uma classificação. Não tem problema se houver indicadores classificados em dois grupos simultaneamente. Após classificar todos os indicadores em pelo menos um grupo, conforme coluna “Procedimentos” desta etapa na Figura 1, devemos obter uma visão geral da quantidade de indicadores em cada grupo, assim como os eventuais cruzamentos entre eles.

Esta análise deve ser considerada porque o total de indicadores desenvolvidos em uma primeira versão provavelmente será elevado e, assim, impraticável quando se trata da avaliação de sites de saúde. Então, o próximo método é realizar uma análise a partir da visão geral. Precisamos de mecanismos para lidar com um volume elevado de indicadores e refletir sobre como podem ser circunscritos para atender aos objetivos de avaliações de sites. Com esta análise, é possível identificar como os indicadores estão distribuídos e registrar indicativos de exclusão ou união entre eles para as etapas seguintes. O intuito de

construir indicadores não é reunir todas as evidências disponíveis sobre determinado tema de saúde. Assim, o objetivo é reduzir o total de indicadores para um conjunto viável e com informação essencial para avaliar sites destinados aos usuários.

### **3.6 Análise dos tópicos**

No mesmo sentido de circunscrever a primeira versão de indicadores desenvolvida, pode ser necessário analisar novamente os tópicos incluídos como amostra para construção dos indicadores. Para reduzir o total de indicadores a um conjunto que disponha de informação essencial para avaliar sites de saúde destinados aos usuários, podem ser estabelecidos critérios de exclusão para os tópicos encontrados no *DynaMed Plus* e selecionados na terceira etapa (Figura 1).

Os seguintes métodos podem ser empregados. Uma ideia é excluir tópicos sobre assuntos relacionados à pequenos grupos de pessoas, casos excepcionais ou, ainda, tópicos considerados de baixa relevância para a equipe envolvida no desenvolvimento dos indicadores. Outra possibilidade é excluir tópicos que contém informação muito técnica com detalhes sobre métodos diagnóstico, dosagens de medicamentos, etimologia, etc.

Com a aplicação de critérios de exclusão, conforme coluna “Procedimentos” desta etapa na Figura 1, o total de indicadores tende a ser reduzido. Caso isto não ocorra, a nova quantidade de indicadores ainda pode ser elevada e provavelmente difícil de ser empregada na avaliação de sites de saúde destinados aos usuários. Se assim acontecer, recomendamos seguir para os métodos da próxima etapa. Independentemente dessa possibilidade, a redação dos indicadores é aquela da primeira versão considerada um resultado bruto. Estas redações precisam ser revisadas, pois muitos indicadores podem ser corrigidos, reformulados, unidos ou até excluídos. Portanto, a próxima etapa consiste na reflexão sobre cada um dos indicadores.

### 3.7 Análise dos indicadores

Durante o desenvolvimento da primeira versão de indicadores, não há a preocupação de redigir os indicadores em suas formas finais. Nossa preocupação naquele primeiro momento é mais próxima de uma coleta de dados para obter os indicadores. Muitos deles podem ser redigidos quase como traduções dos textos clínicos originais encontrados no *DynaMed Plus*. A intenção é que os indicadores contenham informação passível de observação nos sites de saúde.

Esta análise permite aplicar três métodos distintos que desencadeiam na reformulação, união ou exclusão de indicadores incluídos até esta etapa (Figura 1). O primeiro procedimento é a exclusão de mais indicadores segundo critérios de exclusão. O segundo é a união de indicadores com conteúdos semelhantes ou complementares, fazendo com que o total de indicadores diminua ainda mais. O terceiro procedimento também diz respeito à união de conteúdos, mas especificamente na permutação de conteúdos que podem ficar melhor redigidos ao transferir um trecho de um indicador para outro, sem que nenhum seja excluído. Detalhamos esses métodos nesta etapa.

No primeiro procedimento, podem ser definidos critérios de exclusão como: muita informação com termos técnicos; várias ponderações sobre diferentes formas de diagnóstico ou tratamento complexos ou complementares; e informações destinadas aos profissionais ou gestores de sistemas de saúde. Esse tipo de informação provavelmente não traz benefícios se o usuário não tiver conhecimento técnico e, conseqüentemente, a disponibilização delas em sites destinados aos usuários poderia não ter efeito positivo. Além disso, o público-alvo mais adequado para esse tipo de informação é o profissional de saúde. As evidências encontradas em indicadores desse tipo são apropriadas para o contexto da relação entre profissional de saúde e usuário. Assim, o profissional de saúde tem maiores condições de julgar a informação baseada em evidência que determinado usuário precisar conhecer.

O segundo procedimento é a união de indicadores com conteúdos semelhantes ou complementares. A união pode ser feita de 3 formas. A primeira é unir indicadores devido às suas informações estarem contempladas em outro

indicador. A segunda forma de união é desenvolver uma nova redação de indicador tendo como base informação proveniente de dois ou mais indicadores. A terceira não reduz o total de indicadores, porque identifica que os indicadores envolvidos ficam melhor redigidos com a transferência de um trecho da evidência de um indicador para o outro.

Durante o procedimento de análise dos indicadores para exclusão ou união, todos devem ser avaliados quanto às suas redações. Nossa intenção é deixar a redação fácil de ser compreendida. Desta forma, os indicadores poderão ser empregados na avaliação de sites de saúde tanto por profissionais de saúde como por usuários. A redação final dos indicadores desenvolvidos através desses métodos, assim como o processo que levou ao seu desenvolvimento, seja ele a união de indicadores e/ou a reformulação da redação, devem ser relatadas nos estudos deste campo de avaliação transparentemente, conforme coluna “Procedimentos” desta etana na Figura 1. A apresentação de cada um dos indicadores de acurácia desenvolvidos também deve estar acompanhada pela explicação de como cada indicador deve ser aplicado na avaliação de sites.

#### **4 Conclusões**

O principal objetivo deste estudo foi desenvolver métodos para construção de indicadores de acurácia para avaliação de informação de saúde na Internet. O conjunto de métodos desenvolvido é composto por sete etapas.

Para alcançar esses resultados, construímos um conjunto de métodos respaldados pelos fundamentos da ‘medicina baseada em evidências’ (MBE) e orientados por práticas aprimoradas ao longo de mais de cinco décadas de evolução. Se a prática da MBE é um desafio para os profissionais de saúde, o acesso à informação baseada em evidência e legível aos cidadãos parece um desafio ainda maior. Contudo, o acesso à informação de saúde na Internet por parte do público em geral é uma realidade a lidar.

Nossa proposta é que os sites de saúde direcionados aos cidadãos sejam avaliados com indicadores de acurácia desenvolvidos através das práticas da MBE. Assim, este trabalho pode contribuir tanto para o fomento da MBE dentre

os profissionais de saúde que produzem informação na Internet como para o empoderamento do cidadão por meio do acesso à informação de qualidade. Este trabalho reitera a importância dos indicadores baseados em evidência e nos instiga a refletir sobre implicações para o campo de avaliação da qualidade da informação de saúde disponível na Internet.

Este trabalho se inscreve no debate internacional sobre como definir os indicadores do critério acurácia na avaliação da informação de sites sobre qualquer assunto de saúde. O consenso entre especialistas não apresenta de forma atualizada o conhecimento que é produzido no mundo. Com as várias bases bibliográficas e os métodos orientados pela MBE, é possível chegar a conclusões atualizadas e sintonizadas com o atual estágio do conhecimento científico sobre determinado assunto de saúde.

Portanto, defendemos a atualização do critério acurácia. Propomos a nova definição: *critério para avaliar a conformidade da informação com a melhor e a mais atual evidência científica disponível*. Esta evidência deve ser obtida a partir das práticas da ‘medicina baseada em evidência’. A atualização do conceito tem implicações para o campo de avaliação da qualidade da informação de saúde na Internet.

