

## XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017

### GT-11 – Informação & Saúde

#### RECUPERAÇÃO DA INFORMAÇÃO CIENTÍFICA SOBRE DOENÇAS TROPICAIS NEGLIGENCIADAS: ANÁLISE COMPARATIVA DA SCOPUS, PUBMED E WEB OF SCIENCE

Natanael Vitor Sobral (Universidade Federal da Bahia - UFBA)

Zeny Duarte de Miranda (Universidade Federal da Bahia - UFBA)

#### *RETRIEVAL OF SCIENTIFIC INFORMATION ON NEGLIGENCE TROPICAL DISEASES: comparative analysis of Scopus, PubMed and Web of Science*

#### Modalidade da Apresentação: Comunicação Oral

**Resumo:** Objetiva desenvolver um percurso metodológico de recuperação de informações científicas sobre as Doenças Tropicais Negligenciadas (DTN) nas bases de dados internacionais (Scopus, Web of Science e PubMed). Para isto, utiliza as técnicas desenvolvidas em Sobral (2015), que construiu um modelo para busca de informações científicas a partir do uso dos termos padronizados definidos nos Descritores em Ciências da Saúde da Biblioteca Virtual de Saúde (DeCS/BVS). Os procedimentos metodológicos realizados neste trabalho foram os seguintes: a) identificação das principais doenças tropicais segundo a Organização Mundial da Saúde (OMS); b) conversão dos nomes das doenças em termos padronizados a partir do DeCS/BVS; c) busca de informações nas principais bases de conhecimento anteriormente citadas; d) validação da metodologia através da análise da recuperação dos resultados. Enquanto resultado principal percebeu-se que o uso da busca padronizada aumentou a revocação do processo de recuperação da informação, sendo que a Scopus foi a base que apresentou um aumento mais significativo na recuperação de resultados (11,28%). A PubMed teve um incremento na revocação de 8,92% e a Web of Science de 2,64%. Ademais, propõe-se neste trabalho um fluxo para a recuperação da informação nas referidas bases, que pode ser útil a todos os pesquisadores que desejarem obter informação científica sobre este tema.

**Palavras-Chave:** Medicina Tropical; Bases de Dados; Ciência da Informação; Informação Científica; Recuperação da Informação.

**Abstract:** It aims to develop a methodological pathway for retrieving scientific information on Neglected Tropical Diseases (NTDs) in international databases (Scopus, Web of Science and PubMed). For this, it uses the techniques developed in Sobral (2015), which constructed a model to search for scientific information using the standardized terms defined in the Health Sciences Descriptors of the Virtual Health Library (DeCS). The methodological procedures

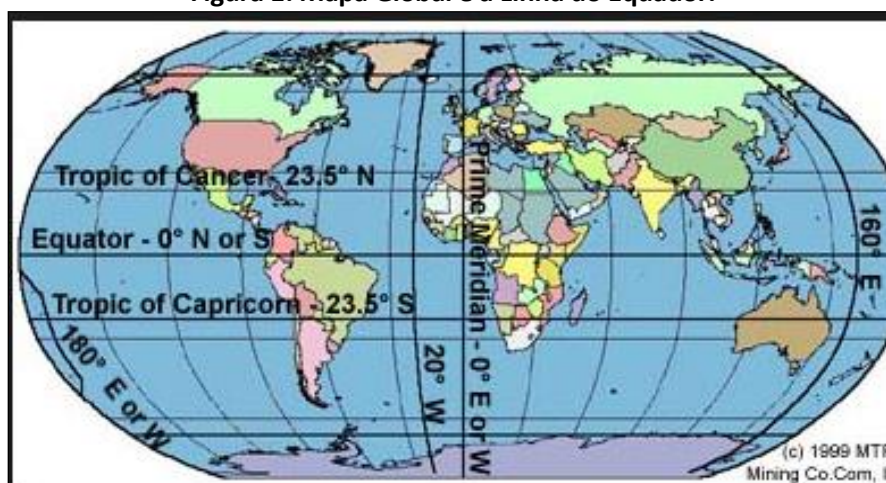
performed in this work were: a) identification of the main tropical diseases according to the World Health Organization (WHO); b) conversion of disease names in standard terms from the DeCS; main knowledge bases previously mentioned d) validation of the methodology through the analysis of the recovery of the results. As a main result, it was noticed that the use of the standardized search increased the recall of the information retrieval process, and Scopus was the basis that presented a more significant increase in the recovery of results (11.28%). PubMed had an increase in recall of 8.92% and Web of Science of 2.64%. In addition, it is proposed in this work a flow for the retrieval of information in these databases, which may be useful to all researchers who wish to obtain scientific information on this subject.

**Keywords:** Tropical Medicine; Databases; Information Science; Scientific Information; Information Retrieval.

## 1 INTRODUÇÃO

Considerando o atual contexto da saúde tropical do Brasil, em relação aos países da Europa, identifica-se que o financiamento e investimento em pesquisas ligadas às doenças tropicais devem ser entendidos como prioridade pelos governantes, tendo em vista os fatores ambientais e sociais que favorecem a disseminação destas doenças, principalmente na população menos abastada. Ao explicar esta questão, Camargo (2008) afirma que as doenças tropicais se referem a doenças infecciosas se proliferam em condições climáticas quentes e úmidas, típicas de países situados entre às proximidades do Equador, entre os trópicos de Câncer e Capricórnio, exatamente onde está situada grande parte do Brasil, o que cria as condições necessárias para a proliferação rápida destas enfermidades, exigindo respostas rápidas e constantes da comunidade científica.

**Figura 1: Mapa Global e a Linha do Equador.**



Fonte: Roebuckclasses Cartography – 1999.

