

A importância das palavras-chave dos artigos científicos da área das Ciências Farmacêuticas, depositados no Estudo Geral: estudo comparativo com os termos atribuídos na MEDLINE

The importance of keywords in scientific articles in Pharmaceutical Sciences, submitted in Estudo Geral: a comparative study of terms assigned in MEDLINE

Ana Miguéis

Serviço Integrado das Bibliotecas da Universidade de Coimbra (SIBUC)
E-mail: evamigueis@sib.uc.pt

Bruno Neves

Serviço Integrado das Bibliotecas da Universidade de Coimbra (SIBUC)
E-mail: bneves@sib.uc.pt

Ana Luísa Silva

Serviço Integrado das Bibliotecas da Universidade de Coimbra (SIBUC)
E-mail: als@uc.pt

Álvaro Trindade

Serviço Integrado das Bibliotecas da Universidade de Coimbra (SIBUC)
E-mail: atrindade@sib.uc.pt

José Augusto Bernardes

Serviço Integrado das Bibliotecas da Universidade de Coimbra (SIBUC)
E-mail: augusto@ci.uc.pt

Resumo

Objetivos: Este trabalho tem como objetivo analisar as palavras-chave usadas pelos autores da Universidade de Coimbra, nos artigos publicados na área das Ciências Farmacêuticas, comparando-as com os termos da linguagem documental Medical Subject Headings (MeSH), empregues na análise de conteúdo desses mesmos artigos. No âmbito dos artigos científicos, as palavras-chave expõem a abrangência de um assunto e os seus conceitos principais, que se podem revelar úteis para a indexação em mecanismos de pesquisa ou para a categorização do texto. A investigação sobre a importância e características das palavras-chave tem incidido sobre vários aspetos, como o da eficiência na recuperação da informação; o uso para a extração automática a partir de diferentes metodologias e algoritmos; o uso por parte dos autores e editores; mais recentemente, tem sido considerada ainda a sua utilização nos comportamentos de etiquetagem (*metatags*); e a comparação com os títulos, resumos e textos integrais, o que confirma do interesse que despertam. Mas, na revisão da literatura, foram poucos os trabalhos encontrados que abordassem a comparação das palavras-chave fornecidas pelos autores de artigos científicos e os descritores empregues pelos indexadores, analisando o contributo que podem dar para a escolha dos termos de indexação.

Metodologia: Após a revisão bibliográfica exploratória na literatura especializada, delimitou-se a amostra, constituída por 182 artigos científicos publicados em revistas internacionais, da comunidade da Faculdade de Farmácia, depositados no Estudo Geral, editados entre 1996 e 2012, e depositados até final de dezembro de 2012. O acesso a estes artigos científicos, armazenados em repositórios de acesso aberto, faz-se a partir dos metadados e do conteúdo de texto integral, aplicando técnicas que representam e expõem as relações semânticas entre documentos, em cujo âmbito os dados são relacionados. Os objetivos específicos a alcançar com este trabalho visam: (i) identificar e aferir o número de palavras-chave colocadas

na versão depositada no Estudo Geral e na versão publicada pela editora; e o número de descritores atribuídos pela MEDLINE; e (ii) verificar as palavras-chave, usadas pelos autores, que coincidem ou se relacionam com os termos do MeSH. No início da recolha da informação definiram-se algumas condições prévias, nomeadamente, fazer uma primeira comparação com as palavras-chave presentes na versão publicada pelo editor da revista, identificando-se, primeiro, os nomes das publicações periódicas a que pertenciam e confirmando-se, de seguida, o número de palavras-chave existentes em ambos os recursos. Num momento posterior e reunidas todas as palavras-chave existentes para cada um dos artigos, compararam-se, então, com os descritores MeSH atribuídos aos títulos indexados na base de dados MEDLINE. As palavras-chave recolhidas a partir dos metadados dos registos dos artigos depositados no repositório e da versão final da editora, foram organizadas, para cada artigo, numa grelha, juntamente com os termos eleitos pela MEDLINE. Para além destes elementos, foram adicionados campos de descrição, como, título, autor e ano de publicação. Na determinação quantitativa das relações identificadas entre as palavras-chave e os descritores MeSH, foram consideradas as relações de equivalência (idênticas), e as relações associativas (relacionadas).