Ao longo dos anos, foram desenvolvidos vários métodos para a prática da MBE. Neste trabalho, nossa posição se configura como a de consumidores das evidências. Compreendemos suas formas, onde e como obtê-las. Elas funcionam como insumos para construção dos indicadores de acurácia. A MBE orienta as melhores práticas para alcançar esses objetivos. Adquirir ‘evidência pré avaliada’ facilitou esse processo. Estruturamos um percurso metodológico que pode ser replicado.

Estudos futuros que se proponham a avaliar a qualidade da informação de saúde na Internet utilizando o critério acurácia devem adquirir as melhores e atuais evidências científicas na área de interesse para desenvolver seus indicadores de avaliação. Neste artigo, com as práticas da MBE, demonstramos que é possível desenvolver indicadores dessa forma.

Pesquisas futuras podem desenvolver instrumentos para medir a qualidade de estudos que avaliam a informação encontrada em sites de saúde. Para observar os métodos dos estudos, esses instrumentos podem incorporar critérios que verifiquem se os indicadores de acurácia foram desenvolvidos de acordo com as práticas da MBE. Os estudos que continuarem utilizando como padrão de qualidade o consenso entre especialistas, livros ou artigos selecionados sem explicitar e justificar tais escolhas, podem ser considerados de baixa qualidade.

Esperamos que outra contribuição deste trabalho seja a inscrição do tema da avaliação da qualidade da informação de saúde na Internet nas agendas de pesquisa e de atuação das instituições de saúde brasileiras, sejam elas públicas, privadas ou associações profissionais.

### **Financiamento**

Fundação Oswaldo Cruz.

### **Referências**

ALBARRAK, A. I. et al. The impact of obesity related websites on decision making among students in Saudi Arabia. **Saudi Pharmaceutical Journal**, [s.l.], v. 24, n. 5, p. 605–610, 2016.

ALPER, B. S.; HAYNES, R. B. EBHC pyramid 5.0 for accessing preappraised evidence and guidance. **Evidence-based medicine**, London, v. 21, n. 4, p. 123–5, 1 ago. 2016.

ATALLAH, Á. N. Medicina baseada em evidências [editorial]. **Diagn Tratamento**, v. 23, n. 2, 2018.

BETTING, L. E. et al. Tratamento de Epilepsia: Consenso dos Especialistas Brasileiros. **Arquivos de Neuro-Psiquiatria**, São Paulo, v. 61, n. 4, p. 1045–1070, dez. 2003.

BRASIL. **Ministério da Saúde. Tuberculose**. 2021 Disponível em: <<http://antigo.saude.gov.br/saude-de-a-z/tuberculose>>. Acesso em: 5 ago. 2021.

BURWELL, D. S. Acquiring the evidence: How to find current best evidence and have current best evidence find us. In: STRAUS, S. E. et al. (Eds.). .

**Evidence-based medicine: how to practice and teach EBM.** 5. ed. [S.l.]: Elsevier Health Sciences, 2019. p. 57–90.

CHAPA, D. et al. Using preappraised evidence sources to guide practice decisions. **Journal of the American Association of Nurse Practitioners**, Austin, v. 25, n. 5, p. 234–243, 1 maio 2013.

DICENSO, A.; BAYLEY, L.; HAYNES, R. B. Accessing pre-appraised evidence: fine-tuning the 5S model into a 6S model. **Evidence-based nursing**, London, v. 12, n. 4, p. 99–101, 1 out. 2009.

DYNAMED, (INTERNET). **About DynaMed | Clinical Decision Support Tool.** 2019a. Disponível em: <<https://www.dynamed.com/home/about>>. Acesso em: 7 maio. 2019.

DYNAMED, (INTERNET). **Executive Leadership | Clinical leaders| Clinical Influencers | DynaMed Plus.** 2019b. Disponível em: <<https://www.dynamed.com/home/about/executive-leadership>>. Acesso em: 27 set. 2019.

DYNAMED, (INTERNET). **DynaMed Partnerships | DynaMed.** 2019c. Disponível em: <<https://www.dynamed.com/home/about/partners>>. Acesso em: 27 set. 2019.

DYNAMED, (INTERNET). **Conflict of Interest Policy | DynaMed.** 2019d. Disponível em: <<https://www.dynamed.com/home/editorial/conflict-of-interest-policy>>. Acesso em: 27 set. 2019.

DYNAMED, (INTERNET). **Record No. T116300, Pulmonary Tuberculosis.** 2019e. EYSENBACH, G. et al. Empirical Studies Assessing the Quality of Health Information for Consumers on the World Wide Web. **Journal of the American Medical Association**, Chicago, v. 287, n. 20, p. 2691–2700, 22 maio 2002.

EYSENBACH, G. The impact of the Internet on cancer outcomes. **Cancer Journal For Clinicians**, [s.l.], v. 53, n. 6, p. 356–371, 2003.

HAYNES, R. B. Of studies, summaries, synopses, and systems: the “4S” evolution of services for finding current best evidence. **Evidence Based Mental Health**, London, v. 4, n. May, p. 37–41, 2001.

HAYNES, R. B. Of studies, syntheses, synopses, summaries, and systems: the 5S evolution of information services for evidence-based healthcare decisions. **Evidence-based medicine**, London, v. 11, n. 6, p. 162–4, 1 dez. 2006.

LAGAN, B. M.; SINCLAIR, M.; KERNOHAN, W. G. What Is the Impact of the Internet on Decision-Making in Pregnancy? A Global Study. **Birth-issues in Perinatal Care**, [s.l.], v. 38, n. 4, p. 336–345, 2011.

- MCMASTER, U. ACCESSSS | About. 2019a. Disponível em:  
<<https://www.accessss.org/Pages/About>>. Acesso em: 30 maio. 2019.
- MORENO, J. M. et al. A quality evaluation methodology for health-related websites based on a 2-tuple fuzzy linguistic approach. **Soft Computing**, [s.l.], v. 14, n. 8, SI, p. 887–897, jun. 2010.
- NADANOVSKY, P. Epidemiologia aplicada a clínica: um enfoque científico do uso da informação médica - Medicina Baseada em Evidência. **Estudos em Saúde Coletiva**, IMS/UERJ - Rio de Janeiro - Br, v. 185, p. 1–27, 1999.
- PAOLUCCI, R.; PEREIRA NETO, A. Methods for evaluating the quality of information on health websites: Systematic Review (2001-2014). **Latin American Journal of Development**, Curitiba, v. 3, n. 3, p. 994–1056, 14 maio 2021.
- PARK, H. et al. Can a health information exchange save healthcare costs? Evidence from a pilot program in South Korea. **International Journal of Medical Informatics**, [s.l.], v. 84, n. 9, p. 658–666, 2015.
- PEREIRA NETO, A.; PAOLUCCI, R. Evaluation of the Quality of Health Information on the Internet: An Analysis of Brazillian Initiatives. In: PEREIRA NETO, A.; FLYNN, M. B. (Eds.). **The Internet and Health in Brazil**. [s.l.]: Springer International Publishing, 2019. p. 181–209.
- POWELL, J. et al. The Characteristics and Motivations of Online Health Information Seekers: Cross-Sectional Survey and Qualitative Interview Study. **Journal of Medical Internet Research**, Toronto, v. 13, n. 1, p. e20, 23 fev. 2011.
- PROROK, J. C. et al. The quality, breadth, and timeliness of content updating vary substantially for 10 online medical texts: An analytic survey. **Journal of Clinical Epidemiology**, Philadelphia, v. 65, n. 12, p. 1289–1295, 2012.
- SACKETT, D. L. et al. Evidence based medicine: what it is and what it isn't. **British Medical Journal (Clinical research ed.)**, London, v. 312, n. 7023, p. 71–2, 13 jan. 1996.
- SILVA-JEREZ, N. S.; FURNIVAL, A. C. Informação sobre saúde: letramento informacional e acesso aberto à literatura científica. **Em Questão**, Porto Alegre, v. 24, n. 2, p. 260, 19 abr. 2018.
- SMITH, R.; RENNIE, D. Evidence-based medicine: an oral history. **Journal of the American Medical Association**, Chicago, v. 311, n. 4, p. 365, 22 jan. 2014.
- SPOELMAN, W. A. et al. Effect of an evidence-based website on healthcare usage: an interrupted time-series study. **British Medical Journal Open**, London, v. 6, n. 11, 2016.