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017  
23 a 27 de outubro de 2017 – Marília – SP**

Compreendendo que nos países desenvolvidos estão alocados os recursos humanos, financeiros e tecnológicos mais avançados da área de Medicina, entende-se que é primordial o desenvolvimento de parcerias entre os países em desenvolvimento com os desenvolvidos, visando avançar nas pesquisas da área de Medicina Tropical, e principalmente, estimular as redes de colaboração científica para a produção de medicamentos e soluções efetivas para estes problemas. De um lado, encontram-se recursos consolidados para lidar com problemas avançados no campo da saúde, do outro, encontram-se os desafios e pesquisadores ávidos por parcerias profícuas que contribuam para a viabilização das soluções criativas pensadas nos países em desenvolvimento para os países em desenvolvimento. Tal cenário pressupõe possibilidades de intercâmbio de conhecimentos e recursos positivos para ambos os lados.

De forma geral, admite-se que o envolvimento dos países desenvolvidos nas soluções dos desafios ligados à Medicina Tropical, comumente, está associado aos seguintes fatores: 1) possibilidade de lucrar sobre os países em desenvolvimento com a venda de vacinas e medicamentos; 2) surtos que afetam os seus territórios esporadicamente, como por exemplo, os recentes casos de Zika em solo estadunidense, que trouxe como consequência problemas congênitos em 5% dos bebês e fetos das mulheres acometidas pela doença (ESTADOS UNIDOS DA AMÉRICA, 2017).

Da indústria farmacêutica, fortemente influenciada pelas forças do mercado, nota-se uma postura inerte no que tange a produção de medicamentos para as doenças tropicais. Segundo Pedrique et al. (2013), dos 850 medicamentos e vacinas aprovados para todas as doenças entre 2000 e 2011, apenas 4% foram para as Doenças Tropicais Negligenciadas (DTN), incluindo a malária, tuberculose, doenças diarreicas e outras enfermidades da pobreza.

Enquanto área interdisciplinar, a Ciência da Informação (CI) estabelece relações com as mais diversas áreas do conhecimento, visando contribuir para a solução de problemas distintos, porém, centrados nas questões informacionais. Deste modo, similarmente às afirmações encontradas em Sobral (2015a), evidentemente, não se pretende em um trabalho da área de CI apresentar soluções para as doenças tropicais, todavia, busca-se compreender questões vinculadas aos processos de busca, monitoramento e análise de informações situadas no campo da Medicina Tropical.

Assim, parte-se do pressuposto que a CI tem como objeto a produção, seleção, organização, interpretação, armazenamento, recuperação, disseminação, transformação e uso da informação (GRIFFITH, 1980), delimitando neste escopo a sua contribuição para os mais

variados temas e áreas do conhecimento, inclusive para si mesma. Enquanto problemática, numa perspectiva diacrônica, verifica-se que a discussão das relações entre o acúmulo de informações e sua posterior seleção e utilização constituem-se como fenômeno de interesse da CI ao longo de sua curta história. No contexto das DTN, tal problemática interfere diretamente na maneira e na velocidade como as informações científicas são geradas, sistematizadas, recuperadas e utilizadas pelos pesquisadores, sobretudo, considerando que em muitos casos, há o envolvimento de cientistas externos à área médica que desejam estudar o tema, porém, possuem pouco conhecimento sobre os nomes das doenças, suas variações terminológicas e as alternativas e critérios para a recuperação de informações científicas sobre as patologias que compõem o escopo das DTN.

Desta feita, o propósito deste estudo é desenvolver estratégias para a recuperação de informações científicas sobre a Medicina Tropical nas principais bases de informações científicas internacionais que indexam conteúdo dessa área. Em suma, esta pesquisa objetiva desenvolver um percurso metodológico de recuperação de informações científicas sobre as DTN nas bases de dados: Scopus, Web of Science e PubMed.

Elegeram-se as bases supracitadas pelo eminente prestígio que estas possuem no meio acadêmico-científico, e por seu vasto acervo de informações científicas da área de Medicina com alta qualidade e impacto, constituindo-se assim, em um conjunto informacional indispensável para a realização de pesquisas que envolvem a área médica. Tal estratégia deverá avançar para outras bases e repositórios de conhecimento científico à medida que o projeto de doutorado a qual está vinculada evolua e realize experimentos em outros contextos.

## **2 A MEDICINA TROPICAL E AS POLÍTICAS PÚBLICAS**

A Medicina Tropical é uma área do conhecimento científico que investe grande parte de seus esforços no estudo das doenças negligenciadas, típicas de países em desenvolvimento, situados nas proximidades da linha do Equador, especificamente localizados em lugares com acentuada situação de pobreza. Tal linha de estudos, em geral, não faz parte da agenda dos países desenvolvidos, detentores de maior poder econômico e científico, devido ao fato destes não se veem acometidos frequentemente por tais enfermidades. Assim, recai sobre os países em desenvolvimento a responsabilidade maior de realização destas pesquisas.

Conforme aponta Dujardin et al. (2010), as regiões da América Latina, Caribe, África e Ásia são as mais afetadas pelas doenças tropicais. Isto ocorre, principalmente, pelas seguintes

causas: enorme diversidade ecológica, mudanças ambientais em curso, migrações massivas, questões socioeconômicas, proximidade da linha do equador (clima quente e úmido), condições sanitárias, e ausência de políticas efetivas orientadas ao enfretamento destas doenças.

Aqui, chama-se a atenção para o fato de que além das DTN há uma ciência tropical negligenciada, o que reforça as dificuldades dos países latino-americanos, africanos e asiáticos. Porém, admite-se que sozinhos esses países encontrarão dificuldades para alcançarem as soluções que necessitam, tendo em vista o seu estágio atual de desenvolvimento em Ciência, Tecnologia & Inovação (CT&I). Para atenuar o problema, a recomendação da OMS (2015) é compor um esforço coletivo que envolva trabalhadores de saúde das comunidades locais e pesquisadores interessados em doenças tropicais de todo o mundo. Sem um esforço global organizado, sistematizado e acompanhado é improvável que sejam alcançadas as soluções necessárias para os problemas de DTN.

Considerando tal realidade, expõe-se que o combate aos problemas de saúde pública e tropical deve ser realizado em âmbito estratégico, visando atacar as causas e mitigar as variáveis que ocasionam as doenças. Dessa maneira, a figura 2 e o quadro 1, baseados em Sobral (2015b) apontam os níveis de atuação das políticas públicas em saúde, concernente a sua efetividade.