Resultados e discussão: Os resultados obtidos confirmam o entendimento de que as palavras-chave recolhidas nos metadados dos artigos depositados no Estudo Geral não correspondem a todas as palavras-chave presentes na versão publicada pela editora da revista, o que determinou que a análise dos resultados incidisse apenas nas palavras-chave da versão da editora, comparando-as com os descritores MeSH atribuídos aos títulos indexados na base de dados MEDLINE. Nesta comparação, verificou-se que o número de palavras-chave é significativamente mais baixo do que os termos empregues pela MEDLINE. No que se refere à frequência do número de palavras-chave por artigo, a maioria dos artigos não apresenta mais que cinco (5) palavras-chave, sendo bem mais aumentada a frequência da atribuição dos descritores da MEDLINE. Do ponto de vista semântico, no que respeita às relações de equivalência e de associação entre as palavras-chave e os descritores da MEDLINE, perto de metade das palavras-chave apresentavam relações de equivalência ou associativas, em proporções semelhantes. Do mesmo modo, a frequência das relações estabelecidas pelas palavras-chave situa-se, maioritariamente, entre 1 a 2 conceitos equivalentes e 1 a 2 conceitos associados por artigo.

Conclusões: Face aos resultados apurados, pode-se afirmar que as palavras-chave representam uma fonte de acesso vantajosa aos artigos científicos, o que valoriza a importância das palavras-chave escolhidas pelos autores para a representação da informação dos artigos indexados pela base de dados MEDLINE. Assim, nos 182 artigos analisados, o número de relações equivalentes e associativas é aproximado e, ainda que as percentagens não sejam muito elevadas, permite concluir por uma influência direta ou indireta na escolha dos descritores. Tal resultado pode promover a utilização das palavras-chave no ensino da indexação por assuntos e nos esforços a desenvolver para automatizar o processo de indexação. Ao utilizar o Estudo Geral como suporte a este trabalho, pretende-se chamar a atenção para o repositório institucional da Universidade de Coimbra e para a importância dos autores, enquanto intervenientes ativos no processo de representação e recuperação da informação, através das palavras-chave registadas nos seus artigos. Reconhecer a importância das palavras-chave na representação e recuperação de informação é uma forma de promover o autodepósito e estimular o completo preenchimento dos metadados dos documentos depositados; e contribuir para uma maior visibilidade, probabilidade de aumento de citações, e maior impacto da produção científica dos autores depositantes e da própria instituição.

Palavras-chave: Recuperação da informação; Linguagens de indexação; Palavras-chave; Artigos científicos; Medical Subject Headings; MEDLINE

Abstract

Objectives: To analyze the keywords used by authors at the University of Coimbra in papers published in the field of the Pharmaceutical Sciences, comparing them to Medical Subject Headings (MeSH) through the use of content analysis. Keywords used in scientific papers show the scope of a subject and its main concepts and are useful in indexing search mechanisms and categorizing texts. Investigation into the importance and characteristics of keywords has focused on various aspects such as efficiency in retrieving information, automatic extraction using different methods and algorithms; the use of keywords by authors and editors; and more recently, their use in metatags and comparison with titles, abstracts and full texts, confirms the interest in the topic. Few papers, however, were found in the literature comparing keywords provided by authors of scientific papers and the descriptors used by indexers, in order to analyze the authors' contribution in choosing indexing terms.

Method: After an exploratory bibliographic review in specialized literature, the sample was composed of 182 scientific papers produced by the community of the School of Pharmacy, University of Coimbra, deposited in Estudo Geral and edited between 1996 and 2012 (up to December 2012) and eventually published in international journals. Access to these papers, stored in open-access repositories, was accomplished by analyzing metadata and the content of full texts using techniques that represent and expose semantic relationships among documents, under which data are related. This paper's specific objectives were: (i) to identify and measure the number of keywords used in the versions deposited in Estudo Geral and in the versions published by editors; and the number of descriptors assigned by MEDLINE; and (ii) verify the keywords used by the authors that coincide with or relate to the MeSH terms. Some conditions were established prior to data collection, namely to first compare the keywords used in the version published by the journal's editor first identifying the names of the periodicals to which they belonged and then confirm the number the keywords in both resources. Afterwards, having all the keywords in each of the papers together, we compared the MeSH descriptors assigned to the titles indexed in the MEDLINE database. The keywords collected from the metadata of the records of papers deposited in the repository and the editor's final versions were organized for each paper in a grid together with the MEDLINE's terms. Blank spaces for title, author and year of publication were also added. When quantitatively determining the relationships identified among the keywords and MeSH descriptors, we considered equivalence (identical) and associative (related) relationships.

Results and discussion: The results confirm that the keywords collected from the metadata of the papers deposited in the Estudo Geral do not correspond to all the keywords present in the published versions, which determined that the analysis of results focused only on the keywords presented in the editors' versions, comparing them with the MeSH descriptors assigned to the titles indexed in the MEDLINE database. This comparison demonstrated that the number of keywords is significantly lower than the number of terms used by MEDLINE. In regard to the number of keywords per paper, most papers do not present more than five (5) key words, while the number of MEDLINE descriptors is much higher. From a semantic point of view, regarding equivalence and association relationships among the keywords and MEDLINE descriptors, about half of the keywords presented equivalence or associative relationships in similar proportions. Additionally, the relationships established by the keywords mostly present from 1 to 2 equivalent concepts and from 1 to 2 associate concepts.