ZHANG, Y.; SUN, Y.; XIE, B. Quality of health information for consumers on the web: A systematic review of indicators, criteria, tools, and evaluation results. **Journal of the Association for Information Science & Technology**, [s.l.], v. 66, n. 10, p. 2071–2084, 26 may 2015

Recebido: 05/02/2021

Aceito: 08/06/2021

### **Declaração de autoria**

**Concepção e elaboração do estudo:** Rodolfo Paolucci Pimenta, André de Faria Pereira Neto, Paulo Nadanovsky

**Coleta de dados:** Rodolfo Paolucci Pimenta

**Análise e interpretação de dados:** Rodolfo Paolucci Pimenta, André de Faria Pereira Neto

**Redação:** Rodolfo Paolucci Pimenta, André de Faria Pereira Neto

**Revisão crítica do manuscrito:** Rodolfo Paolucci Pimenta

### **Como citar**

PAOLUCCI, Rodolfo; PEREIRA NETO, André Faria; NADANOVSKY, Paulo. Avaliação da acurácia da informação em sites de saúde: Métodos para construção de indicadores baseados em evidência. **Em Questão**, Porto Alegre, v. 27, n. 4, p. 137-188, 2021. DOI: <http://doi.org/10.19132/1008-5245274.137-188>

## Evaluation of information accuracy in health sites: methods for constructing evidence-based indicators

### **Rodolfo Paolucci**

PhD; Foundation for the Support of the Technical School, Rio de Janeiro, RJ, Brazil;

rpprjbr@gmail.com; ORCID:<https://orcid.org/0000-0003-3986-1118>

### **André de Faria Pereira Neto**

PhD; Sergio Arouca National School of Public Health, Rio de Janeiro, RJ, Brazil;

andrepereiraneo@gmail.com; ORCID:<https://orcid.org/0000-0003-3631-8857>

### **Paulo Nadanovsky**

PhD; State University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil;

nadanovsky@ims.uerj.br; ORCID:<https://orcid.org/0000-0003-3345-9873>

**Abstract:** The problem of the quality of health information on the Internet has been pointed out in the last two decades. To address this problem, the evaluation

of websites by professionals, researchers and institutions has been configured as a new field of research and action. One of the main issues in this field is related to the evaluation criteria. Accuracy is the most commonly used criterion, but we have identified a problem in its definition and application: admitting consensus among specialists as a quality parameter. However, the consensus does not guarantee a current and correct result. In this article, we present and discuss methods guided by evidence-based medicine for the construction of accuracy indicators. The set of methods developed comprises seven steps: search strategy, selection of evidence-based information sources, collection and selection of topics, development of the first version of indicators, and analysis of groups, topics and indicators. Our methods ensure the development of indicators according to the best and current scientific evidence. They solve the problem of the accuracy criterion for the field of evaluation of information quality on health sites and propose a new definition. The results of this study fill gaps in the knowledge of the field and guide future research and actions because they are replicable for any health subject.

**Keywords:** Evidence-based medicine; Internet and access to information; /Criteria; Indicators of accuracy; Health Information; Health Sites.

## **1 Introduction**

Access to quality health information on the Internet can have positive consequences for both health system managers and citizens. For health managers, access to quality information reduces the costs of health systems because it decreases the search and use of services offered (PARK et al., 2015; SPOELMAN et al., 2016). For citizens, for example, almost two decades ago, the benefits obtained by most cancer patients who accessed information online were reported (EYSENBACH, 2003). The author discusses the impact of information obtained from health sites. When this information is of good quality, patients can obtain knowledge, be less confused, have realistic expectations and thus share decisions, adhere to treatments and effectively undertake self-care (EYSENBACH, 2003). Quality information promotes the development of skills that give citizens greater decision-making power about their health and self-care (ALBARRAK et al., 2016; LAGAN; SINCLAIR; KERNOHAN, 2011).

However, the quality of health information on the Internet is a problem requiring more research. In the first systematic review on the subject, most of the studies analyzed (70%) concluded that information quality is a problem

(EYSENBACH et al., 2002). A similar result was found in another review published 13 years later. Zhang, Sun and Xie (2015) found that 55.2% of the studies evaluated health information on the Internet as negative and 37% as of varying quality. According to Lagan, Sinclair and Kernohan (2011), when wrong information is accessed, it can cause feelings of fear, anxiety and paranoia. For Powell et al. (2011), incorrect, incomprehensible or outdated information can trigger decision-making harmful to health. Therefore, access to low-quality information can have negative consequences for the health of citizens. Therefore, how to deal with this problem?

The main proposal to address the problem of health information quality on the Internet comprises evaluations initiated by public or private institutions. They have existed for more than twenty years, and the first experiences are predominantly European and North American (MORENO et al., 2010). They can perform rigorous evaluations involving professionals and users in judging the quality of health information on the Internet. Some of these initiatives give a quality seal to sites that meet their criteria (PEREIRA NETO; PAOLUCCI, 2019). The seal is a brand that communicates reliability to the users of the sites and represents institutionally approved information. It is a solution and a service offered to most people who are unable to judge the quality of health information on the Internet, a problem substantiated by Silva-Jerez and Furnival (2018).

One of the main issues discussed in this field is related to the criteria used in the evaluation of health sites (PAOLUCCI; PEREIRA NETO, 2021). The predominant evaluation criteria in the specialized literature are present in three systematic reviews.

Eysenbach et al. (2002) performed the first systematic review. This study gathered the methods and processes adopted in the evaluation of the quality of health information on the Internet. This review analyzed 79 studies published until 2001. The authors identified 86 different quality criteria. Among them, many were employed according to the same rules but were reported under different names. To normalize these nomenclatures and present a methodological framework, they synthesized these 86 denominations into five quality criteria (EYSENBACH et al., 2002). This is one of the main

contributions of this pioneering study in this field. The five criteria are technical, *design*, comprehensiveness, legibility and accuracy.

The second systematic review was published thirteen years later (ZHANG; SUN; XIE, 2015). The sample of this review included studies published exclusively in English between 2002 and 2013. It consisted of 165 different studies. The authors identified 11 evaluation criteria that were grouped into the categories “content” and “design”. The “content” category has criteria contemplated by those summarized in the first systematic review, as they deal with the same rules as the technical criteria, accuracy, comprehensiveness and readability. The same can be stated in relation to the *design* criterion found in the two reviews.

The third systematic review was also completed in 2015 (PAOLUCCI; PEREIRA NETO, 2021). One of the objectives of this study was to update the data found in the first review on the subject (EYSENBACH et al., 2002). To select its sample, the study used a search strategy similar to that of the first review of 2002. The period covered was from 2001 to mid -2014. Different from the second review published in the same year (ZHANG; SUN; XIE, 2015), there was no restriction regarding the languages of the studies that composed the sample analyzed by Paolucci and Pereira Neto (2021). The sample included 279 studies. Paolucci and Pereira Neto (2021) found that the criteria summarized in the first systematic review are still valid. They found that the five criteria were used by most studies in the following proportion: accuracy (64%), readability (38%), comprehensiveness (27%), technical (24%) and *design* (17%) (PAOLUCCI; PEREIRA NETO, 2021). According to this result, the accuracy criterion remains the most commonly used criterion in studies that aim to evaluate the quality of health information on the Internet.

However, we identified a problem in the definition of the rules that constitute the accuracy criterion. In the first systematic review, accuracy was defined as the “[...] degree of agreement of the information provided with the best evidence or with generally accepted medical practice” (EYSENBACH et al., 2002, p. 2695). The degree of agreement of the information with medical practice is usually determined through expert consensus. This process does not

guarantee that the result of this consensus is current and correct. In contrast, the professionals involved may be outdated in relation to the best scientific evidence. They can also present opinions and behaviors that are in disuse.