**Figura 2: Níveis de atuação das políticas públicas em saúde.**



Fonte: Sobral – 2015b.

Conforme verifica-se na imagem acima, as ações de natureza operacional possuem como característica o curtíssimo prazo, em geral, atuam no âmbito da consequência do problema, e devem servir como paliativo enquanto às ações táticas e estratégicas estão sendo desenvolvidas. As ações táticas empregam um grau de inteligência maior do que a operacional.

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017  
23 a 27 de outubro de 2017 – Marília – SP**

Ainda que não sirva ao propósito de solução, mas sua natureza está intimamente relacionada a processos de geração de conhecimento, utilização de tecnologias, modernização, e técnicas de planejamento e gestão, que permitem avançar em direção ao controle pleno de um determinado problema, e também, mitigar as consequências causadas pelas doenças. Quanto ao nível estratégico, visivelmente sua atuação está ligada às causas do problema, e não aos efeitos. Sua essência é orientada à construção de soluções efetivas e duradouras, que atinjam, precisamente, as condições de produção do problema.

No intuito de ilustrar tais níveis, tornando-os didáticos, Sobral (2015b) desenvolveu um quadro de exemplos de ações para cada nível de política pública, focando em soluções para o problema da Dengue no Brasil.

**Quadro 1: Exemplo de ações para cada nível de atuação das políticas públicas em saúde – combate à Dengue.**

<b>Operacional</b>	<b>Tático</b>	<b>Estratégico</b>
Disseminar o uso de repelentes.	Construir sistemas de informação para monitorar os casos de Dengue.	Investir em C&T para desenvolver uma vacina contra a Dengue.
Utilizar carro de fumaça para afastar os mosquitos.	Reforçar o efetivo dos Hospitais, modernizar e construir novas unidades.	Desenvolver instrumentos de limpeza urbana que inviabilizem o acúmulo de água.
Distribuir o pó que combate o mosquito da Dengue.	Contratar e desenvolver equipes de agentes de saúde para massificar o combate à Dengue.	Conscientizar a população sobre como eliminar o mosquito.

**Fonte: Sobral - 2015b.**

Desta feita, nota-se que a primeira coluna preocupa-se em apresentar ações de urgência para prevenir e tratar os efeitos da proliferação do mosquito causador da doença, realizando atos de natureza simples e imediata. A segunda coluna foca em ações gerenciais e de controle, que objetivam ampliar as condições de prevenção e tratamento, fortemente amparadas pelas técnicas administrativas e ferramentas de tecnologia da informação e comunicação. Enquanto isto, a terceira coluna enfatiza as ações que combatem a causa do problema, tendo em vista que a criação de vacinas, a eliminação das condições de reprodução do mosquito e a conscientização da população, se atingidas plenamente, eliminariam ou reduziriam drasticamente os casos de Dengue no Brasil.

De uma maneira geral, os esquemas apresentados na figura 2 e no quadro 1 podem ser utilizados para a proposição de políticas públicas em Saúde Tropical. Admite-se aqui, que a atuação nos três níveis apresentados é importante para o combate aos problemas de saúde existentes, deste modo, mesmo as ações operacionais são consideradas importantes, pois ajudam a ganhar tempo enquanto as ações mais elaboradas, efetivas e duradouras são desenvolvidas.

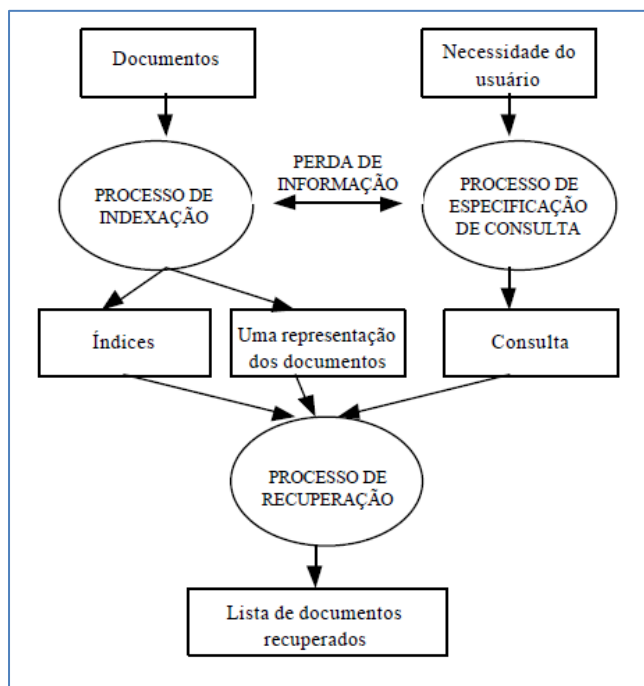
### **3 RECUPERAÇÃO DA INFORMAÇÃO CIENTÍFICA E SUAS RELAÇÕES COM OS ESTUDOS MÉTRICOS DA INFORMAÇÃO**

A estratégia de busca é o processo de traduzir uma pergunta de pesquisa em um formato que o motor de busca consiga entender (GILLESPIE; GILLESPIE, 2003). Tal estratégia está inserida no processo de Recuperação da Informação. Manning et al. (2008) definem a Recuperação da Informação como o processo de encontrar informações, em geral, textos em grandes coleções, satisfazendo necessidades de informação dos usuários. Mooers (1951), autor fundante da Recuperação da Informação, a conceituou como o processo ou método pelo qual um potencial usuário de informações é capaz de converter sua necessidade de informações em uma lista real de informações sobre documentos do seu interesse.

Sob o aspecto prático, Gey (1992) apresentou os itens que compõem um Sistema de Recuperação de Informação (Figura 3), que na interpretação de Cardoso (2004), tais sistemas incluem documentos, necessidades do usuário que desencadeiam a formulação de consultas, e finalmente, a Recuperação da Informação, que depende do alinhamento entre o processo de indexação dos documentos e a busca realizada. Como produto disto, uma lista de documentos considerados relevantes é apresentada ao usuário solicitante. Havendo significativa divergência entre os termos de indexação e a busca dos usuários, as possibilidades de perda de informações são alavancadas, e o processo de Recuperação da Informação tende a ser menos eficaz.