Conclusions: The results show that keywords are an advantageous source of access to scientific papers, highlighting the relevance of keywords chosen by authors to represent information provided by papers indexed in the MEDLINE database. Therefore, the number of equivalent and associative relationships in the 182 papers analyzed is close and even though the percentages are not very high, they enable us to conclude that authors directly or indirectly influence the choice of descriptors. Such a result may encourage the use of keywords to teach indexing per subject and improve efforts toward the automation of the indexing process. The use of Estudo Geral to support this study is intended to call attention to the institutional repository at the University of Coimbra and to the importance of authors as active actors in the process of the representation and retrieval of information through the use of keywords they choose for their papers. Acknowledging the importance of keywords in the representation and retrieval of information is a way to promote self-depositing and to encourage the completion of metadata of the documents deposited; this would contribute to greater visibility, an increased number of citations, and greater impact on the part of scientific production by the institution's authors and the institution itself.

Keywords: Information retrieval; Indexing languages; Keywords; Scientific papers; Medical Subject Headings; MEDLINE

1. Introdução

Este trabalho surgiu do interesse pelo uso das palavras-chave, enquanto instrumento de representação da informação contida nos documentos. A organização da informação e a sua representação são aspetos essenciais da ciência da informação, que tem como uma das suas finalidades a partilha da informação, enquanto recurso estratégico na cadeia que gera o conhecimento (TONELLO; LUNARDELLI; ALMEIDA JÚNIOR, 2012).

Hoje coexistem vários modelos de representação e de recuperação da informação, num ambiente dinâmico e em constante mutação, como o recurso a termos extraídos diretamente da linguagem natural, em que se inserem as expressões utilizados pelos próprios autores no título, resumo e texto integral. Neste modelo não existe qualquer controlo-formal ou semântico dos termos, sendo esta uma linguagem que se afirma e consolida em áreas específicas e limitadas, que particularizam e identificam um código de comunicação nas diferentes áreas científicas (SIMÕES, 2008), materializando-se em listas de palavras-chave, ou seja, termos vocabulares retirados dos documentos e que correspondem a uma escolha dos autores.

Os editores de revistas científicas, quando da submissão dos artigos, solicitam aos autores que coloquem as palavras-chave nos seus textos (GONÇALVES, 2008), que referem, habitualmente, a abrangência de um assunto e os seus conceitos principais. Este procedimento, no âmbito dos artigos científicos, pode-se revelar útil para a indexação em mecanismos de pesquisa ou categorização do texto (ERCAN; CICEKLI, 2007).

O uso das palavras-chave potencia o acesso ao conteúdo dos documentos, para além da informação que é representada pelo título e resumo; traduz o pensamento dos autores, e mantém o contacto com a realidade da prática quotidiana, acompanhando a evolução científica e tecnológica, que é refletida pelos documentos.

A investigação sobre a importância e características das palavras-chave tem incidido sobre vários aspetos, como o da eficiência na recuperação da informação; a extração automática a partir de diferentes metodologias e algoritmos; o uso por parte dos autores e editores; a utilização nos comportamentos de etiquetagem (*metatags*); e a comparação com os títulos, resumos, textos e descritores atribuídos.

No entanto, na revisão da literatura, identificaram-se poucos trabalhos que abordassem a comparação das palavras-chave, fornecidos por autores de artigos científicos, com os descritores empregues pelos indexadores e o contributo que podem dar para a escolha dos termos de indexação¹. Com base no pressuposto de conhecer a relação conceptual entre as palavras-chave e os descritores empregues, preparou-se a presente comunicação², que analisa, comparativamente, as palavras-chave empregues pelos autores nos seus artigos e os termos da linguagem de indexação que lhe são associados.

O repositório científico da Universidade de Coimbra, designado Estudo Geral, tem a missão de divulgar e preservar a produção científica de autores ligados à Universidade de Coimbra e afirma-se como um instrumento que preserva e maximiza o impacto da investigação realizada no seio da instituição. Assim sendo, foi com base no Estudo Geral que se realizou a escolha do conjunto de documentos a analisar, e que incidiu sobre a coleção “Artigos em Revistas Internacionais” da comunidade “Faculdade de Farmácia”. Esta opção resulta do facto de os resultados da investigação científica realizada serem divulgados através desta tipologia documental (o artigo científico), estando esta prática fortemente implementada e apoiada por critérios internacionais de avaliação. O acesso a estes artigos científicos, armazenados em repositórios de acesso aberto, faz-se a partir dos metadados e do conteúdo de texto integral. Para este trabalho foram identificadas e recolhidas todas as palavras-chave introduzidas pelos respetivos autores e comparadas com os termos de indexação atribuídos aos artigos da base de dados MEDLINE, pela National Library of Medicine (NLM) dos Estados Unidos.