The concern with practices supported by expert consensus is not new. At the end of the 1990s, there was a growing offer of continuing education courses for physicians in an attempt to keep them updated with the best scientific evidence (NADANOVSKY, 1999). This offer is justifiable because the newly graduated professional may have had access to knowledge through books that are often outdated. However, according to Nadanovsky (1999), continued training did not solve the problem of updating physicians. He states that the consensus of experts is based on the knowledge of the professionals involved, individuals inevitably influenced by personal factors (i.e. beliefs, political positions, their physical and mental health, emotional state, etc.) and by different working conditions (infrastructure, income, motivation, etc.).

The consensus of experts is still advocated in certain situations. Betting et al. (2003, p. 1068), despite claiming that “therapeutic decisions should be based on evidence presented in the literature”, argue that in the face of “[...] the persistence of unanswered clinical questions or that remain controversial [...] the opinion of experts becomes great value” (BETTING et al., 2003, p. 1046). However, the evidence, in addition to reaffirming the “inability of the physician to keep up with the important advances”, demonstrates that there are “[...] extraordinary variations in the clinical behavior of the physician and in the rates of beneficial, useless and harmful interventions” (NADANOVSKY, 1999, p. 22). For these reasons, the promotion of evidence-based medicine (EBM) has intensified. For Atallah (2018, p. 43), “Supporting the practices in medicine based on the best evidence from rigorous methodologies that give them comparative validity is a civilizing process”, because, when observing the history of humanity, such practices allowed discovering cures for lethal diseases and saved thousands of lives.

Given this scenario, maintaining the generally accepted medical practice in the definition of the accuracy criterion is a problem for the field of evaluation of the quality of health information on the Internet. Accuracy indicators were

constructed through expert consensus, scientific technical manuals, medical guidelines, textbooks or literature.

In this sense, the objective of this article is to present and discuss the methods guided by the EBM that we developed for the construction of accuracy indicators. As a secondary objective, this article aims to update the definition of the accuracy criterion for the field evaluating the quality of health information on the Internet.

## 2 Methodological framework

David Sackett is considered the father of EBM because, in the late 1960s at McMaster University, he and colleagues in the department began teaching clinical disciplines with a new proposal (SMITH; RENNIE, 2014). The proposal called *critical appraisal* incorporated epidemiology and statistics and used an approach focused on patient problems. In 1996, Sackett et al. (1996) report that EBM has evolved since its foundation and remained an important topic of discussion for a wide audience composed of health professionals, managers, investors and the general public. They clarify what the EBM is and what it is not. They have the following definition:

Evidence-based medicine is the conscious, explicit and judicious use of the best current evidence in decision-making about the care of individual patients. The practice of evidence-based medicine means integrating individual clinical knowledge with the best available external clinical evidence from systematic research. (SACKETT et al., 1996, p. 71)

They explain what they are calling individual clinical knowledge and the best available external clinical evidence. The first is defined by them as the knowledge that clinicians acquire from the experience of clinical practice. They state that this knowledge results in more assertive diagnoses and a more sensitive posture towards patient's rights and choices. The second is defined by these authors as evidence from clinical research considered relevant, such as basic science research and patient-centered research on diagnostic tests, prognoses, and treatment and prevention interventions. For example, in the case of tuberculosis, there is evidence from clinical research conducted in 2005

stating that coughing for two or more weeks is indicative of the disease, and this measure should be used to make an early diagnosis (DYNAMED, 2019a). In Brazil, the Ministry of Health also reports that coughing for three or more weeks is a symptom of tuberculosis (BRASIL, 2021). For Sackett et al. (1996, p. 72, translation in ssa), the best available external clinical evidence “invalidates previously accepted diagnostic tests and treatments and replaces them with new ones that are more powerful, more accurate, more effective and safer”.

It is important to note that binomial clinical experience and better external evidence are both essential in EBM. The following excerpt summarizes the interdependence of this binomial for the practice of EBM:

Without *clinical expertise*, the practice runs the risk of being tyrannized by the evidence because even excellent external evidence may be inapplicable or inadequate for a particular patient. Without the best current evidence, the practice runs the risk of becoming quickly outdated to the detriment of patients. (SACKETT et al., 1996, p. 72)

Sackett et al. (1996) emphasize the integrative character contained in the definition of EBM, but the patient's role in decision-making seems to be restricted to the manifestation of preferences and values. EBM is focused on the health professional and seems to keep him/her solely responsible for the knowledge and for providing information to patients. In our view of the world, we believe that patients can play a more active role in their health. They can integrate the best and most current scientific evidence with their personal experience, that is, with their experience in the health issue that affects them. Thus, patients are better able to discuss decision-making with health professionals.

## 2.1 Preappraised evidence

According to Dicenso, Bayley and Haynes (2009), in the past, to obtain the best evidence, we needed to learn techniques to search the specialized literature and develop critical assessment skills to identify and understand quality clinical studies. However, the authors explain that many resources were created to facilitate the collection of evidence. They characterize these resources as

*preappraised* because they provide *preappraised evidence*. These resources go through “[...] a filtering process to include only the highest quality studies and are updated regularly so that the evidence we access through these resources are current” (DICENSO; BAYLEY; HAYNES, 2009, p. 99).

For Burwell (2019), the filtering processes of the resources of *preappraised evidence* are explicit, and the parameters for judging the quality of the studies are based on scientific merit and the ease with which the evidence can be applied in clinical decisions. Characterizing such evidence as pre-evaluation highlights its potential for the practice of EBM. The author points out three reasons that justify its potential. The first concerns the fact that professionals do not always know how to perform critical evaluations of studies. If the professional knows criteria to judge the quality of the studies, the second reason is not to have these criteria available when necessary. The third reason is to erroneously apply these criteria. In these cases, *preappraised evidence* is the best option for health professionals, and they should perform more extensive searches in the specialized literature only if they do not find what they need previously evaluated. (BURWELL, 2019).

Ten years ago, there were already resources that offered *preappraised evidence* for various clinical problems (DICENSO; BAYLEY; HAYNES, 2009). Some of them provide summaries of evidence. For Chapa et al. (2013), resource summaries of *preappraised evidence* allow health professionals to quickly access relevant information on conditions and interventions for both specific patients and population profiles. Burwell (2019) presents a guide that helps us to observe whether a resource is actually based on evidence: *Guides for judging whether an online clinical text is evidence-based and current*. The guide consists of 14 items that can be observed to evaluate clinical texts online. The author’s proposal consists of the application of these items in the evaluation of online clinical texts, indicating whether each item is absent or present. For him, a text with fewer than five items present cannot be considered evidence-based text (BURWELL, 2019).

Burwell (2019) presents the *meta-search feature* that assists the effort spent to find sites that offer *preappraised evidence* in online clinical text format.

*Meta-search* is a type of search widely used in Internet services to gather the results of searches on different websites into a single platform. This type of search has become a way of dealing with the large and growing volume of dispersed information that is available on the Internet, including scientific and technological information.

There are numerous sites dedicated to the production and availability of *preappraised evidence*. This volume tends to increase due to the promotion of EBM and its expansion into areas of knowledge beyond medicine. With this growth, sites can be found with critical summaries of randomized clinical studies, systematic reviews, *guidelines* or even sites that consolidate summaries of the *preappraised evidence* of all these types of materials in the form of online clinical texts.

## 2.2 Pyramid 5.0 of evidence -based health care

To teach the practice of EBM, models have been developed and improved. These models guide health professionals in the search and selection of evidence-based information that can be integrated with their experiences and *expertise* in the context of clinical practice. Alper and Haynes (2016) present a summary of these models.

