

A ADEQUAÇÃO DE DESCRITORES NA REPRESENTAÇÃO DE ARTIGOS CIENTÍFICOS: UMA ANÁLISE SOBRE O TEMA “ESTUDOS MÉTRICOS” NA MEDICINA

ADEQUATION OF DESCRIPTORS ON REPRESENTATION OF SCIENTIFIC PAPERS: AN ANALYSIS ABOUT “METRIC STUDIES” ON MEDICINE

Leilah Santiago Bufrem
Walter Moreira
João batista Ernesto de Moraes
Juliana Larazzotto Freitas

Resumo: Considera que os termos descritores de pesquisa nem sempre representam adequadamente um objeto de estudo, temática ou conceito a eles relacionado. Analisa o desempenho de descritores na produção científica de medicina sobre estudos métricos, extraídos da base Scopus em um período de cinco anos. Parte de uma amostra de 30% do *corpus*, definida por números aleatórios para a análise de conteúdo. Corroborar com o pressuposto de que a qualidade da representação das pesquisas abarcada por suas dimensões morfológicas tem o papel de propiciar a melhor compreensão de seu conteúdo no processo de comunicação científica. Verifica os descritores dos campos de busca correspondentes às palavras-chave, aos títulos e aos resumos dos artigos para perceber se representam as temáticas centrais e seus conceitos determinantes conforme critérios definidos a priori. Como resultados, pontua os aspectos problemáticos dos descritores encontrados e estabelece um quadro analítico para visualização de incidências de descritores adequados e inadequados. Compara os dados empíricos obtidos com os valores relativos à significância de recuperação da informação nesses campos, conforme pesquisa de Freitas, Bufrem e Gabriel Júnior (2010) sobre atribuição de pesos a campos de busca em bases de dados. Conclui que a preocupação com a representação adequada de títulos especialmente e de palavras-chave deve ser intensificada a fim de propiciar maior facilidade de recuperação destes estudos. Constata número considerável de representações adequadas, ou seja, 92,68%, principalmente em resumos. Pretende estimular a preocupação com a definição de descritores para títulos e palavras-chave nas ciências médicas e estimular a realização de futuras pesquisas em domínios ou especialidades da medicina relativas à representação temática e conceitual.

Palavras-chave: Representação temática. Representação de artigos. Descritores de busca. Estudos métricos na medicina.

Abstract: This paper considers that the search descriptor terms do not always adequately represent an object of study, theme or concept related to them. It analyzes the performance of descriptors in the scientific production of medicine about metric studies, extracted from the Scopus database over a period of five years. It starts from a sample of 30% of the *corpus* defined by random numbers for the content analysis. The paper corroborates the assumption that the quality of research representation encompassed by their morphological dimensions has the role of providing a better understanding of its contents in the scientific communication process. The study examines the articles descriptors to verify if they represent the central themes and their determinants concepts according to criteria defined previously. As a result, the study scores the problematic aspects of the founded descriptors and compares the empirical data obtained with the values of weights assigned to each search field of a methodology for assigning weights to different search fields. It concludes that the concern for adequate title and keyword representations should be strengthened to provide greater ease recovery of these studies. The research notes considerable number of appropriate

representations, mainly for the abstracts (92.68 %). Finally, it aims to stimulate the concern about defining descriptors for papers at medical sciences and to encourage the achievement of future research in medicine expertise domains considering the thematic and conceptual representation.

Keywords: Thematic representation. Papers representation. Descriptors search. Metric studies in Medicine.

1 INTRODUÇÃO

Considerando que a linguagem documentária é parte constituinte de sistemas de organização e de comunicação da informação, Lara (2004) argumenta que independentemente do universo ao qual se aplique (bases bibliográficas, *sites* na WEB, conteúdos de manuais técnicos), o uso da terminologia teórica e da terminologia concreta contribui à consecução desses objetivos, à medida que fornece princípios para a identificação dos domínios, delimitação de conceitos e termos, estabelecimento de relações entre conceitos apoiadas em definições, além de prover referência concreta aos descritores.

Em relação à distinção entre termos e descritores, quando se faz referência a termos utiliza-se a linguagem da terminologia e quando se refere a descritores, utiliza-se a linguagem da linguística documentária. No que concerne ao seu propósito, segundo Barité (2013), termo pode ser definido como “palavra ou conjunto de palavras que representa um conceito em um âmbito especializado do saber” enquanto que o descritor, para o autor, é “termo normalizado em sua forma e seu alcance semântico, para sua inclusão em tesauro ou em uma lista de descritores”. Barité (2013) explicita que o descritor corresponde formalmente à etiqueta de um conceito, e é a unidade mínima de significado que integra um tesauro ou uma lista de descritores. O descritor pode ser chamado de termo preferencial e é o termo pelo qual se indexará tendo em vista a recuperação da informação. Logo, os termos não preferenciais não podem ser denominados descritores (BARITÉ, 2013). Com os descritores possibilita-se maior uniformidade e consistência na constituição de unidades de informação.