**Figura 3: Componentes de um sistema de recuperação de informação.**





Fonte: Gey - 1992.

Historicamente, os estudos ligados à Organização e Recuperação da Informação estão presentes no centro/núcleo da CI. Na Web of Science, os primeiros trabalhos vinculados ao tema foram publicados por Mooers (1954), que estudou os mecanismos de recuperação da informação e as relações da teoria da comunicação com a teoria da recuperação. Em 1956, o mesmo Mooers já expressava preocupação com o rápido desenvolvimento dos dispositivos de recuperação da informação, afirmando que, antes de um bibliotecário ou administrador de pesquisa entender um processo, vários novos métodos já haviam sido anunciados, panorama similar ao encontrado na atualidade.

No final da década de 1970, Hawkins (1978) desenvolveu trabalho pioneiro relacionando a Recuperação da Informação aos Estudos Métricos da Informação (mais especificamente a bibliometria), porém, tal trabalho teve a finalidade de estudar a literatura sobre Recuperação da Informação Online, e não necessariamente compreender as inter-relações entre os assuntos. A conclusão do autor supracitado foi que a produção científica sobre o referido tema estava dispersa em várias revistas, muitas delas não dedicadas à Biblioteconomia e CI, o que ressaltou explicitamente a importância do tema para outras áreas do conhecimento.

Mais recentemente, estudos importantes relacionaram aspectos dos estudos métricos da informação com a Recuperação da Informação. Leydesdorff e Bornmann (2016) discutiram a categorização de assuntos que a Web of Science oferece às revistas científicas indexadas em sua



**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017  
23 a 27 de outubro de 2017 – Marília – SP**

base. Dentre as inúmeras preocupações, chamou à atenção dos autores a importância que tal classificação possui, haja vista sua utilização pelos principais rankings científicos do mundo. Deste modo, qualquer problema de classificação de assunto, além de prover prejuízos à Recuperação da Informação, também, pode comprometer a qualidade dos rankings científicos que são, sumariamente, de essência bibliométrica e informétrica.

Mayr e Scharnhorst (2015) lembram que a Recuperação da Informação não é apenas restrita aos contextos de comunicação acadêmica e bases de dados bibliográficos. Sua área de aplicação engloba inteligência artificial, informações comerciais, catálogos de bibliotecas, coleções de museus e bibliotecas, e a rede mundial (web) como um todo (motores de busca). Todavia, reforça-se que no escopo dos Estudos Métricos da Informação dentro da CI, em geral, estuda-se a Recuperação da Informação, principalmente, no universo dos artigos de revistas, e apenas recentemente, abriu-se espaço para a comunicação acadêmica na web (webometria e altmetria).

Glänzel (2015), ao observar as relações entre os Estudos Métricos da Informação e a Recuperação da Informação discute a importância das métricas para ajustar as estratégias de busca, e lembra que a recuperação bibliométrica (*bibliometric retrieval*) é uma ferramenta poderosa para desenvolver e ajustar a estratégia de busca em qualquer nível de agregação, mas que sempre haverá ruídos no processo de busca. Em especial, estes ruídos se dão por problemas de indexação, sinonímias, duplicação, incompletude dos metadados, erros de digitação, e obviamente, pode ocorrer, também, pela falta de destreza dos usuários.

Desse modo, nota-se que o processo de Recuperação da Informação não é trivial, diferenciando-se de um simples processo de busca e consulta, e por isso, depende de alguns conhecimentos específicos para que se obtenha um resultado satisfatório. Dentre as diversas competências requisitadas, destacam-se as seguintes:

- Conhecimento da estrutura da base de dados e dos seus requisitos funcionais.
- Compreensão do tema e da variedade de termos que o representa.
- Entendimento da política de indexação e do método utilizado para atribuição de termos.
- Percepção da interface de busca.
- Destreza com os operadores e uso de busca avançada.
- Utilização de filtros de pesquisa.

- Noção das possibilidades de representação gráfica das informações recuperadas.
- Aferição das estatísticas informacionais para fins de validação da busca e definição de novos recortes.

No mais, enquanto alternativa automatizada existe a possibilidade de realizar a busca e monitoramento de informações de maneira constante via API (*Application Programming Interface*). Esta tecnologia, que se utiliza de comandos de query, solicitam informações das bases para extração de dados, registros e informações. Tais monitoramentos podem servir a propósitos diversos, que vão desde o acompanhamento de atualizações sobre uma determinada produção até a obtenção de indicadores em tempo real.

Na Web of Science, por exemplo, é possível acessar informações formatadas e atualizadas para aprimorar o repositório da instituição; consultar automaticamente e em tempo real vários registros, eliminando a necessidade de pesquisas manuais. Por meio desta técnica, pode-se também elaborar estratégias de busca avançadas e solicitar das bases de dados, por meio de rotinas informatizadas, pedidos de informações em que aplicações externas terão acesso eletrônico ao sistema requerente.

#### **4 PROCEDIMENTOS METODOLÓGICOS**

Este estudo, de caráter metodológico, teve por objetivo desenvolver um conjunto de procedimentos para recuperação da informação nas principais bases internacionais de informação científica (Web of Science, Scopus e PubMed). Para isto, utilizaram-se etapas do processo proposto por Sobral (2015a), que incluem:

- a) identificação das principais DTN;
- b) padronização dos termos que representam os nomes das doenças;
- c) busca de informações nas principais bases de informação científica: Scopus, Web of Science e PubMed;
- d) comparação dos resultados obtidos.

Para a compreensão do processo de busca e recuperação da informação realizou-se uma pesquisa bibliográfica, consultando literaturas nacionais e internacionais que discutiam o assunto. Visando estruturar o resultado em um esquema visualmente entendível, aplicou-se a técnica de fluxogramas, fortemente trabalhada no *Business Process Management* (Gerenciamento de Processos de Negócio). Para este propósito, fez-se uso da ferramenta Bizagi®.