The first model was proposed by Haynes (2001). He proposed a hierarchical structure model in the form of a pyramid called “4S”, the first letter of the names of each level. From the base to the top of the pyramid, this model has four levels of organization of the evidence found in the studies: *Studies*, *Syntheses*, *Synopses* and *Systems*. Five years after this publication, the same author published an editorial update of the model (HAYNES, 2006). In this update, he added the “Summaries” level in the pyramid between systems and synopses. According to the author, the purpose of this summary level is to integrate the evidence from the lower three layers. He explains that the three lower levels represent evidence on only one aspect related to a health problem or condition such as a drug. In addition, its configuration requires that health professionals integrate the findings and, in the case of first-level studies, make

critical evaluations on their own. The new summary level allows a comprehensive view of a health problem or condition.

In the following years, evidence-based information services continued to evolve, in addition to the increase in the volume of publications of both original studies, systematic reviews and even review reviews. Other models were proposed to meet the evolution with new levels for the pyramid, such as synopses for syntheses. Faced with a scenario permeated by “increasing alliterations supplanting definitions (for example, synopses of syntheses of systematically derived recommendations) and complexity surpassing the usefulness of such models”, Alper and Haynes (2016) developed a new simplified version of the pyramid.

This model can be considered a mature version after 15 years of evolution. Alper and Haynes (2016) explain the ideas that enriched the pyramid model. The first is that the first three levels present a filtered and *preappraised* view of the evidence from original studies, systematic reviews and *guidelines*. In addition to this view, these levels may also have synopses constructed to facilitate access to evidence. The second idea is that each level develops from the levels below, making the information found at higher levels a better guide for clinical decision-making. The third idea is the configuration of the fourth level of the pyramid that includes summaries integrating the three lower levels to be a clinical reference.

The fourth level of the pyramid may be the one by which every health professional should start a search because it presents a summary of evidence-based information from any of the lower-level study types (original study, systematic review or *guideline*). For Alper and Haynes (2016), the development of a search tool that organizes the results by levels of the pyramid allows finding the best answers to a problem or health condition.

We decided to appropriate this model in the development of methods for building accuracy indicators to evaluate the quality of health information on the Internet. In general, Dicenso, Bayley and Haynes (2009) guide how to follow and apply the ideas of this model to acquire the best and current evidence. However, our main reference for obtaining the best and current evidence and

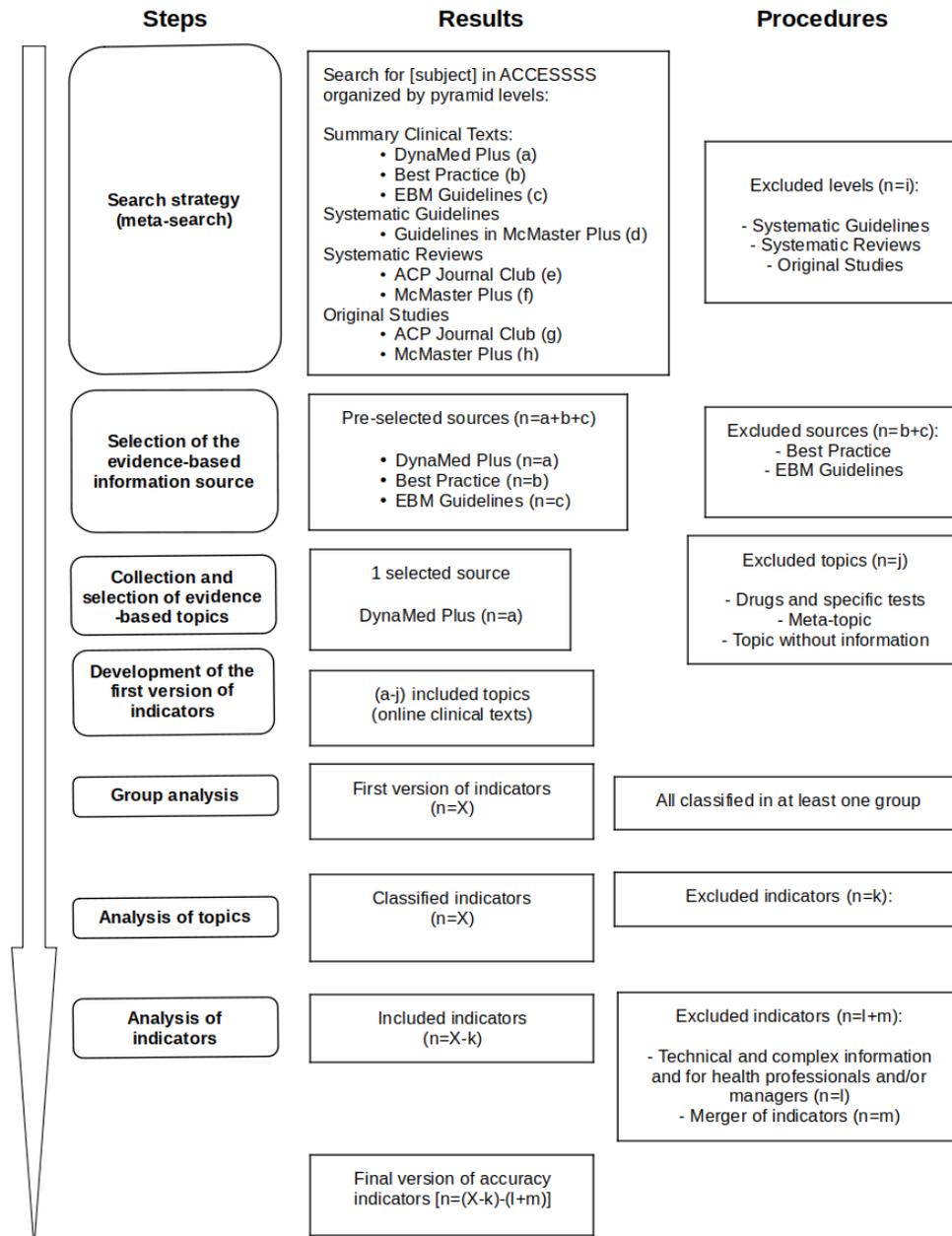
developing methods for the construction of accuracy indicators comes from Burwell (2019).

### **3 Results and discussion: methods for constructing indicators**

Our main reference for the development of methods guided by EBM for the construction of accuracy indicators is the chapter by Burwell (2019) entitled “Acquisition of evidence: How to find the best and current evidence and be found by them”. This is chapter 2 of one of the main references in the field of EBM, the book *Evidence-Based Medicine: How to practice and teach EBM*. The first edition of this book dates from 1997 and is currently in its fifth edition.

The set of methods developed consists of seven steps: search strategy, selection of the evidence-based information source, collection and selection of evidence-based topics, development of the first version of indicators, group analysis, analysis of the topics and analysis of the indicators. Figure 1 shows the methodological flow diagram containing the seven steps, the possible results to be obtained and the possibilities of procedures applied in each of them. Each step is explained in detail in this section.

Figure 1 - Methodological flow diagram



Source: Prepared by the authors

### 3.1 Search strategy

Initially, it is necessary to define where the search will be performed, i.e. which evidence-based information services will be used as a source of information. The objective of studies to develop accuracy indicators used to evaluate health sites should not be to gather all available evidence and present a state-of-the-art on a given subject. In this field of assessment, the objective should be to develop

evidence-based accuracy indicators to evaluate health sites that contain information of interest to their target audiences. Therefore, instead of searching in various services, there is a better option that meets this goal, the use of a *meta-search service*. It is necessary to define where to perform this type of search as the initial task in this methodological step.

Burwell (2019) presents three *evidence-based meta-search* services but highlights one of them: ACCESSSS. It is maintained by Canadian *McMaster University*. According to the official website of the tool (MCMASTER, 2019a), available evidence comes from more than 120 clinical journals through a rigorous process of critical evaluation performed by experts. This partnership created the *McMaster Premium Literature Service* (McMasterPLUS™). Thus, they ensure that new evidence is quickly processed and incorporated.