Segundo van der Laan *et al.* (2004, p.337) “se não houver sintonia entre os termos utilizados pelo indexador para representar e os termos utilizados pelo usuário para buscar, as informações que esses termos representam ficarão irremediavelmente perdidas”. Evidenciando a importância da padronização de termos e do uso de vocabulários controlados, Lima e Boccato (2009, p.131) afirmam que as linguagens documentárias, construídas de acordo com “princípios e métodos estabelecidos pela terminologia, propiciam a compatibilidade entre a linguagem do usuário e a utilizada pelo sistema de recuperação da informação”. Entretanto, o que se questiona aqui não são as divergências entre as unidades de indexação e os modos de busca estabelecidos pelos usuários, e sim, a que aspectos remetem as

representações atribuídas pelos autores produtores de conhecimento e/ou fontes responsáveis pela representação de títulos, palavras-chave e/ou descritores e resumos em suas descrições.

Parte-se da premissa de que os descritores nem sempre apresentam desempenho terminológico adequado como representantes de um objeto de estudo ou conceito a eles relacionado e que a certeza de se obter o core de documentos sobre a temática que se busca não é tão exata como se espera, considerando que os autores e/ou periódicos responsáveis pela representação da informação nem sempre atribuem palavras-chave e/ou descritores que possibilitem a recuperação de seus textos pela temática central, assim como não padronizam termos utilizados em seus resumos e títulos a fim de promover uma indexação e recuperação de informação adequada.

Confirma-se a questão de que as falhas na recuperação de artigos científicos, muitas vezes imperceptíveis ao usuário devido a essas considerações, não podem ser atribuídas exclusivamente ao mecanismo de busca e à consideração da relevância dos campos de busca, mas também, aos modos de representação utilizados pelos autores, revistas ou fontes responsáveis pelo processo.

Nesse sentido, retoma-se Guimarães e Sales (2010), quando afirmam que a área de organização e representação da informação é responsável pela mediação entre os contextos de produção e uso da informação, em especial naquilo que tange à dimensão dos conteúdos, “no mais das vezes denominada como Tratamento Temático da Informação”. Segundo os autores, essa área desdobra-se em três vertentes teóricas, a catalogação de assunto (*subject cataloguing*) de matriz norte-americana, a indexação (*indexing*) de matriz inglesa e a análise documental (*analyse documentaire*) de matriz francesa. (GUIMARÃES, 2008, 2009 *apud* GUIMARÃES; SALES, 2010).

A área de tratamento temático da informação, em qualquer uma de suas vertentes, coincide com a posição de Barité (1997, p.124), ao centrar-se nas questões relativas

[...] à análise, descrição e representação do conteúdo dos documentos, bem como suas inevitáveis interfaces com as teorias e sistemas de armazenamento e recuperação da informação em cujo âmbito desenvolvem-se processos, valendo-se de instrumentos para a geração de produtos.

Esse processo ocorre em todas as áreas do conhecimento científico, uma vez que se refere a um conjunto de operações cuja lógica interna permite a execução de etapas com a finalidade de representar o conteúdo de um documento de forma distinta da original, de modo a subsidiar o processo de recuperação da informação.

No domínio das ciências da saúde, Brandau, Monteiro e Braile (2005), autores de trabalho publicado em periódico sobre cirurgia cardiovascular, chamam a atenção dos

profissionais da área para a importância do uso de descritores a fim de auxiliá-los a fazer a busca corretamente, visto que as publicações científicas mais bem avaliadas costumam especificar a necessidade da indicação de descritores (ou palavras-chave) e em que base de dados devem ser consultados.

Esta problemática evidencia a necessidade de se analisarem os termos de representação na produção científica da medicina para verificar se sua eleição é coerente com o tema central dos artigos e/ou objetos de estudo, assim como, se há coerência entre título, palavras-chave ou descritores e resumo de um mesmo artigo.

É importante clarificar que a ideia de palavra-chave já remete a um vocabulário não controlado, mais próximo da linguagem natural, normalmente escolhida livremente pelo autor do trabalho, ainda que possa dispor de instrumentos de controle vocabular. Mas, optou-se por deixar palavras-chave e descritores como elementos de análise deste estudo porque algumas revistas permitem e outras não a liberdade de atribuição de palavras-chave, do que decorre não serem sempre termos considerados descritores.

Questiona-se se as considerações desses autores/produtores transcendem seus objetos de estudo e os conceitos relacionados a esses objetos. Como propósito, pretende-se verificar o desempenho terminológico dos termos descritores correspondentes às palavras-chave, aos títulos e aos resumos de um *corpus* de artigos da Medicina sobre o tema estudos métricos, para perceber se representam os temas centrais e os conceitos determinantes dos textos.