¹ Entre eles, encontram-se dois artigos de Gil-Leiva e Alonso-Arroyo (2005 e 2007), que analisam as palavras-chave dos autores e as comparam com os descritores usados em várias bases de dados, concluindo que existe uma relação direta, ou indireta, na seleção dos descritores

² Esta comunicação surgiu como forma de dar continuidade ao trabalho já iniciado e publicado, disponível em: <<http://www.portalseer.ufba.br/index.php/revistaici/article/view/8045>>.

2. Metodologia

A amostra delimitada para este estudo é constituída por 182 artigos científicos publicados entre 1996 e 2012 em revistas internacionais da área das Ciências Farmacêuticas, depositados na comunidade da Faculdade de Farmácia da Universidade de Coimbra do repositório Estudo Geral, que compreendia, no seu conjunto, um total de 292 artigos científicos depositados até final de Dezembro de 2012.

Foi pela importância atribuída à MEDLINE – principal base de dados bibliográficos de artigos científicos da área das Ciências da Saúde, com cerca de 19 milhões de referências, publicadas desde 1948 – e à linguagem controlada que utiliza para indexar os seus artigos – o Medical Subject Headings (MeSH) –, que a escolha recaiu sobre este recurso de informação. Assim, definiram-se algumas condições prévias, no início da recolha da informação, e que consistiram em incluir apenas: os artigos com acesso ao texto integral; os artigos com palavras-chave em língua inglesa; os artigos editados em publicações periódicas referenciadas na base de dados MEDLINE.

No decurso deste levantamento, houve a perceção de que o número de palavras-chave recolhidas nos metadados dos artigos depositados no Estudo Geral eram em número reduzido e decidiu-se, então, fazer uma primeira comparação com as palavras-chave presente na versão publicada pelo editor da revista, identificando-se, desse modo, os nomes das publicações periódicas a que pertenciam e confirmando-se, de seguida, o número de palavras-chave existentes em ambos os recursos. Num momento posterior, e reunidas todas as palavras-chave existentes para cada um dos artigos, compararam-se, então, com os títulos existentes na base de dados MEDLINE e com os termos de indexação do MeSH.

As palavras-chave recolhidas a partir dos metadados dos registos dos artigos depositados no repositório, e da versão final da editora, foram organizadas numa grelha, juntamente com os termos controlados eleitos pela MEDLINE. Para além destes elementos, foram adicionados campos de descrição do artigo, como, título, autor e ano de publicação. Na determinação quantitativa das relações identificadas entre as palavras-chave e os termos MeSH, foram consideradas as relações de equivalência, correspondendo à classificação de 1 ponto, e as relações de semelhança a meio ponto (0,5), conforme demonstra a tabela seguinte.

Tabela 1 – Exemplo dos critérios definidos nas relações

| Palavras-chave (PCH) | Descritores MeSH (DE) | Equivalentes (E) | Associadas (A) |
|----------------------|--|------------------|----------------|
| Alginate | Alginates | 1 | |
| Il-1 | Interleukin-1 | 1 | |
| β -lactamases | beta-Lactamases | 1 | |
| FISH | In Situ Hybridization, Fluorescence | 1 | |
| Mitochondria | Mitochondria, Liver | | 0,5 |
| Ethanol content | Ethanol | | 0,5 |
| Essential oil | Oils, Volatile | | 0,5 |
| AIDS | HIV-1 | | 0,5 |

Para o reconhecimento das palavras-chave equivalentes aos descritores foram definidos os seguintes critérios: termos simples com a mesma grafia; termos com variantes em gênero ou número, termos compostos com a mesma grafia, ainda que com ordem invertida, e termos com o mesmo significado do ponto de vista semântico. Foram, também, considerados os termos representados por acrônimos ou símbolos químicos desdobrados, e os elementos textuais do alfabeto grego, transliterados para o alfabeto português.

Em todas as restantes palavras-chave em que se identificou algum tipo de relação semântica com os termos escolhidos pelo MeSH, foram consideradas como termos associados.

Foi criada uma segunda tabela (que se exemplifica na Tabela 2) com os resultados obtidos para todos os 182 artigos, registrando o número de relações (equivalências e associações) entre as palavras-chave usadas pelos autores e os descritores da MEDLINE e com a classificação atribuída a cada tipo de relação.