ACCESSSS also stands out because it is based on the “Pyramid 5.0” model (ALPER; HAYNES, 2016; BURWELL, 2019). In addition to performing the *meta-search* in evidence-based information services, the results are presented in an orderly manner according to the structure of “Pyramid 5.0” (BURWELL, 2019). On a single page, we found an overview of a subject. This view provides support for decision-making on the next steps to build accuracy indicators. Therefore, ACCESSSS is the *meta-search* tool that composes our set of methods.

In this sense, this step consists of searching for the health subject of interest in the ACCESSSS. The search result is presented according to the column “Results” of Figure 1. At the “Original Studies” level (base of the pyramid), the total found in each source of information based on evidence from synopses of original studies is presented. In the next level, “systematic reviews” provides the total number of systematic review studies found in each source. In the third “*Systematic Guidelines*” lists total found in each source on *guidelines*. In the fourth level, “*Summary Clinical Texts*”, ACCESSSS presents the total number of topics found in each source from summaries of clinical texts. This level presents integrated summaries of the first three layers of the pyramid. This means the following: if the best evidence on a particular aspect of the health subject is in a clinical study, in a systematic review or in a *guideline*, it is

probably available in summary form in this fourth level of the pyramid, regardless of the type of material pre-evaluated (BURWELL, 2019).

Thus, primarily, our method consists of using the results found in the fourth level to select the source of information to be consulted for the construction of the indicators. For this reason, the procedures set out in this step in Figure 1 should be applied.

### **3.2 Selection of the evidence-based information source**

The total number of topics resulting from the fourth level of the search performed in ACCESSSSS can be extensive, dispersed in different services and little circumscribed for a particular subject of a health site. According to the terms chosen to perform the search, the *meta-search* can present more than one service of summarized clinical texts, according to the column “Results” of this step in Figure 1. In this case, it will be necessary to select the best service for each case. This methodological step presents methods to make this selection.

The first possibility is to verify studies that have evaluated the services found. The study by Prorok et al. (2012) presented results of the evaluation of services of this type in relation to the punctuality in updating the content, the breadth of coverage and the quality of the evidence offered. The study analyzed ten services and developed a *ranking* according to a scale of 1 to 10 for each of the three criteria observed. On this scale, the number 1 is the best evaluation. The service called *DynaMed Plus* received the best overall evaluation and ranked first in the *ranking* developed by Prorok et al. (2012). The results of the evaluation performed by studies such as Prorok et al. (2012) is the first parameter that can guide the choice of the source of information for the construction of accuracy indicators.

As *DynaMed Plus* stands out among the evidence-based information services found in the fourth level of “Pyramid 5.0” in the search performed in ACCESSSSS, we investigated this source and how it is structured. To visit *DynaMed Plus*, we visited the official website. The service is conducted by professionals linked to universities located in the United States (DYNAMED,

2019b). It has the following partners: *American College of Physicians, American Medical Association, American Association of Nurse Practitioners, Guidelines International Network, McMaster University, Micromedex® Solutions, Wikipedia* and *Zynx Health* (DYNAMED, 2019c). In addition, the website has a page on its conflict of interest policy (DYNAMED, 2019d).

However, within the scope of the MBE, we promote the idea that:

[...] Explicit processes for finding, evaluating, incorporating and updating evidence related to the diagnosis, course, cause, prevention, treatment and rehabilitation of health problems are more important than the academic position occupied. (BURWELL, 2019, p. 61)

In this sense, we performed an evaluation of *DynaMed Plus* in relation to its processes for such purposes, which configures a second method to provide parameters for selecting a source. We used the tool presented in the methodological framework, the “*Guides for judging whether an online clinical text is evidence-based and current*” (BURWELL, 2019). The result was the fulfillment of 11 of the 14 criteria. Thus, we can conclude that the topics of *DynaMed Plus* are based on evidence because they follow a standard that meets almost all the criteria suggested by Burwell (2019).

There is also a third selection method. Another relevant aspect that an evidence-based information service should have is a policy of levels of evidence for its texts and the methods used to establish such levels. In the case of *DynaMed Plus*, an online PDF document is available that describes how the evidence can be classified into three levels: *level 1 (likely reliable) evidence, level 2 (mid-level) evidence* and *level 3 (lacking direct) evidence*.

Therefore, we consider that *DynaMed Plus* is currently the best service and incorporate into our approach. It should be the first option of an evidence-based information source for the construction of accuracy indicators used to evaluate information quality in health sites. Therefore, the other sources can be excluded according to the “Procedures” column of this step in Figure 1.

### 3.3 Collection and selection of evidence-based topics

The search performed in the ACCESSSS provides topics found in the *DynaMed Plus* service (Figure 1). According to the researched subject, few or many topics can be found. If few topics are found and they meet the informational needs for the construction of accuracy indicators on a particular subject of interest, the study may proceed to the collection of evidence-based information. This is the ideal scenario. However, searches can result in a large number of topics that provide information on the desired subject. In this case, it is necessary to select the appropriate topics for the indicators that we intend to develop.

To address a large volume of topics, because some are strictly technical, we developed four exclusion criteria that can guide the circumscription of interesting topics for the generation of evaluation indicators of health sites: (1) specific drugs that appear to be labeled: such information goes into many details that probably will not bring benefits if arranged on websites intended for end users; (2) specific diagnostic test: a topic containing technical information about a test such as the types of measurements used; (3) meta-topic: *DynaMed Plus* presents topics containing lists of *links* to others; and (4) does not provide information: *DynaMed Plus* presents topics on tools such as health calculators.

In our method, to apply these four exclusion criteria, the topics should be accessed, and the first section of each page should be evaluated; in most cases, the “*Overview and Recommendations*” section is common in *DynaMed Plus* topics. This procedure allows excluding topics from the research samples according to the column “Procedures” of this step in Figure 1. With the inclusion of a more circumscribed sample of topics, we can proceed to the fourth methodological step and build the first version of accuracy indicators for the evaluation of the quality of health sites.

### 3.4 Development of the first version of indicators

The topics found are online clinical texts to provide accessible information on the best available scientific evidence. We understand that such information is accessible by summarizing the evidence, making it possible for the user,

whether a health professional or not, to obtain information about the most effective treatment and diagnostic options and more efficient forms of disease prevention. Thus, the user does not need to have advanced knowledge to understand statistical methods such as those commonly found in systematic review studies. However, the objective of these online clinical texts is not to construct indicators to evaluate health sites.

We developed a method to develop accuracy indicators using the evidence-based information available in the *DynaMed Plus* topics. First, it is necessary to understand the structure of the topics and how the information is arranged on this site. This source has the advantage of being standardized. The evidence-based information found on each topic in *DynaMed Plus* is organized into sections. Generally, the first section of the left side menu of a topic page presents the option “*Overview and Recommendations*”, which, when expanded, shows the three subsections “*Background*”, “*Evaluation*” and “*Management*”. This first section contains a summary of the evidence. This is a clear and direct clinical text. The user can have an overview of the different aspects that involve the researched topic. In this sense, each section and subsection contains contextualized information. In general, in the subsection “*Background*” provides the following information: the definition of the topic, epidemiological statistics, risk factors and modes of transmission. When there are specific cases about some information, it is common to find considerations and recommendations on the need to better know the evidence to adequately consider it in each case. In “*Evaluation*”, the summary evidence addresses information related to symptoms and existing forms of diagnosis. In this subsection, considerations about the different symptoms and diagnostic possibilities are also found. The last subsection, “*Management*”, contains information on the best proven treatments, as well as their side effects.

According to this information pattern for each topic, the initial procedure of our method is to analyze the possibility of its contents being used in the composition of accuracy indicators, considering that they will be applied in the evaluation of sites intended for users interested in a particular health topic. In

most cases, there is relevant information in this first section. Depending on the health topic, it is necessary to resort to other sections, such as “*Diagnosis*”, “*Treatment*” and “*Prevention and Screening*”. These provide detailed information about the evidence. It is possible to find statistical data, options for different populations and the studies that were used to support the summarized clinical texts.