Pondera-se que os descritores devem se ater à descrição do objeto de estudo e dos conceitos relativos a este objeto. Entretanto, em muitos casos, destacam-se aspectos metodológicos da pesquisa, ou termos específicos que não representam o tema de estudo nas palavras-chave. Em outros, utiliza-se de uma linguagem não padronizada também evidente nos campos título e resumo dos artigos. Os autores, em determinadas ocasiões, utilizam como palavras-chave as áreas do conhecimento em que são realizadas as pesquisas, ou até mesmo adotam atributos e especificidades do objeto como componentes deste campo de descrição, não remetendo de modo adequado à temática principal dos artigos e também há situações em que os artigos não são delimitados temporal e geograficamente quando necessário para a clareza da representação.

Representando aqui uma das áreas de estudo da CI, a linguística documentária, para organização e representação da informação, consideram-se as questões que se relacionam com a metalinguagem linguística. Nesse caso, a linguística documentária não releva aspectos do nível semiótico, segundo Hjelmslev (*apud* BADIR, 2005), mas sim, revela a estrutura e forma

desta linguagem. Logo, consideram-se as linguagens documentárias como mediadoras da informação, ou seja, linguagens de representação da informação.

Pondera-se a possibilidade de promover tal mediação com contribuições da linguística, em moldes pragmáticos ou não, conforme as finalidades de cada domínio. A linguagem documentária apresenta um plano de expressão e um plano de conteúdo (HJELMSLEV, 1975 *apud* LARA, 1993, p.74), do mesmo modo que a linguagem natural. Esses planos podem ser considerados como planos dependentes entre si. Logo, pode-se dizer que uma linguagem documentária é estruturada segundo os campos lexicais e semânticos de uma área de conhecimento, entretanto, a ausência de definições que remetam a contextos determinados, ou o uso indevido de suas expressões, comprometem a sua função comunicativa. (LIMA, 2007, p.125). Destaca-se como incumbência da linguística documentária:

[...] compor os quadros de referência para a análise, avaliação e construção da linguagem documentária, entendida como linguagem de informação, associando os níveis sintático-semântico-pragmático para identificar com clareza a inserção do signo documentário no plano sistêmico e no plano funcional, objetivando-o no tempo, no espaço e na cultura. (TÁLAMO; LARA, 2006, p.206).

O estruturalismo linguístico de Gardin tem permeado os modos de representação de texto na CI, conforme Kobashi (2007), para quem a informação é indexada por palavras (justapostas, relacionadas graficamente em mapas estáticos ou dinâmicos), utilizadas para busca, ou seja, para indexar a pergunta do usuário.

Considerando o fato de que o conhecimento e suas representações se expressam pela linguagem, Kobashi (2007) argumenta que a criação de linguagens para operar em contextos de produção e de busca de informação é parte constitutiva da preocupação com a funcionalidade dos sistemas de informação. A análise e a construção das linguagens comportam inúmeras abordagens, segundo as perspectivas políticas, ideológicas teóricas e metodológicas adotadas. (KOBASHI, 2007). Por considerarem supérfluas ou pouco úteis as operações de tratamento da informação e as linguagens de organização da informação, muitos repositórios não recorrem às linguagens documentárias, para filtrar informação, além do que, segundo a autora, a informação participa de diferentes estruturas de significação, o que motiva a reflexão permanente sobre os métodos de elaborar linguagens apropriadas para os diferentes contextos e seus públicos. Lara (1993, p.77), em relação ao rigor metodológico na elaboração de linguagens documentárias, afirma que este impõe, necessariamente, a consideração do contexto no qual as palavras se inserem.

O fato é que sem as linguagens documentárias, o aprimoramento e rigor metodológico demandado em suas práticas para que estas contemplem plenamente os contextos em foco não

pode haver comunicação e fluxo de informações documentárias, pois o acesso à informação depende destas linguagens para ensejar a intercomunicação entre sistema e usuário. “Desse modo, qualquer que seja a perspectiva teórica adotada, o porquê, o para quê e o para quem se organiza informação determinam sua construção”. (KOBASHI, 2007).

O objeto deste estudo, portanto, define-se a partir dessas considerações, pelas quais se pretende estimular a preocupação com a definição de descritores na medicina, como a evidenciada no estudo de Brandau, Monteiro e Braile (2005). Segundo os autores, os pesquisadores devem atentar para a definição dos descritores ou palavras-chave na submissão de um trabalho para publicação, visto que, tais termos são de grande valor para a indexação. “Muitos pesquisadores da área de saúde, apenas para delimitar um campo da ciência, os utilizam na busca de informações para pesquisar sobre doenças, técnicas cirúrgicas ou mesmo escrever um trabalho”. E caso esses termos não estejam de acordo com a nomenclatura das bases de dados, os artigos procurados podem não ser encontrados e, muitas vezes, nem citados por conta disso. (BRANDAU *et al.*, 2005).