Tabela 2 – Exemplo do nº de relações (equivalentes e associadas) identificadas em cada artigo

| Artigo | PCH | DE | #PCH | #DE | #Usadas | #E | #A | Total (E+A) |
|-----------------------------|---|--------------------------------|------|-----|---------|----|----|-------------|
| Gomes-da-Silva, 2012 | <i>Breast cancer</i> | <i>Breast Neoplasms</i> | 5 | 19 | 2 | 1 | 1 | 1,5 |
| | Dual-targeted delivery | Cell Line | | | | | | |
| | Ligand-mediated targeting | Cell Line, Tumor | | | | | | |
| | siRNA | Down-Regulation | | | | | | |
| | Stable nucleic acid lipid particles (SNALP) | Endothelial Cells | | | | | | |
| | | Female | | | | | | |
| | | Green Fluorescent Proteins | | | | | | |
| | | Humans | | | | | | |
| | | Liposomes | | | | | | |
| | | Lysosomes | | | | | | |
| | | Nanoparticles | | | | | | |
| | | Neovascularization, Pathologic | | | | | | |
| | | Peptides | | | | | | |
| | | Phosphoproteins | | | | | | |
| | | Polyethylene Glycols | | | | | | |
| | RNA-Binding Proteins | | | | | | | |
| | RNA, Messenger | | | | | | | |
| | RNA, Small Interfering | | | | | | | |
| | Tumor Microenvironment | | | | | | | |

3. Resultados e discussão

Dos 182 artigos selecionados, 43 não possuíam qualquer informação nos metadados referentes ao “assunto” no registo dos artigos depositados no Estudo Geral, embora essa informação constasse no ficheiro anexado. Em termos quantitativos, identificaram-se 700 palavras-chave nos metadados do registo dos artigos no Estudo Geral, e 976 na versão da editora. Aos mesmos artigos, indexados na MEDLINE, atribuíram-se 2326 descritores MeSH, conforme referido na Figura 1.

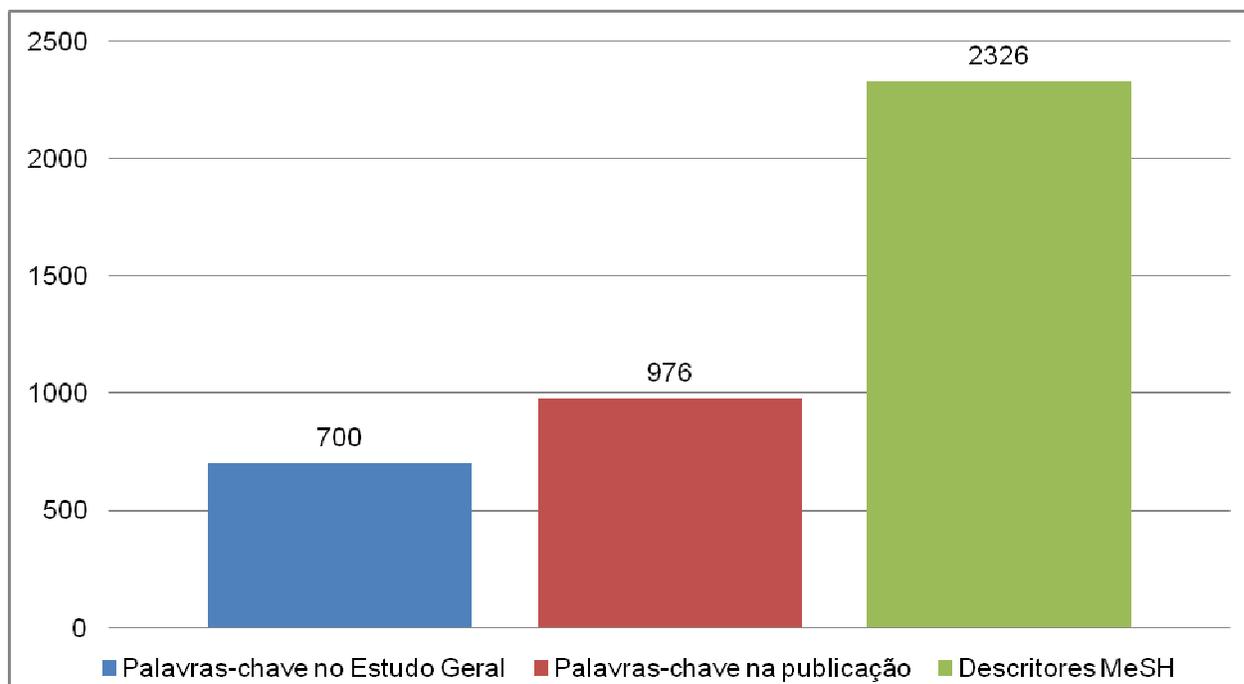


Figura 1 - Número de palavras-chave dos artigos e descritores da MEDLINE (n=182)

Os dados obtidos traduzem a percepção de que as palavras-chave retiradas dos metadados dos artigos depositados no Estudo Geral, ainda que sendo as mesmas, são em número mais reduzido do que as que se encontram na editora da publicação, representado apenas 72% das palavras-chave desta versão. Este facto determinou que a análise dos resultados seguintes incidisse apenas nas palavras-chave existentes na versão da editora, comparando-as com os descritores MesH atribuídos aos títulos indexados na MEDLINE.