With an understanding of the information structure of sections and subsections for the topics, the next step is to develop the indicators. This consists of the following. For each topic included (column “Results” of Figure 1), we first read the initial sections. After reading these sections, drafting an initial version that includes the evidence base and the *DynaMed Plus* reference must be completed.

In addition, another method must be previously applied at this stage, and it is important to classify the indicators into groups. Initially, some groups can be specified. Others may be generated through the classification resulting from the reading and abstraction of the context of the evidence. Each group represents a subject within the health topic being evaluated. For example, in the case of tuberculosis presented above, an indicator can be written as follows: symptoms suggestive of pulmonary tuberculosis are fever, fatigue, weight loss, night sweats, hemoptysis or cough for two or more weeks (DYNAMED, 2019e). This indicator can be classified into a group of indicators called “Symptoms”.

The first version of indicators can be considered a gross result that needs to be reviewed, corrected, and adapted, that is, to go through an improvement process. The purpose is to arrive at a viable set of indicators to be applied in the evaluation of health sites intended for users. In this sense, the next steps comprise the steps in this analysis processes.

### **3.5 Group analysis**

In this first stage of analysis, the first procedure of our method is to verify whether all indicators of the first version (Figure 1) are in at least one of the groups created. If any indicator is not within the scope of a group, we should

reflect and consider a classification. There is no problem if there are indicators classified into two groups simultaneously. After classifying all indicators in at least one group, according to the “Procedures” column of this step in Figure 1, we should obtain an overview of the number of indicators in each group, as well as the possible crossings between them.

This analysis should be considered because the total number of indicators developed in the first version is likely to be high and thus impractical in regard to the evaluation of health sites. Then, the next method is to perform an analysis from the overview. We need mechanisms to deal with a high volume of indicators and reflect on how they can be circumscribed to meet the objectives of website evaluations. With this analysis, it is possible to identify how the indicators are distributed and to record indicators of exclusion or union between them for the following steps. The aim of constructing indicators is not to gather all the available evidence on a given health topic. Thus, the objective is to reduce the total number of indicators to a viable set with essential information to evaluate sites intended for users.

### **3.6 Analysis of topics**

In the same sense of circumscribing the first version of indicators developed, it may be necessary to again analyze the topics included as a sample for the construction of the indicators. To reduce the total number of indicators to a set that has essential information to evaluate health sites for users, exclusion criteria can be established for the topics found in *DynaMed Plus* and selected in the third stage (Figure 1).

The following methods can be used. One idea is to exclude topics on subjects related to small groups of people, exceptional cases or even topics considered of low relevance for the team involved in the development of the indicators. Another possibility is to exclude topics that contain very technical information with details on diagnostic methods, drug dosages, etymology, etc.

With the application of exclusion criteria, according to the “Procedures” column of this step in Figure 1, the total number of indicators tends to be

reduced. If this does not occur, the new number of indicators may still be high and probably difficult to use in the evaluation of health sites intended for users. If this is the case, we recommend that you proceed to the methods of the next step. Regardless of this possibility, the wording of the indicators is that of the first version considered a gross result. These drafts need to be revised because many indicators can be corrected, reformulated, combined or even excluded. Therefore, the next step is to reflect on each of the indicators.

### **3.7 Analysis of indicators**

During the development of the first version of the indicators, there is no concern about drafting the indicators in their final forms. Our concern in this first instance is closer to data collection to obtain the indicators. Many of them can be written almost as translations of the original clinical texts found in *DynaMed Plus*. The intention is that the indicators contain information that can be observed on health sites.

This analysis allows the application of three different methods that trigger the reformulation, union or exclusion of indicators included until this stage (Figure 1). The first procedure is the exclusion of more indicators according to exclusion criteria. The second is the combination of indicators with similar or complementary content, causing the total number of indicators to decrease further. The third procedure also concerns the merging contents, but specifically in the permutation of contents that can be better written when transferring a passage from one indicator to another, without any being excluded. We detail these methods in this step.

In the first procedure, exclusion criteria can be defined as follows: lots of information with technical terms; various considerations on different forms of diagnosis or complex or complementary treatment; and information intended for health system professionals or managers. This type of information probably does not bring benefits if the user does not have technical knowledge and, consequently, making it available on websites intended for users would not have a beneficial effect. In addition, the most appropriate target audience for this type

of information is the health professional. The evidence found in indicators of this type is appropriate for the context of the relationship between health professionals and users. Thus, health professionals are better able to judge the evidence-based information that a particular user needs to know.

The second procedure is merging indicators with similar or complementary contents. The merger can be made in 3 ways. The first is to combine indicators because their information is included in another indicator. The second form is to develop a new indicator wording based on information from two or more indicators. The third does not reduce the total number of indicators because it identifies that the indicators involved are better formulated with the transfer of a piece of evidence from one indicator to the other.

During the analysis of the indicators for exclusion or combination, all should be evaluated regarding their wording. Our intention is to make the writing easy to understand. Thus, the indicators can be used in the evaluation of health sites by both health professionals and users. The final wording of the indicators developed using these methods, as well as the process that led to their development, be it the combination of indicators and/or the reformulation of the wording, should be reported in the studies of this field of evaluation, according to column “Procedures” of this step in Figure 1. The presentation of each of the accuracy indicators developed should also be accompanied by an explanation of how each indicator should be applied in the evaluation of websites.

#### **4 Conclusions**

The main objective of this study was to develop methods for constructing accuracy indicators for the evaluation of health information on the Internet. The set of methods developed consists of seven steps.

To achieve these results, we built a set of methods supported by the foundations of evidence-based medicine (EBM) and guided by improved practices over more than five decades of evolution. If the practice of EBM is a challenge for health professionals, access to evidence-based and legible information to citizens seems to be an even greater challenge. However, access

to health information on the Internet by the general public is a reality to be addressed.

Our proposal is that health sites targeted at citizens be evaluated with accuracy indicators developed through EBM practices. Thus, this study can contribute both to the promotion of EBM among health professionals who produce information on the Internet and to citizen empowerment through access to quality information. This study reiterates the importance of evidence-based indicators and urges us to reflect on implications for the evaluation of the quality of health information available on the Internet.

This study is part of the international debate on how to define the indicators of the accuracy criterion in the evaluation of information on websites on any health subject. Expert consensus does not up-to-date knowledge available in the world. With the various bibliographic bases and methods guided by the EBM, it is possible to reach up-to-date conclusions and stay abreast with the current stage of scientific knowledge on a given health subject.

Therefore, we advocate updating the accuracy criterion. We propose the new definition: *criterion to evaluate the conformity of information with the best and most current scientific evidence available*. This evidence should be obtained from the practices of 'evidence-based medicine'. Updating this concept has implications for the field of evaluation the quality of health information on the Internet.

Over the years, several methods have been developed for the practice of EBM. In this study, our position is configured as that of consumers of evidence. We understand their forms, where and how to obtain them. They act as inputs for the construction of accuracy indicators. EBM guides the best practices to achieve these goals. Acquiring *preappraised evidence* facilitated this process. We structured a methodological path that can be replicated.

Future studies that seek to evaluate the quality of health information on the Internet using the accuracy criterion should acquire the best and current scientific evidence in the area of interest to develop their evaluation indicators.

In this article, with the practices of the MBE, we demonstrate that it is possible to develop indicators in this way.

Future research may develop instruments to measure the quality of studies that evaluate the information found on health websites. To observe the study methods, these instruments may incorporate criteria that verify whether the accuracy indicators were developed according to the practices of EBM. Studies that continue to use expert consensus as well as selected books or articles without explaining and justifying such choices may be considered of low quality.

We hope that another contribution of this study is the inclusion of the topic of evaluation of the quality of health information on the Internet in the research and action agendas of Brazilian health institutions, whether public, private or professional associations.

### **Financing**

Oswaldo Cruz Foundation.