Como objetivo geral do estudo, verifica-se a adequação dos termos de representação da produção científica periódica da medicina sobre estudos métricos quanto à sua coerência na representação de informação científica. Para tanto, coleta-se um *corpus* de artigos da área e por meio de um quadro ilustrativo representam-se os usos inadequados dos seus termos de representação. Por fim, caracterizam-se os modos de representação do domínio em questão em relação aos elementos títulos, palavras-chave e resumo.

Em relação ao domínio selecionado, parte-se de Hjørland e Albrechtsen (1995, p.400) para defini-lo como uma comunidade discursiva, considerando-o como uma nova interface para investigação na área de CI, especificamente para a organização e representação do conhecimento. (ARBOIT; GUIMARÃES, 2013).

A análise de domínio possibilita a delimitação de contextos no âmbito da análise documentária e do tratamento temático da informação, de modo a permitir a representação de comunidades discursivas e o estabelecimento de relações conceituais do modo mais fiel possível da realidade temática focada.

Relacionando-se o campo da linguística documentária com a análise de domínio, revela-se que estudos realizados na área podem contribuir para melhorar as práticas de descrição e representação da informação científica, integrar métodos assim como facilitar a recuperação da informação de artigos científicos.

2 METODOLOGIA

Este estudo descritivo, que utilizou a análise de conteúdo como modo de organizar e categorizar a produção científica apoia-se em um *corpus* de artigos extraídos da base indexadora Scopus na área de Medicina em um período de cinco anos (2008-2012).

Esta base de dados indexadora foi lançada em 2004 e, hoje, é considerada a maior base de dados multidisciplinar de resumos, citações e textos completos da literatura científica mundial. (GRÁCIO; OLIVEIRA, 2012, p.6). A base cobre 27 áreas do conhecimento e indexa mais de 21.000 títulos, dos quais 296 são brasileiros, sendo que 83 correspondem à área de Medicina e três da Ciência da Informação. (SCOPUS, 2013).

O *corpus* de artigos foi recuperado por meio de uma estratégia de busca avançada limitando-se à área de Medicina. Buscou-se pelos campos título, palavra-chave e resumo pelos termos: bibliometr, scientometr, informetr, infometr, webometr, patentometr, “scientific collaboration”, co-authorship, “citation analysis”, “co-citat”, "impact factor", "h index", "Bradford's law", "Zipf's law", obsolescence e "scientific policy".

A estratégia de busca especificou a tipologia documental “artigos científicos” para a área eleita e também se estabeleceu que somente fossem recuperados artigos que tivessem ao menos um autor brasileiro para delimitar ainda mais o *corpus*. O *corpus* total foi extraído de 30 revistas diferentes, das quais 11 são brasileiras.

O *corpus* consistiu em 150 artigos sobre o tema estudos métricos, do qual se analisou uma amostra de 30%, selecionada por meio de números aleatórios. A análise foi realizada nos campos de busca título, palavra-chave e resumo e se fundamentou em questionamentos como: o tema central fica evidente no título? As palavras-chave descrevem o objeto de estudo e os conceitos principais trabalhados? O resumo apresenta elementos essenciais conforme a norma de resumo?

Nos critérios para classificação dos artigos quanto ao resumo, foram consideradas a objetividade e clareza na sua redação; sua completude, isto é, se apresentou aspectos introdutórios, objetivo, metodologia, apontamento de resultados e considerações finais, independentemente do modelo adotado pelo periódico que o publicou. Quanto ao modelo, há resumos informativos com as seções especificadas que são comumente adotados pelos periódicos das áreas de ciências da saúde e de ciências exatas e, também, resumos sem as seções especificadas, apesar de apresentarem o conteúdo relativo a cada seção.

Para a análise das palavras-chave, verificou-se se remetiam ao tema e ao objeto central do artigo, à delimitação geográfica e temporal do tema quando se fizesse necessário e

também, se estavam de acordo com o título e o resumo. Em alguns casos foi consultado o DeCS - Descritores de Ciências da Saúde para confirmar a validade do descritor. Os periódicos, em sua maioria, quando não exigem, ao menos sugerem que os autores consultem algum instrumento de controle vocabular.

Neste processo de classificação, realizou-se a leitura dos textos completos, a partir da qual se verificou se havia coerência entre os elementos título e palavras-chave de um mesmo artigo por meio do estabelecimento de relações e comparações entre estes elementos e o objeto de estudo e objetivos do mesmo artigo, que se concentravam geralmente na sua introdução. A introdução dos artigos analisados foi relacionada aos campos de busca já referidos, observando assim, se estes representavam o tema central do artigo. Em relação à análise da coerência dos resumos, observou-se se estes mencionavam os aspectos essenciais de cada seção principal dos artigos: introdução, metodologia, análise, resultados e considerações finais.