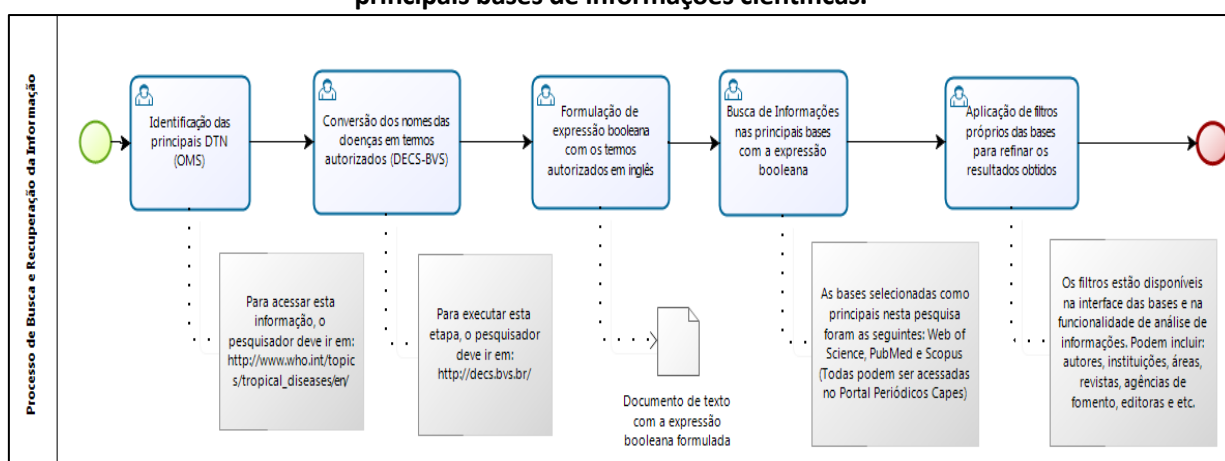
Após a estruturação do fluxo e realização do processo de recuperação foi possível comparar os resultados e verificar o aumento percentual de informações recuperadas em cada base analisada, a partir do parâmetro “busca padronizada com uso de tesouro” versus “busca não padronizada”.

## 5 RESULTADOS E DISCUSSÃO: PESQUISA EM ANDAMENTO

O presente trabalho está inserido no contexto de uma tese de doutorado em andamento, deste modo, os resultados apresentados são preliminares, e futuramente serão utilizados para a Recuperação de Informações sobre DTN, visando aplicações bibliométricas e informétricas, que propiciem panoramas comparativos da produção científica nacional e internacional sobre o tema em questão. A proposta inicial da pesquisa é a prospecção de métodos de Recuperação da Informação em bases de dados e a classificação da informação científica em Medicina Tropical, porém, tal alvo é apenas meio para objetivos maiores.

Buscou-se aqui, desenvolver estratégias eficientes para a recuperação de informações científicas sobre a Medicina Tropical nas principais bases internacionais que indexam conteúdo dessa área. Desse modo, os resultados obtidos pelo estudo estão sumarizados em dois tópicos: 1) sistematização do fluxo metodológico de recuperação de informações (Figura 4); e 2) comparação da busca e recuperação com e sem termos padronizados (Figura 5).

**Figura 4: Processo de busca e recuperação de informações em Medicina Tropical nas principais bases de informações científicas.**



Fonte: Os autores - 2017.

O esquema apresentado na figura 4 evidencia-se como um processo técnico para pesquisadores de qualquer área do conhecimento que desejarem buscar conhecimento

científico sobre o tema de DTN, seja para finalidades bibliométrica/informétricas, ou para a recuperação de informações e documentos. A literatura de Medicina Tropical carece de uma delimitação clara sobre quais são as doenças pertencentes ao rol de DTN. Sobre isto, Camargo (2008) afirma que todas as doenças humanas, a priori, são tropicais, uma vez que a espécie humana se originou nos trópicos, e, com ela, suas doenças, exceto as doenças que a humanidade foi adquirindo, ao longo de sua história de companheiros de jornada como cães, gatos, roedores, aves e mesmo de parentes próximos, os primatas.

A título de exemplo, Silva e Domingues (2011) lembram que é preciso observar as diferentes dimensões (política, econômica e social) para determinar, se de fato, uma doença é de fato tropical. Em linhas gerais, entende-se que os fatores causadores das doenças são diversos, o que dificulta a determinação taxativa da enfermidade enquanto uma “doença tropical”. Assim, nota-se que há classificações diversas para o assunto, sendo um processo em constante atualização, haja vista o surgimento frequente de novas doenças. De todo modo, em concordância com Camargo (2008), entende-se que a utilização do termo “doenças negligenciadas” surge como o mais adequado, pois não singulariza os trópicos como fator causador, e remete fortemente à questão da pobreza, que é um fator-chave para a ausência de prevenção e tratamento dessas doenças.

Desta feita, pensa-se que a sistematização proposta pela OMS é um bom ponto de partida sob o aspecto classificatório das doenças, sobretudo, pela posição egrégia da instituição enquanto fomentadora e incentivadora do conhecimento neste tema. Assim, no processo expresso na figura 4, deu-se preferência ao mapeamento da lista de doenças pertencentes ao rol de DTN a partir da visão da OMS.

Para solidificar a estratégia de busca, optou-se pela utilização do tesouro dos Descritores em Ciências da Saúde da Biblioteca Virtual em Saúde (DeCS-BVS). Frequentemente utilizado em trabalhos científicos, destaca-se aqui o estudo de León et al. (2015), que utilizou este instrumento para padronizar os termos, e realizar buscas estruturadas em algoritmos baseados em linguagem booleana, similarmente ao presente trabalho.

