

METODOLOGIA DE COLETA E PADRONIZAÇÃO DE ARTIGOS PUBLICADOS PELA FIOCRUZ

Para recuperação das publicações acadêmicas da Fiocruz, o Observatório em Ciência, Tecnologia e Inovação em Saúde desenvolveu uma metodologia própria e específica para esta análise. Esta foi dividida em quatro etapas, conforme diagrama abaixo:

1) Seleção das bases de dados e elaboração da estratégia de busca

Para recuperação dos artigos científicos foram utilizadas três bases de dados disponíveis no Portal Capes: Web of Science (Clarivate Analytics), Scopus (Elsevier) e SciELO. A estratégia usada em cada uma das bases foi a seguinte.

Web of Science (WoS): a busca foi feita em campo específico para recuperação de publicações através da instituição de afiliação do autor “OG= Organização Consolidada”, disponível na seção de busca avançada. Neste campo buscou-se pelos termos “*Center for Technological Development in Health (CDTS) OR Fundação Oswaldo Cruz*”. Além dessa lista consolidada pela WoS, utilizou-se uma lista de referência já previamente compilada pelo Observatório (Anexo). Estas foram inseridas no campo “OO=Organização”.

Scopus: a busca foi feita no campo de afiliação dos autores (AFFIL). Neste utilizou-se a lista de referência já previamente compilada pelo Observatório (a mesma utilizada na WoS) e a lista consolidada de organizações disponível na WoS.

SciELO: a busca foi feita no campo “OG=Organização”, utilizando-se a lista consolidada de organizações disponível na WoS e a lista de referência compilada pelo Observatório.

É importante salientar que a lista de palavras-chaves relativas à instituição foi redundante no intuito de aumentar a sensibilidade da estratégia de busca.

2) Download dos registros, importação para o software e validação da estratégia de busca

As publicações recuperadas em cada uma das três bases foram inseridas no software VantagePoint® utilizando filtros específicos, e reunidas por meio da função “data fusion”.

Para validar a estratégia de busca, utilizou-se uma lista de publicações da Fiocruz como referência. A lista continha dados de título, autor e revista de 291 publicações do CDTs no período analisado (2012/maio 2018). Observou-se que 19 dos 291 artigos listados não foram recuperados com a estratégia elaborada.

Todavia, os artigos não recuperados foram publicados em revistas não indexadas por nenhuma das três bases de dados utilizadas. Esta validação demonstrou que a estratégia adotada foi robusta, consistente e reprodutível para monitoramento periódico das publicações da Fiocruz.

3) Eliminação de duplicatas e harmonização dos registros

Para organização do banco de dados do Observatório, foi necessário remover as publicações em duplicata. Inicialmente, as duplicatas foram removidas utilizando-se como critério os campos DOI, título, revista e ano. Registros com esses quatro campos iguais foram removidos e apenas um registro foi mantido.

Observou-se que algumas revistas, por estarem indexadas em bases diferentes, não tinham seu título apresentado da mesma forma, por isso uma harmonização manual para padronização desse campo foi feita com base no ISSN das revistas.

A mesma padronização foi necessária para os seguintes campos:

- Títulos das publicações (padronização do idioma e forma de escrita);
- País de filiação dos autores (traduzido para o idioma português); e

4) Limpeza final e geração dos indicadores

A harmonização das publicações foi finalizada com a exclusão de publicações que não possuíam pelo menos um autor afiliado à Fiocruz. Como a estratégia de busca utilizada teve caráter abrangente, havia a possibilidade de inclusão de instituições que não faziam parte do universo de análise.

Anexo: Lista de referência compilada pelo Observatório

("Aggeu Magalhaes" OR "Bio Manguinhos" OR "Biomanguinhos" OR "Bio-Manguinhos" OR "Biomanguinlios" OR "Carlos Chagas" OR "Centro de Criação de Animais de Laboratório" OR "Centro de Desenvolvimento Tecnológico em Saúde" OR "Ctr Desenvolvimento Tecnol Saude" OR "Escola Nacional de Saude Publica" OR "Escola Nacl Salude Publ" OR "Escola Nacl Saude Publ" OR "Escola Nacl Saude Publica" OR "Escola Natl Saude Publ" OR "Evandro Chagas" OR "Evandro Chagoas" OR "Far Manguinhos" OR "Farmanguinhos" OR "Fernades Figueira" OR "Fernandes Figueira" OR "FICORUZ" OR "FIO CRUZ" OR "FIO-CRUZ" OR "Fiocntz" OR FIOCUZ OR "**Fundação Oswaldo Cruz**" OR "**Fiocruz**" OR "Fundacao Fiocruz" OR "FIOCRUZ" OR "Goncalo Moniz" OR "Gonçalo Moniz" OR "Helio Fraga" OR "Inst Oswald Cruz" OR "Inst Oswaldo Fdn" OR "Inst Tecnol Immunobiol" OR "Instituto de Comunicação e Informação Científica e Tecnológica em Saúde" OR "Instituto de Ciência e Tecnologia em Biomodelos" OR "Instituto de Tecnologia em Imunobiológicos" OR "Instituto Nacional de Controle de Qualidade em Saúde" OR "Joaquim Venancio" OR "Leonidas & Maria Deane" OR "Leonidas & Maria Deane" OR "Leônidas e Maria Deane" OR "Leonidas Maria Deane" OR "Osaldo Cruz" OR "Osawaldo Cruz" OR "Osqaldo Cruz" OR "Oswaldo Cruz" OR "Oswaldo Cruz" OR "Oswald Cruz" OR "Oswald Cruz" OR "Oswald Cruz" OR "OSWALD CRUZ" OR "Oswaldo" OR "Oswaldo Crus" OR "OSWALDO CRUS" OR "Oswaldo Crusz" OR "Oswaldo Cruz" OR "Oswaldo Curz" OR "OSWALDO CURZ" OR "Oswalo Cruz" OR "OSWCRU" OR "Owaldo Cruz" OR "Owaldo Cruz" OR "Programa Comp Cient Qswald Cruz" OR "Rene Rachou" OR "René Rachou" OR "Sergio Arouca" OR INCQS OR CECAL OR CPqAM OR CPqGM OR CPqLMD OR CPqRR OR CESTEHE OR CPqRR)