Em vista da necessidade da representação para recuperação de artigos, já que, se a representação destes não for coerente com os seus conteúdos, os usuários que realizam buscas podem ser prejudicados em sua revocação temática, buscou-se, por meio da amostra representativa, verificar se os artigos evidenciavam seu tema central e seus conceitos principais com base no cotejamento descrito acima. Os estudos caracterizados como exploratórios, cujos títulos pareciam abrangentes e os resumos sintetizados demasiadamente, dificultavam a compreensão do tema e dos conceitos relacionados a ele, fato que exigia uma segunda leitura destes textos completos.

Vale dizer que cinco artigos recuperados aleatoriamente foram descartados após a leitura do resumo, já que não se enquadravam na temática. O primeiro artigo utilizou o termo descritor “h index” referindo-se à expressão PLM/h *index*, utilizado para investigar o efeito da L-dopa no índice de PLM /h de medula espinhal de sujeitos feridos. O segundo artigo foi recuperado pela expressão h index de *wealth index*, tratando da desnutrição em crianças. O terceiro foi eliminado por conter o termo descritor “fator de impacto” relacionado à temática poluição. O quarto artigo também foi recuperado pelo termo “fatores de impacto”, porém relacionado à fertilidade da mulher. Por último, o termo recuperado no quinto artigo foi antropometria, sobre a avaliação do estado nutricional das crianças índias do Alto Xingu, não se adequando ao tema “estudos métricos”. Estes artigos foram substituídos por outros do *corpus*, por meio de números aleatórios para seleção.

3 ANÁLISE E RESULTADOS

Na análise não se identificou a presença de aspectos metodológicos que evidenciassem as trajetórias das pesquisas nas suas palavras-chave, como por exemplo, um método utilizado na pesquisa e que também foi descrito como palavra-chave. Entretanto, nas palavras-chave de alguns artigos destacaram-se as seguintes tipologias documentais: artigo de periódico, artigo científico, artigo de pesquisa, publicação científica, artigo de revista. Estes últimos termos foram localizados no DeCS como termos preferenciais ou descritores.

Já, o título foi considerado aqui como campo de maior significância, por concentrar o maior peso na recuperação da informação, conforme proposta de atribuição de peso a campos de busca de uma base de dados (FREITAS *et al.*, 2010). Com a análise dos títulos comprovou-se se abrangiam a amplitude do tema tratado no artigo e se apresentavam termos padronizados em relação aos outros campos de busca.

Os quadros 1A e 1B ilustram a adequação ou não dos campos de busca ao tema central do artigo, conforme critérios estabelecidos na metodologia. Quanto aos quadros, é importante ressaltar que o 1A exemplifica como foi operacionalizada a análise e o 1B traz resultados expressivos de incidências das adequações dos campos para cada artigo. Considera-se que os artigos que não tiveram seus resumos e/ou palavras-chave disponibilizados no documento original não foram contabilizados na porcentagem relativa à totalidade de incidência de cada campo no 1A. Logo, 92,68% dos resumos foram considerados adequados, 86,67% dos títulos e 83,78% das palavras-chave adequaram-se. Essas informações levam a crer que de modo geral, os pesquisadores da área da saúde apresentam cautela na atribuição de seus descritores ou termos, em grande parte, decorrente da maturidade e reconhecimento científico desta ciência e de sua evolução em relação a aspectos estudados no domínio da CI. Corroborar-se tal afirmação quando mencionadas as demandas de bases indexadoras e dos próprios periódicos científicos para publicação, hoje, mais rigorosas quanto à boa qualidade das publicações, com vistas à indexação e internacionalização da ciência. (BRAILE *et al.*, 2007; GOMES, 2010; REGO, 2014).

Embora a comunidade científica e os profissionais da área possam usufruir de um instrumento facilitador das atividades de representação e de recuperação de pesquisas por meio do vocabulário estruturado e trilingue DeCS, que serve como uma linguagem única de indexação, não se deve ignorar que o rigor no sistema nacional de avaliação das publicações científicas estimula autores e editores na busca pelo reconhecimento e visibilidade de suas pesquisas. Para conquistar uma publicação reconhecida no meio é fundamental a boa

qualidade da estruturação das pesquisas. Um estudo recente, que constituiu a amostra analisada, tratou justamente do aspecto estrutural das pesquisas, sobre o controle de qualidade na estruturação de resumos de artigos não experimentais da área de cirurgia. (GUIMARÃES *et al.*, 2013). O estudo é de autoria de professores da área de cirurgia em colaboração com um profissional da informação. Este fato retrata a importância da Ciência da Informação como ciência normativa para o desenvolvimento e crescimento da produção científica de diferentes áreas. No estudo mencionado, a qualidade dos resumos foi mensurada com base em oito critérios de avaliação que foram divididos em 32 categorias. Segundo Guimarães *et al.* (2013), a qualidade de artigos de seis revistas não experimentais brasileiras de cirurgia foi classificada como boa.