Assim, no que se refere à frequência do número de palavras-chave por artigo, a sua maioria apresenta entre 5 a 6 palavras-chave, valores registados em 63% dos artigos, com uma média de 5,36 palavras-chave. A frequência do número de descritores MeSH varia entre 11 a 14 descritores, o que corresponde a 33% dos artigos, com uma média de 12,78 descritores (Tabela 3).

Tabela 3 - Frequência do número de palavras-chave na publicação e de descritores na MEDLINE

| | PCH | Freq. Absoluta | Freq. Relativa (%) | DE | Freq. Absoluta | Freq. Relativa (%) |
|--------------|-------------|----------------|--------------------|--------------|----------------|--------------------|
| | 1 | 2 | 1% | 4 | 3 | 2% |
| | 2 | 1 | 1% | 5 | 7 | 4% |
| | 3 | 14 | 8% | 6 | 11 | 6% |
| | 4 | 21 | 12% | 7 | 9 | 5% |
| | 5 | 62 | 34% | 8 | 12 | 7% |
| | 6 | 53 | 29% | 9 | 12 | 7% |
| | 7 | 20 | 11% | 10 | 10 | 5% |
| | 8 | 7 | 4% | 11 | 18 | 10% |
| | 9 | 1 | 1% | 12 | 13 | 7% |
| | 13 | 1 | 1% | 13 | 13 | 7% |
| Total | 976 | 182 | 100% | 14 | 16 | 9% |
| Média | 5.36 | | | 15 | 7 | 4% |
| Moda | 5 | | | 16 | 10 | 5% |
| | | | | 17 | 6 | 3% |
| | | | | 18 | 11 | 6% |
| | | | | 19 | 4 | 2% |
| | | | | 20 | 4 | 2% |
| | | | | 21 | 4 | 2% |
| | | | | 22 | 2 | 1% |
| | | | | 23 | 3 | 2% |
| | | | | 24 | 1 | 1% |
| | | | | 25 | 2 | 1% |
| | | | | 26 | 3 | 2% |
| | | | | 29 | 1 | 1% |
| | | | Total | 2326 | 182 | 100% |
| | | | Média | 12.78 | | |
| | | | Moda | 11 | | |

Do ponto de vista semântico, foram identificadas 533 palavras-chave com algum tipo de relação: 257 equivalentes (26%) e 276 associadas (28%), conforme se indica na Figura 2. Regista-se um conjunto de 443 palavras-chave (46%) que não apresentam qualquer tipo de relação.

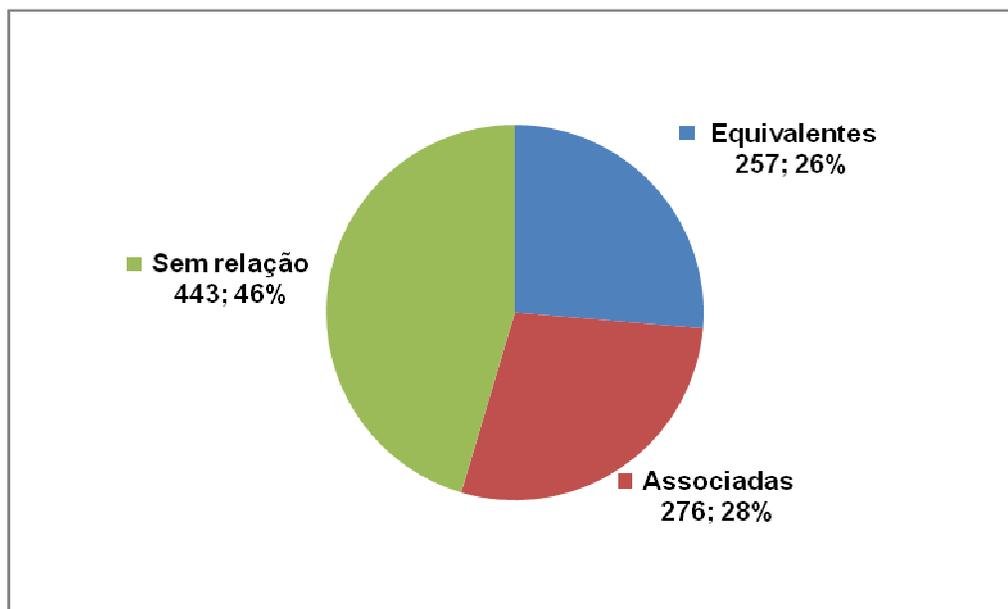


Figura 2 – Relações identificadas na comparação das palavras-chave com os descritores MeSH (n=182)