Com isto, o processo de recuperação da informação se torna mais qualificado, pois além de se basear na classificação de DTN da OMS, ainda utiliza um tesouro para descobrir equivalentes aos termos, objetivando ampliar a revocação da consulta. De uma forma geral, os processos de recuperação são pautados apenas nos termos gerais, tais como “medicina tropical”, “saúde tropical” e “doenças tropicais”, conforme pode ser visto em Sobral, Silva e




**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017**  
**23 a 27 de outubro de 2017 – Marília – SP**

Miranda (2017). Esta estratégia, conforme criticam os próprios autores pode se mostrar ineficiente por restringir a quantidade de informações recuperadas. O ideal é que se utilize além dos termos gerais, os nomes das doenças (termos específicos), seus termos equivalentes (sinônimos), e quando aplicável, os agentes causadores das doenças. Tal estratégia assegura que o máximo de registros será recuperado, evitando a perda de informações pela utilização de uma estratégia de busca restritiva.

Ciente que um dos problemas da recuperação da informação em DTN é a perda de informações pelo uso de expressões de buscas não padronizadas, buscou-se comparar os resultados das buscas com e sem o uso de um tesouro em termos percentuais (ver figura 5). Para evitar a recuperação de resultados do tema consultado em outros domínios, para além da Medicina, limitou-se a busca em cada base ao domínio mais específico de vínculo do tema. Exemplo: Na Web of Science foi possível limitar à área de Medicina Tropical; na Scopus, simplesmente à Medicina; e na PubMed, por já ser essencialmente uma base de saúde, optou-se por não realização delimitações.

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017**  
**23 a 27 de outubro de 2017 – Marília – SP**

**Figura 5: Processo de busca e recuperação de informações em Medicina Tropical nas principais bases de informações científicas.**

<b>Web of Science</b>	<p>Tópico: ("Buruli ulcer" OR "Chagas disease" OR "Dengue" OR "Chikungunya" OR "Dracunculiasis" OR "guinea-worm disease" OR "Echinococcosis" OR "Endemic treponematoses" OR "Yaws" OR "Foodborne trematodiasis" OR "Human African trypanosomiasis" OR "sleeping sickness" OR "Leishmaniasis" OR "Leprosy" OR "Hansen disease" OR "Lymphatic filariasis" OR "Onchocerciasis" OR "river blindness" OR "Rabies" OR "Schistosomiasis" OR "Soil-transmitted helminthiasis" OR "Taeniasis" OR "Cysticercosis" OR "Trachoma")                      Refinado por: Categorias do Web of Science: ( TROPICAL MEDICINE )                      Tempo estipulado: Todos os anos. Índices: SCI-EXPANDED, SSCI, A&amp;HCI, CPCI-S, CPCI-SSH, ESCL.</p> <p><b>Resultados: 23.005</b>                      (de Principal Coleção do Web of Science)</p>	<p>Tópico: ("Buruli Ulcer" OR "Mycobacterium ulcerans Infection" OR "Chagas Disease" OR "South American Trypanosomiasis" OR "Dengue" OR "Dengue Fever" OR "Chikungunya virus" OR "Chikungunya" OR "Dracunculiasis" OR "Dracunculosis" OR "Guinea Worm Infection" OR "Echinococcosis" OR "Hydatid Cyst" OR "Hydatidosis" OR "Cysts, Hydatid" OR "Yaws" OR "Frambesia" OR "Trematoda" OR "Flukes" OR "African Trypanosomiasis" OR "African Sleeping Sickness" OR "Nagana" OR "Leishmaniasis" OR "Leprosy" OR "Hansen's Disease" OR "Filarial Elephantiasis" OR "Lymphatic Filariasis" OR "Bancroftian Elephantiasis" OR "Bancroftian Elephantiasis" OR "Lymphatic Filariasis" OR "Onchocerciasis" OR "Rabies" OR "Hydrophobia" OR "Schistosomiasis" OR "Bilharziasis" OR "Helminthiasis" OR "Nematomorpha Infections" OR "Helminth Infestation" OR "Parasitic Worm Infections" OR "Vermination" OR "Taeniasis" OR "Taenia Infections" OR "Cysticercosis" OR "Taenia" OR "Trachoma" OR "Egyptian Ophthalmia")                      Refinado por: Categorias do Web of Science: ( TROPICAL MEDICINE )                      Tempo estipulado: Todos os anos. Índices: SCI-EXPANDED, SSCI, A&amp;HCI, CPCI-S, CPCI-SSH, ESCL.</p> <p><b>Resultados: 23.612</b>                      (de Principal Coleção do Web of Science)  <b>2,64%</b></p>
<b>Scopus</b>	<p>TITLE-ABS-KEY ( "Buruli ulcer" OR "Chagas disease" OR "Dengue" OR "Chikungunya" OR "Dracunculiasis" OR "guinea-worm disease" OR "Echinococcosis" OR "Endemic treponematoses" OR "Yaws" OR "Foodborne trematodiasis" OR "Human African trypanosomiasis" OR "sleeping sickness" OR "Leishmaniasis" OR "Leprosy" OR "Hansen disease" OR "Lymphatic filariasis" OR "Onchocerciasis" OR "river blindness" OR "Rabies" OR "Schistosomiasis" OR "Soil-transmitted helminthiasis" OR "Taeniasis" OR "Cysticercosis" OR "Trachoma" ) AND ( LIMIT-TO ( SUBJAREA, "MEDI" ) )</p> <p><b>145.991 document results</b></p>	<p>TITLE-ABS-KEY ( "Buruli Ulcer" OR "Mycobacterium ulcerans Infection" OR "Chagas Disease" OR "South American Trypanosomiasis" OR "Dengue" OR "Dengue Fever" OR "Chikungunya virus" OR "Chikungunya" OR "Dracunculiasis" OR "Dracunculosis" OR "Guinea Worm Infection" OR "Echinococcosis" OR "Hydatid Cyst" OR "Hydatidosis" OR "Cysts, Hydatid" OR "Yaws" OR "Frambesia" OR "Trematoda" OR "Flukes" OR "African Trypanosomiasis" OR "African Sleeping Sickness" OR "Nagana" OR "Leishmaniasis" OR "Leprosy" OR "Hansen's Disease" OR "Filarial Elephantiasis" OR "Lymphatic Filariasis" OR "Bancroftian Elephantiasis" OR "Bancroftian Elephantiasis" OR "Lymphatic Filariasis" OR "Onchocerciasis" OR "Rabies" OR "Hydrophobia" OR "Schistosomiasis" OR "Bilharziasis" OR "Helminthiasis" OR "Nematomorpha Infections" OR "Helminth Infestation" OR "Parasitic Worm Infections" OR "Parasitic Worm Infestations" OR "Vermination" OR "Taeniasis" OR "Taenia Infections" OR "Cysticercosis" OR "Taenia" OR "Trachoma" OR "Egyptian Ophthalmia" ) AND ( LIMIT-TO (SUBJAREA, "MEDI" ) )</p> <p><b>162.460 document results</b>  <b>11,28%</b></p>
<b>PubMed</b>	<p>"Buruli ulcer" OR "Chagas disease" OR "Dengue" OR "Chikungunya" OR "Dracunculiasis" OR "guinea-worm disease" OR "Echinococcosis" OR "Endemic treponematoses" OR "Yaws" OR "Foodborne trematodiasis" OR "Human African trypanosomiasis" OR "sleeping sickness" OR "Leishmaniasis" OR "Leprosy" OR "Hansen disease" OR "Lymphatic filariasis" OR "Onchocerciasis" OR "river blindness" OR "Rabies" OR "Schistosomiasis" OR "Soil-transmitted helminthiasis" OR "Taeniasis" OR "Cysticercosis" OR "Trachoma"</p> <p><b>94.992 results</b></p>	<p>TITLE-ABS-KEY ( "Buruli Ulcer" OR "Mycobacterium ulcerans Infection" OR "Chagas Disease" OR "South American Trypanosomiasis" OR "Dengue" OR "Dengue Fever" OR "Chikungunya virus" OR "Chikungunya" OR "Dracunculiasis" OR "Dracunculosis" OR "Guinea Worm Infection" OR "Echinococcosis" OR "Hydatid Cyst" OR "Hydatidosis" OR "Cysts, Hydatid" OR "Yaws" OR "Frambesia" OR "Trematoda" OR "Flukes" OR "African Trypanosomiasis" OR "African Sleeping Sickness" OR "Nagana" OR "Leishmaniasis" OR "Leprosy" OR "Hansen's Disease" OR "Filarial Elephantiasis" OR "Lymphatic Filariasis" OR "Bancroftian Elephantiasis" OR "Bancroftian Elephantiasis" OR "Lymphatic Filariasis" OR "Onchocerciasis" OR "Rabies" OR "Hydrophobia" OR "Schistosomiasis" OR "Bilharziasis" OR "Helminthiasis" OR "Nematomorpha Infections" OR "Helminth Infestation" OR "Parasitic Worm Infections" OR "Parasitic Worm Infestations" OR "Vermination" OR "Taeniasis" OR "Taenia Infections" OR "Cysticercosis" OR "Taenia" OR "Trachoma" OR "Egyptian Ophthalmia" )</p> <p><b>103.466 results</b>  <b>8,92%</b></p>