Em relação à análise aqui realizada, vale dizer que as inadequações encontradas na verificação do desempenho dos descritores de busca, conforme critérios definidos a priori, impossibilitou a marcação do campo título, resumo e/ou palavra-chave como adequado nos artigos referidos abaixo. Tais inadequações serão explicitadas a seguir.

QUADRO 1A – Adequação ou não dos descritores para cada campo de busca conforme a temática central do artigo

Adequação ao tema central			
Artigo	Título	Resumo	Palavra-chave
1	X	X	X
4	X	X	X
7	X	X	X
9	X	X	X
10	X	X	X
14	X	X	X
16	X	X	X
18	X	X	X
21	X	X	X
22	X	X	X
23	X	X	X
26	X	X	X
27	X	X	X
29	X	X	X
30	X	X	X
32	X	X	X
33	X	X	X
34	X	X	X
35	X	X	X
38	X	X	X
40	X	X	X

41	X	X	X
42	X	X	X
44	X	X	X
2	X	X	
3	X	X	
12	X	X	...
13	X	X	...
19	X	X	
20	X	X	
24	X	X	...
31	X	X	...
36	X	X	
37	X	X	...
5	X		X
25	X	...	X
39	X	...	X
11	X
28	X
6		X	X
17		X	X
43		X	X
15		X	
8			X
45			...
Total de incidências	45	41	37
Artigos adequados	39	38	31
Adequados em (%)	86,67%	92,68%	83,78%

Legenda: ... – Descritor ausente; X – Descritor adequado; célula sem marcação – Descritor não adequado.

Fonte: Dados da pesquisa, 2014.

QUADRO 1B – Incidência e (%) de representações adequadas para os campos de busca dos artigos X Valores de cada campo e suas combinações para a recuperação da informação⁶

Combinação de campos	Incidências adequadas	(%) Incidências Adequadas	Peso da soma dos campos de busca⁷
três campos adequados	24	53,30%	7 pontos
título e resumo adequados	10	22,20%	5 pontos
título e palavras-chave adequadas	3	6,60%	6 pontos
título adequado	2	6,60%	4 pontos
resumo e palavras-chave adequadas	3	4,40%	3 pontos
resumo adequado	1	2,20%	1 ponto
palavras-chave adequadas	1	2,20%	2 pontos
nenhum campo adequado	0	0%	0 pontos

Fonte: Dados da pesquisa, 2014.

Foi possível comparar os dados empíricos obtidos na contagem de incidências dos campos adequados e suas respectivas combinações para um mesmo artigo, com os valores relativos à significância de recuperação da informação nesses campos. Esses valores, conforme Freitas *et al.* (2010) são oriundos de uma representação binária de três bits, sendo o primeiro menos significativo (com peso um) atribuído ao resumo, o segundo bit (com peso dois) atribuído às palavras-chave e o terceiro mais significativo (com peso quatro) atribuído ao título. A representação binária possibilita a variação de oito números, de zero a sete.

Conclui-se que a preocupação com a representação adequada de títulos em especial e de palavras-chave deve ser intensificada a fim de propiciar maior facilidade de recuperação da informação destas pesquisas.

Em relação aos resumos, campos de busca mais bem representados, conforme a análise, observou-se que 42% deles apresentaram a especificação das seções que o compunham, isto é, introdução, objetivos, metodologia, resultados e conclusões explicitados como subdivisão; enquanto que 49% não apresentaram estas especificações. Vale dizer que essa exigência é determinada pelos periódicos. Dos artigos que não apresentaram resumo (9%), dois destes tampouco apresentaram palavras-chave.

⁶ Os valores utilizados na última coluna do quadro 1B foram definidos em monografia de graduação que propôs um modelo de atribuição de pesos para campos de busca em base de dados. O estudo gerou o artigo *Proposta de metodologia para a recuperação da produção científica em Ciência da Informação na base Brapci*. Disponível em: <http://www.portalseer.ufba.br/index.php/revistaici/article/viewFile/4629/3564>;

⁷ Conforme modelo de atribuição de pesos para campos de busca. (FREITAS *et al.*, 2010).

Sobre as palavras-chave, 82,2% dos artigos da amostra as apresentaram. O restante, representado por oito artigos, não as especificou no documento original, embora os periódicos em que foram publicados previssem em “Norma para autores” a necessidade de *keyterms* ou *keywords* para o processo de submissão.

Para esta análise ponderou-se o fato de que a temática central das pesquisas e seus conceitos principais variam conforme os objetivos dos estudos.