Na tabela seguinte (Tabela 4) indica-se a distribuição dos 182 artigos por anos civis, sendo de 15 anos, o período temporal abrangido (entre 1996 e 2012); e 166 dos artigos (91%) concentram-se nos anos de 2001 a 2009. De acordo com a pontuação definida para as relações estabelecidas entre as palavras-chave e os descritores MeSH (1 ponto para as relações de equivalência e 0,5 para as associativas), as percentagens obtidas foram de 27% equivalências e 24% associativas, num total de 51%.

Tabela 4 – Dados gerais obtidos por ano de publicação dos artigos

| Ano | #Artigos | #PCH | #Usadas | #DE | E | A | Total (E+A) | %E | %A | Total % (E+A) |
|------|----------|------|---------|-----|----|------|-------------|-----|-----|---------------|
| 1996 | 1 | 3 | 1 | 4 | 1 | 0 | 1 | 33% | 0% | 33% |
| 1998 | 5 | 22 | 13 | 54 | 8 | 2,5 | 10,5 | 36% | 23% | 59% |
| 1999 | 1 | 8 | 3 | 11 | 2 | 0,5 | 2,5 | 25% | 13% | 38% |
| 2000 | 4 | 24 | 10 | 53 | 6 | 2 | 8 | 25% | 17% | 42% |
| 2001 | 14 | 71 | 41 | 194 | 22 | 9,5 | 31,5 | 31% | 27% | 58% |
| 2002 | 10 | 57 | 21 | 114 | 13 | 4 | 17 | 23% | 14% | 37% |
| 2003 | 25 | 130 | 78 | 298 | 43 | 17,5 | 60,5 | 33% | 27% | 60% |
| 2004 | 13 | 81 | 33 | 168 | 18 | 7,5 | 25,5 | 22% | 19% | 41% |
| 2005 | 20 | 108 | 60 | 252 | 29 | 15,5 | 44,5 | 27% | 29% | 56% |
| 2006 | 17 | 86 | 55 | 207 | 25 | 15 | 40 | 29% | 35% | 64% |
| 2007 | 22 | 114 | 65 | 327 | 29 | 18 | 47 | 25% | 32% | 57% |
| 2008 | 33 | 179 | 98 | 426 | 37 | 30,5 | 67,5 | 21% | 34% | 55% |

| | | | | | | | | | | |
|--------------|------------|------------|------------|-------------|------------|------------|------------|-------------|--------------|--------------|
| 2009 | 12 | 65 | 36 | 143 | 14 | 11 | 25 | 22% | 34% | 55% |
| 2011 | 4 | 23 | 17 | 56 | 9 | 4 | 13 | 39% | 35% | 74% |
| 2012 | 1 | 5 | 2 | 19 | 1 | 1 | 1,5 | 20% | 20% | 40% |
| Total | 182 | 976 | 533 | 2326 | 257 | 138 | 395 | | | |
| Média | | | | | | | | 27,4 | 23,93 | 51,27 |

A frequência das relações estabelecidas pelas palavras-chave (Tabela 5) situa-se maioritariamente entre 1 a 2 conceitos equivalentes (58%) e 1 a 2 conceitos associados (70%) por artigo, o que permite dizer que as palavras-chave escolhidas pelos autores são importantes para a representação da informação.

Tabela 5 - Frequência do número de palavras-chave equivalentes e associadas

| | Equivalentes | Freq. Absoluta | Freq. Relativa (%) | Associadas | Freq. Absoluta | Freq. Relativa (%) |
|--------------|--------------|----------------|--------------------|-------------|----------------|--------------------|
| | 0 | 44 | 24% | 0 | 27 | 15% |
| | 1 | 61 | 34% | 1 | 66 | 36% |
| | 2 | 43 | 24% | 2 | 62 | 34% |
| | 3 | 28 | 15% | 3 | 22 | 12% |
| | 4 | 5 | 3% | 4 | 4 | 2% |
| | 5 | 1 | 1% | 5 | 1 | 1% |
| Total | 256 | 182 | 100% | 277 | 182 | 100% |
| Média | 1.40 | | | 1.52 | | |
| Moda | 1 | | | 1 | | |

4. Conclusões

Face aos resultados obtidos neste trabalho, é possível dizer que as palavras-chave representam uma importante fonte de acesso aos artigos científicos. Em média, cada artigo regista cinco palavras-chave, em que uma ou duas apresentam relações com os descritores empregues pela MEDLINE. Nos 182 artigos analisados, 26% das palavras-chave introduzidas são equivalentes e 28% são associadas aos descritores da MEDLINE, o que representa um conjunto de 45% das palavras-chave existentes nos artigos. O número de relações equivalentes (257) e associativas (276) é aproximado, mas o critério aplicado determina que as relações de equivalência têm um valor superior (27%) ao valor atribuído às relações associativas (24%), com um total de 51,27%. Não sendo percentagens muito elevadas, não deixam, contudo, de apresentar

dados significativos e que vão ao encontro de resultados já obtidos em outros trabalhos³, permitindo concluir por uma influência direta ou indireta na escolha dos descritores. Tal, poderá promover a aplicação das palavras-chave no estudo da indexação por assuntos e nos esforços a desenvolver para automatizar o processo de indexação.