### **References**

ALBARRAK, A. I. et al. The impact of obesity related websites on decision making among students in Saudi Arabia. **Saudi Pharmaceutical Journal**, [s.l.], v. 24, n. 5, p. 605–610, 2016.

ALPER, B. S.; HAYNES, R. B. EBHC pyramid 5.0 for accessing preappraised evidence and guidance. **Evidence-based medicine**, London, v. 21, n. 4, p. 123–5, 1 ago. 2016.

ATALLAH, Á. N. Medicina baseada em evidências [editorial]. **Diagn Tratamento**, v. 23, n. 2, 2018.

BETTING, L. E. et al. Tratamento de Epilepsia: Consenso dos Especialistas Brasileiros. **Arquivos de Neuro-Psiquiatria**, São Paulo, v. 61, n. 4, p. 1045–1070, dez. 2003.

BRASIL. **Ministério da Saúde. Tuberculose**. 2021 Disponível em: <<http://antigo.saude.gov.br/saude-de-a-z/tuberculose>>. Acesso em: 5 ago. 2021.

BURWELL, D. S. Acquiring the evidence: How to find current best evidence and have current best evidence find us. In: STRAUS, S. E. et al. (Eds.).

**Evidence-based medicine: how to practice and teach EBM.** 5. ed. [S.l.]: Elsevier Health Sciences, 2019. p. 57–90.

CHAPA, D. et al. Using preappraised evidence sources to guide practice decisions. **Journal of the American Association of Nurse Practitioners**, Austin, v. 25, n. 5, p. 234–243, 1 maio 2013.

DICENSO, A.; BAYLEY, L.; HAYNES, R. B. Accessing preappraised evidence: fine-tuning the 5S model into a 6S model. **Evidence-based nursing**, London, v. 12, n. 4, p. 99–101, 1 out. 2009.

DYNAMED, (INTERNET). **About DynaMed | Clinical Decision Support Tool.** 2019a. Disponível em: <<https://www.dynamed.com/home/about>>. Acesso em: 7 maio. 2019.

DYNAMED, (INTERNET). **Executive Leadership | Clinical leaders| Clinical Influencers | DynaMed Plus.** 2019b. Disponível em: <<https://www.dynamed.com/home/about/executive-leadership>>. Acesso em: 27 set. 2019.

DYNAMED, (INTERNET). **DynaMed Partnerships | DynaMed.** 2019c. Disponível em: <<https://www.dynamed.com/home/about/partners>>. Acesso em: 27 set. 2019.

DYNAMED, (INTERNET). **Conflict of Interest Policy | DynaMed.** 2019d. Disponível em: <<https://www.dynamed.com/home/editorial/conflict-of-interest-policy>>. Acesso em: 27 set. 2019.

DYNAMED, (INTERNET). **Record No. T116300, Pulmonary Tuberculosis.** 2019e. EYSENBACH, G. et al. Empirical Studies Assessing the Quality of Health Information for Consumers on the Worldwide Web. **Journal of the American Medical Association**, Chicago, v. 287, n. 20, p. 2691–2700, 22 maio 2002.

EYSENBACH, G. The impact of the Internet on cancer outcomes. **Cancer Journal For Clinicians**, [s.l.], v. 53, n. 6, p. 356–371, 2003.

HAYNES, R. B. Of studies, summaries, synopses, and systems: the “4S” evolution of services for finding current best evidence. **Evidence Based Mental Health**, London, v. 4, n. May, p. 37–41, 2001.

HAYNES, R. B. Of studies, syntheses, synopses, summaries, and systems: the 5S evolution of information services for evidence-based healthcare decisions. **Evidence-based medicine**, London, v. 11, n. 6, p. 162–4, 1 dez. 2006.

LAGAN, B. M.; SINCLAIR, M.; KERNOHAN, W. G. What Is the Impact of the Internet on Decision-Making in Pregnancy? A Global Study. **Birth-issues in Perinatal Care**, [s.l.], v. 38, n. 4, p. 336–345, 2011.

MCMMASTER, U. **ACCESSSS | About**. 2019a. Disponível em:

<<https://www.accessss.org/Pages/About>>. Acesso em: 30 maio. 2019.

MORENO, J. M. et al. A quality evaluation methodology for health-related websites based on a 2-tuple fuzzy linguistic approach. **Soft Computing**, [s.l.], v. 14, n. 8, SI, p. 887–897, jun. 2010.

NADANOVSKY, P. Epidemiologia aplicada a clínica: um enfoque científico do uso da informação médica - Medicina Baseada em Evidência. **Estudos em Saúde Coletiva**, IMS/UERJ - Rio de Janeiro - Br, v. 185, p. 1–27, 1999.

PAOLUCCI, R.; PEREIRA NETO, A. Methods for evaluating the quality of information on health websites: Systematic Review (2001-2014). **Latin American Journal of Development**, Curitiba, v. 3, n. 3, p. 994–1056, 14 maio 2021.

PARK, H. et al. Can a health information exchange save healthcare costs? Evidence from a pilot program in South Korea. **International Journal of Medical Informatics**, [s.l.], v. 84, n. 9, p. 658–666, 2015.

PEREIRA NETO, A.; PAOLUCCI, R. Evaluation of the Quality of Health Information on the Internet: An Analysis of Brazillian Initiatives. In: PEREIRA NETO, A.; FLYNN, M. B. (Eds.). **The Internet and Health in Brazil**. [s.l.]: Springer International Publishing, 2019. p. 181–209.

POWELL, J. et al. The Characteristics and Motivations of Online Health Information Seekers: Cross-Sectional Survey and Qualitative Interview Study. **Journal of Medical Internet Research**, Toronto, v. 13, n. 1, p. e20, 23 fev. 2011.

PROROK, J. C. et al. The quality, breadth, and timeliness of content updating vary substantially for 10 online medical texts: An analytic survey. **Journal of Clinical Epidemiology**, Philadelphia, v. 65, n. 12, p. 1289–1295, 2012.

SACKETT, D. L. et al. Evidence-based medicine: what it is and what it isn't. **British Medical Journal (Clinical research ed.)**, London, v. 312, n. 7023, p. 71–2, 13 jan. 1996.

SILVA-JEREZ, N. S.; FURNIVAL, A. C. Informação sobre saúde: letramento informacional e acesso aberto à literatura científica. **Em Questão**, Porto Alegre, v. 24, n. 2, p. 260, 19 abr. 2018.

SMITH, R.; RENNIE, D. Evidence-based medicine: an oral history. **Journal of the American Medical Association**, Chicago, v. 311, n. 4, p. 365, 22 jan. 2014.

SPOELMAN, W. A. et al. Effect of an evidence-based website on healthcare usage: an interrupted time-series study. **British Medical Journal Open**, London, v. 6, n. 11, 2016.

ZHANG, Y.; SUN, Y.; XIE, B. Quality of health information for consumers on the web: A systematic review of indicators, criteria, tools, and evaluation results. **Journal of the Association for Information Science & Technology**, [s.l.], v. 66, n. 10, p. 2071–2084, 26 May 2015

Received: 02/05/2021

Accepted: 06/08/2021

#### **Declaration of authorship**

**Conception and elaboration of the study:** Rodolfo Paolucci Pimenta, André de Faria Pereira Neto, Paulo Nadanovsky

**Data collection:** Rodolfo Paolucci Pimenta

**Analysis and interpretation of data:** Rodolfo Paolucci Pimenta, André de Faria Pereira Neto

**Writing:** Rodolfo Paolucci Pimenta, André de Faria Pereira Neto

**Critical revision of the manuscript:** Rodolfo Paolucci Pimenta

#### **How to cite**

PAOLUCCI, Rodolfo; PEREIRA NETO, André Faria; NADANOVSKY, Paulo. Evaluation of information accuracy in health sites: Methods for the construction of evidence-based indicators. **In Questao**, Porto Alegre, v. 27, n. 4, p 137-188, 2021. DOI: <http://doi.org/10.19132/1808-5245274.137-188>