Fonte: Os autores - 2017.

Conforme indica a figura 5, a Scopus foi a base que apresentou um aumento mais significativo na recuperação de resultados com o uso dos termos padronizados (11,28%). Em seguida, PubMed (8,92%) e Web of Science (2,64%). De uma forma geral, com a estratégia proposta foi possível obter um incremento na recuperação da informação. Assim, pressupõe-se que a busca sem o uso de um tesouro gera perda de informações, e com o uso de um tesouro gera ganhos, tendo em vista que o tesouro amplia as possibilidades de revocação a partir da indicação de um conjunto de termos sinônimos que representam um determinado tema, e também, certifica que as palavras-chave utilizadas são legítimas e presentes na literatura.

## 6 CONSIDERAÇÕES FINAIS

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017  
23 a 27 de outubro de 2017 – Marília – SP**

A estratégia proposta se mostrou eficaz para pesquisadores que têm interesse na área de saúde tropical, mas são de outros campos do conhecimento e não conhecem com profundidade os termos utilizados pela área. A dificuldade de obtenção de um conjunto de termos-chave que representassem a produção científica sobre Medicina Tropical motivou essa pesquisa, dada a necessidade futura do projeto, que é, coletar toda essa massa informacional para a realização de estudos métricos envolvendo o tema DTN.

Em suma, Scopus e PubMed apresentaram significativamente as vantagens de utilização de um tesouro para a busca de informações. Não se deve desprezar o resultado alcançado na Web of Science, haja vista que dependendo do contexto de utilização, um documento a mais pode fazer grande diferença para um pesquisador, por isso, o máximo de informações relevantes sobre um determinado tema deve ser sempre o alvo de uma estratégia eficaz.

Em estudos futuros pretende-se realizar uma análise mais qualitativa dos resultados encontrados, visando entender em quais itens ocorre a perda de informações a partir da análise individual de cada palavra-chave utilizada. Outra estratégia futura é a utilização de outros tesouros para além do DeCS/BVS, sugere-se aqui a exploração mais aprofundada do MeSH (*Medical Subject Headings*), vocabulário controlado dos artigos indexados na PubMed, que a propósito, fundamenta o DeCS/BVS.

Por fim, ressalta-se que os resultados foram gerados no início de 2017, e atualmente, estão sendo aperfeiçoados, monitorados e testados em outras bases de dados, no intuito de construção de indicadores ligados à produção científica brasileira e internacional sobre Medicina Tropical, inclusive, com o uso da ferramenta ScriptLattes e de técnicas automáticas de geração e tratamento de palavras-chave que comporão capítulos da tese de doutorado do autor da pesquisa. Certamente, os primeiros achados aqui expostos orientarão o avanço do projeto, estabelecendo um ponto de partida para a obtenção do corpus que será estudado.

## **REFERÊNCIAS**

CAMARGO, E. P. Doenças tropicais. **estudos avançados**, v. 22, n. 64, p. 95-110, 2008. Disponível em: <<http://www.producao.usp.br/handle/BDPI/11791>>. Acesso em: 02 mar. 2017.

CARDOSO, O. N. P. Recuperação de Informação. **INFOCOMP Journal of Computer Science**, v. 2, n. 1, p. 33-38, 2004. Disponível em: <<http://infocomp.dcc.ufla.br/ojsfiles/journals/1/articles/46/submission/proof/46-1-64-1-10-20140917.pdf>>. Acesso em: 02 mar. 2017.