Utilizar o Estudo Geral, como suporte a este trabalho, teve a intenção de chamar a atenção para os repositórios institucionais e para a importância dos autores, intervenientes ativos no processo de representação e recuperação da informação através das palavras-chave registadas nos seus artigos. Evidenciar a importância das palavras-chave na representação e na recuperação da informação é uma forma de promover o autodepósito e de estimular o completo preenchimento dos metadados dos documentos depositados. Deste modo contribui-se para uma maior visibilidade, mais probabilidades de citação e um conseqüente aumento do impacto da produção científica dos autores depositantes e da própria instituição.

Referências

ERCAN, G.; CICEKLI, I. Using lexical chains for keyword extraction. **Information processing and management**, v. 43, p. 1705-1714, 2007. Disponível em: <<http://dx.doi.org/10.1016/j.ipm.2007.01.015>>. Acesso em: 05 jun. 2013.

GIL-LEIVA, I.; ALONSO-ARROYO, A. Keywords given by authors of scientific articles in database descriptors. **Journal of the American Society for Information Science**, v. 58, n. 8, p. 1175–1187, 2007. Disponível em: <<http://dx.doi.org/10.1002/asi.20595>>. Acesso em: 03 jun. 2013

GIL-LEIVA, I.; ALONSO-ARROYO, A. La relación entre las palabras clave aportadas por autores de artículos de revista y su indización en las bases de datos ISOC, IME e ICYT. **Revista Española de Documentación Científica**, v. 28, n. 1, p. 62-79, 2005. Disponível em: <<http://redc.revistas.csic.es/index.php/redc/article/view/165/219>>. Acesso em: 05 jun. 2013.

GONÇALVES, A. L. Uso de resumos e palavras-chave em Ciências Sociais: uma avaliação. **Encontros Bibli**, v. 13, n. 26, 2008. Disponível em: <<http://www.redalyc.org/articulo.oa?id=14712794006>>. Acesso em: 05 jun. 2013.

MIGUÉIS, A.; NEVES, B. Uma abordagem à linguagem de indexação dos artigos científicos

³ Veja-se Miguéis e Neves (2013); Gil-Leiva e Alonso-Arroyo (2005; 2007)

depositados no repositório científico da Universidade de Coimbra. **Ponto de Acesso**, v. 7, n. 1, p. 81-115, 2013. Disponível em: <<http://www.portalseer.ufba.br/index.php/revistaici/article/view/8045>>. Acesso em: 01 set. 2013.

U.S. NATIONAL LIBRARY OF MEDICINE. **Fact Sheet Medical Subject Headings (MeSH®)**. Bethesda, Maryland: NLM. Disponível em: <[URL:http://www.nlm.nih.gov/pubs/factsheets/mesh.html](http://www.nlm.nih.gov/pubs/factsheets/mesh.html)>. Acesso em: 02 set. 2013.

_____. **MEDLINE**. Bethesda, Maryland: NLM. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed>>. Acesso em: 02 set. 2013.

_____. **MeSH**. Bethesda, Maryland: NLM. Disponível em: <[URL:http://www.nlm.nih.gov/mesh.html](http://www.nlm.nih.gov/mesh.html)>. Acesso em: 02 set. 2013.

SIMÕES, M. G. **Da abstração à complexidade formal**: relações conceptuais num tesauro. Coimbra: Almedina, 2008.

TONELLO, I.; LUNARDELLI, R.; ALMEIDA JUNIOR, O. Palavras-chave: possibilidades de mediação da informação. **PontodeAcesso**, v. 6, p. 21-34, 2012. Disponível em: <<http://www.portalseer.ufba.br/index.php/revistaici/article/view/4524>>. Acesso em: 01 mar. 2013.

UNIVERSIDADE DE COIMBRA. **Estudo Geral**. Coimbra: UC. Disponível em: <<https://estudogeral.sib.uc.pt/>>. Acesso em: 01 set. 2013.

DOI: [10.11606/issn.2178-2075.v4i2p112-125](https://doi.org/10.11606/issn.2178-2075.v4i2p112-125)