Dentre os títulos considerados inadequados pontuam-se as seguintes questões observadas: a ausência da delimitação temática do artigo e da subárea sobre a qual tratava um dos artigos; em outro caso, evidenciou-se o uso inadequado do verbo avaliar, pois o estudo não se constituiu em avaliação e sim em uma análise, já que um estudo avaliativo apresenta critérios ou parâmetros a partir de princípios de avaliação; um terceiro caso utilizou uma expressão equivocada para designar um conceito distinto no título. A expressão usada foi “Desenho dos estudos” para referir-se à qualidade metodológica dos artigos, temática central do artigo original. Em outro caso, o título “Saúde do trabalhador no Brasil: pesquisa na pós-graduação” pareceu não retratar o propósito do artigo. Na verdade, os autores realizaram um estudo específico na produção científica da área de medicina sobre o tema saúde do trabalhador no Brasil. A última incidência para o título se referiu à ausência de área temática para a qual se voltou à pesquisa, considerando que foi realizada uma proposta de índice cientométrico para determinada subárea da medicina, utilizando a análise da produção científica desta mesma área.

Embora os resumos tenham tido melhor desempenho em sua estruturação e representação do que os títulos e palavras-chave dos artigos, conforme se pode verificar no Quadro 1A e 1B, sendo 92,68% deles classificados como adequados em detrimento dos títulos adequados (86,67%) e palavras-chave (83,78%) encontraram-se alguns casos problemáticos nos resumos, como a utilização de qualificações, juízos de valor que não condizem com a linguagem científica, na qual a demanda por informações embasadas em quantificações e dados empíricos confere credibilidade às pesquisas. Foram utilizados termos como “completamente inadequado”, “grandes distorções” em resumo sintético que não pôde ser caracterizado como informativo por não apresentar aspectos metodológicos e resultados. Um segundo resumo somente apresentou o objetivo e uma breve introdução, não mencionando, novamente, aspectos metodológicos e resultados finais.

Sobre as palavras-chave, destacam-se casos em que uma das palavras é “países em desenvolvimento”, posto que o artigo tratou especificamente de sistema nacional de avaliação de periódicos científicos brasileiros, além de que o seu título especificou que foi uma análise

realizada no Brasil. Verifica-se outro caso em que a palavra “linguagem” é utilizada como descritor, porém o estudo tratou especificamente do idioma inglês. Questões relativas à hierarquia de termos parecem ser ignoradas para a atribuição de descritores, embora sejam de grande importância para a realização de uma representação temática coerente com a pesquisa. Houve um caso em que foi atribuída palavra-chave que não se relacionava ao tema investigado.

Resgatando os aspectos da hierarquia e especificidade de termos, surgiu mais uma situação em que o tema central do artigo não figurou especificamente como palavra-chave, o tema era índice H. Houve apenas uma situação em que o número de descritores foi excessivo, chegando em 13 palavras-chave. Nesse contexto é provável que o periódico que publicou o artigo não tenha regras específicas para o controle da quantidade de palavras-chave ou não as aplique rigorosamente.

Deve-se enfatizar que os periódicos, nos quais os artigos analisados do *corpus* foram publicados, apresentam uma credibilidade significativa, visto que estão contemplados na base de dados indexadora Scopus e os brasileiros indexados apresentam estrato Qualis entre A1 e B1, em ao menos uma área do conhecimento. Cumpre ressaltar que futuras análises relativas à representação de descritores em qualquer área do conhecimento são de grande importância para o direcionamento dos modos de se fazer pesquisa e de representá-las adequadamente conforme seu conteúdo e tema central.

4 CONSIDERAÇÕES FINAIS

Procurou-se com este estudo possibilitar um panorama das representações temáticas utilizadas nos principais campos de busca de artigos da medicina em relação aos estudos métricos. A pesquisa empírica em uma amostra aleatória da realidade concreta de um *corpus* de artigos permitiu o reconhecimento, ainda que superficial, das opções recorrentes nesta área do conhecimento para representação temática, conforme explicitado nas análises e resultados.

Além disso, foram elencados aspectos problemáticos que permeiam o processo de eleição de termos por autores ou fontes responsáveis por sua representação e indexação como periódicos e/ou bases de dados. O desempenho dos descritores utilizados nos campos de busca título, resumo e palavra-chave foi mensurado por meio da quantificação das inadequações encontradas em cada artigo para cada campo.

Embora, partir da análise realizada, tenha se constatado um número considerável de representações adequadas, considera-se essencial que os autores dediquem maior atenção com a atribuição de títulos, já que estes representam o campo principal para descrição clara e

objetiva da temática de um documento e também que, os comitês editoriais dos periódicos reforcem a aplicação de suas normas de publicação para os autores. É importante ressaltar a importância da comunidade científica de Ciência da Informação como orientadora da representação da informação na produção de textos científicos.

Reafirma-se então, que a boa qualidade referente à estruturação e os modos de representação, abarcados pela dimensão morfológica da pesquisa, tem o papel fundamental de propiciar a melhor compreensão de seu conteúdo no processo de comunicação científica.

Por fim, busca-se motivar a realização de futuras pesquisas em domínios ou especialidades da medicina e da própria Ciência da Informação, para analisar o desempenho dos descritores e termos na representação de temáticas e conceitos.