**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017**  
**23 a 27 de outubro de 2017 – Marília – SP**

DUJARDIN J-C. et al. Research Priorities for Neglected Infectious Diseases in Latin America and the Caribbean Region. **PLoS Negl Trop Dis**, v.4, n.10, out. 2010. Disponível em: <<http://dx.doi.org/10.1371/journal.pntd.0000780>>. Acesso em: 02 mar. 2017.

ESTADOS UNIDOS DA AMÉRICA (País). Centers for Diseases Control and Prevention. U.S. Department Of Health & Human Services. **CDC analysis of data from US territories finds serious birth defects in about 1 in 12 fetuses or infants of pregnant women with Zika infection in the first trimester**. 2017. Disponível em: <<https://www.cdc.gov/media/releases/2017/p0608-zika-data-first-trimester.html>>. Acesso em: 19 jun. 2017.

GEY, F. **Models in Information Retrieval**. *Folders of Tutorial Presented at the 19<sup>th</sup> ACM Conference on Research and Development in Information Retrieval*, 1992.

GILLESPIE, L. D.; GILLESPIE, W. J. Finding current evidence: search strategies and common databases. **Clinical orthopaedics and related research**, v. 413, p. 133-145, 2003.

GLÄNZEL, W. Bibliometrics-aided retrieval: where information retrieval meets scientometrics. **Scientometrics**, v. 102, n. 3, p. 2215-2222, 2015.

GRIFFITH, B. C. **Key papers in information science**. New York: Knowledge Industry Publ, 1980.

HAWKINS, D. T. Bibliometrics of the online information retrieval literature. **Online Review**, v. 2, n. 4, p. 345-352, 1978.

LEÓN, B. et al. Análisis de la producción científica publicada entre 2008 y 2013 sobre suicidio en niños, niñas y adolescentes mediante un estudio bibliométrico. **Revista del Hospital Psiquiátrico de La Habana**, v. 12, n. 2, 2015. Disponível em: <<http://www.medigraphic.com/pdfs/revhospsihab/hph-2015/hph152i.pdf>>. Acesso em: 02 mar. 2017.

LEYDESDORFF, L.; BORNMANN, L. The operationalization of “fields” as WoS subject categories (WCs) in evaluative bibliometrics: The cases of “library and information science” and “science & technology studies”. **Journal of the Association for Information Science and Technology**, v. 67, n. 3, p. 707-714, 2016.

MANNING, C. D. et al. **Introduction to information retrieval**, v. 1. 2008. 496 p.

MAYR, P; SCHARNHORST, A. Scientometrics and information retrieval: weak-links revitalized. **Scientometrics**, v. 102, n. 3, p. 2193-2199, 2015.

MOOERS, C. N. Choice and coding in information retrieval systems. **Transactions of the IRE Professional Group on Information Theory**, v. 4, n. 4, p. 112-118, 1954.

MOOERS, C. N. Zatocoding and developments in information retrieval. In: **Aslib proceedings**. MCB UP Ltd, 1956. p. 3-22.

MOOERS, C. N. Zatocoding applied to mechanical organization of knowledge. **Journal of the Association for Information Science and Technology**, v. 2, n. 1, p. 20-32, 1951.

ORGANIZATION, World Health (OMS). **Neglected tropical diseases**. 2016. Disponível em: <[http://www.who.int/neglected\\_diseases/diseases/en/](http://www.who.int/neglected_diseases/diseases/en/)>. Acesso em: 10 ago. 2016.

ORGANIZATION, World Health (OMS). **The Millennium Development Goals Report: 2015**. Genebra: OMS, 2015. 73 p. Disponível em:

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017**  
**23 a 27 de outubro de 2017 – Marília – SP**

<[http://www.un.org/millenniumgoals/2015\\_MDG\\_Report/pdf/MDG%202015%20rev%20\(July%201\).pdf](http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf)>. Acesso em: 01 ago. 2016.

PEDRIQUE, B. et al. The drug and vaccine landscape for neglected diseases (2000–11): a systematic assessment. **The Lancet Global Health**, London, v. 1, n. 6, p. e371-e379, 2013.

SILVA, P. C. V.; DOMINGUES, A. L. C. Aspectos epidemiológicos da esquistossomose hepatoesplênica no Estado de Pernambuco, Brasil. **Epidemiol. Serv. Saúde**, Brasília, v. 20, n. 3, p. 327-336, set. 2011. Disponível em:

<[http://scielo.iec.pa.gov.br/scielo.php?script=sci\\_arttext&pid=S1679-49742011000300007&lng=pt&nrm=iso](http://scielo.iec.pa.gov.br/scielo.php?script=sci_arttext&pid=S1679-49742011000300007&lng=pt&nrm=iso)>. Acesso em: 05 jul. 2017.  
<http://dx.doi.org/10.5123/S1679-49742011000300007>.

SOBRAL, N. V. **Alinhamento da produção científica do programa de pós-graduação em medicina tropical da UFPE às necessidades sociais de saúde tropical em Pernambuco**: análise cientométrica. 2015. 146 f. Dissertação (Mestrado) - Curso de Ciência da Informação, Programa de Pós-graduação em Ciência da Informação, Universidade Federal de Pernambuco, Recife, 2015a. Disponível em:

<<http://repositorio.ufpe.br/handle/123456789/13842>>. Acesso em: 01 nov. 2016.

SOBRAL, N. V. **Responsabilidade Social e Políticas Públicas no campo da Saúde**. Recife: Imip, 2015b. 32 slides, color. Palestra ministrada no IMIP. (A & B).

SOBRAL, N. V.; SILVA, F. M.; MIRANDA, Z. D. Perfil da produção científica em Medicina Tropical na América Latina: análise do termo “Tropical Medicine” na Web of Science. **Em Questão**, v. 23, p. 31-49, 2017. Disponível em:

<<http://www.redalyc.org/pdf/4656/465650499003.pdf>>. Acesso em: 01 abr. 2017.