REFERÊNCIAS

ARBOIT, A. E.; GUIMARÃES, J. A. C. Conhecimento e linguagem na organização do conhecimento: aspectos dialógicos a partir da concepção de Bakhtin. In: XIV Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB 2013), Florianópolis. **Anais...** Disponível em:

<<http://enancib.sites.ufsc.br/index.php/enancib2013/XIVenancib/schedConf/presentations>>. Acesso em: 9 dez. 2013.

BADIR, S. A noção de texto em Hjelmslev. **Cadernos de Semiótica Aplicada**, v. 3, n. 2, dez. 2005. Disponível em: <<http://www.fclar.unesp.br/grupos/casa/CASA-home.html>>.

BARITÉ, M. et al. **Diccionario de Organización del Conocimiento. Clasificación, Indización, Terminología**. 5 ed. Montevideo: PRODIC, 2013.

BARITÉ, M. Referenciales teóricos vigentes en el área de tratamiento temático de la información y su expresión metodológica. Porto Alegre: ABEED, 1998. **Relatório técnico do II Encontro de Dirigentes dos cursos superiores de Biblioteconomia dos países do Mercosul**, Buenos Aires, nov. 1997.

BRAILE, D. M. et al. A importância da indexação para as revistas científicas. **Revista Brasileira de Cirurgia Cardiovascular**, São José do Rio Preto, v. 15, n. 4, dez. 2007.

BRANDAU, R. M. *et al.* Importância do uso correto dos descritores nos artigos científicos. **Revista Brasileira de Cirurgia Cardiovascular**, São José do Rio Preto, v. 20, n. 1, jan./mar. 2005.

FREITAS, J. L. *et al.* Proposta de metodologia para a recuperação da produção científica em Ciência da Informação na Base Brapci. **Ponto de Acesso**, Salvador, v. 4, n. 3, p. 45-67, dez 2010.

GOMES, V. P. O editor de revista científica: desafios da prática e da formação. **Informação & Informação**, Londrina, v. 15, n. 1, p. 147-172, jan./jun, 2010.

GRACIO, M. C. C.; OLIVEIRA, E. F. T. A inserção e o impacto internacional da pesquisa brasileira em "Estudos Métricos" uma análise na base Scopus. **Tendências da Pesquisa**

Brasileira em Ciência da Informação, v. 5, n. 1, 2012. Disponível em:
<<http://inseer.ibict.br/ancib/index.php/tpbci/article/view/71/113>>.

GUIMARÃES, C. A. et al. Brazilian Scientific Journals in Surgery: quality control in the abstract structure of non-experimental articles. **Acta Cirúrgica Brasileira**, v. 28, n. 1, 2013.

GUIMARÃES, J. A. C.; SALES, R. Análise documental: concepções do universo acadêmico brasileiro em Ciência da Informação. **DataGramZero**, Rio de Janeiro, v. 11, n. 1, fev. 2010.

HJØRLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science: domain-analysis. **Journal of the American Society for Information Science**, New York, v. 46, n. 6, p. 400-425, 1995.

KOBASHI, N. Y. Fundamentos semânticos e pragmáticos da construção de instrumentos de representação de informação. **DataGramZero**, v. 8, n. 6, dez. 2007. Disponível em
<http://www.dgz.org.br/dez07/Art_01.htm>. Acesso em: 18 jul. 2011.

LARA, M. L. G. de. Linguagem documentária e terminologia. **Transinformação**, Campinas, v. 16, n. 3, p. 231-240, set./dez. 2004.

LARA, M. L. G. de. Linguagens documentárias, instrumentos de mediação e comunicação. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 26, n. 1/2, p. 72-80, jan./jun. 1993.

LIMA, V. M. A. A informação documentária: codificação e decodificação. **Transinformação**, Campinas, v. 19, n. 2, p. 119-127, maio/ago., 2007.

LIMA, V. M. A.; BOCCATO, V. R. C. O desempenho terminológico dos descritores em Ciência da Informação do Vocabulário Controlado do SIBi/USP nos processos de indexação manual, automática e semi-automática. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 14, n. 1, jan./abr. 2009.

REGO, T. C. Produtivismo, pesquisa e comunicação científica: entre o veneno e o remédio. **Educação e Pesquisa**, São Paulo, v. 40, n. 2, p. 325-346, abr./jun. 2014.

SCOPUS. **Content over view Scopus**. 2013.

TÁLAMO, M. F.G. M; LARA, M. L. G. O campo da Lingüística Documentária. **Transinformação**, Campinas, v. 18, n. 3, p. 203-211, set./dez. 2006.

VAN DER LAAN, R. H. *et al.* Avaliação de descritores relativos às Ciências da Informação: relato de pesquisa. **Em Questão**, Porto Alegre, v. 10, n. 2, p. 337-347, jul./dez. 2004.