

## **METODOLOGIA DE MAPEAMENTO COMPUTACIONAL DE INFORMAÇÕES MÉDICAS: APLICAÇÃO EM LAUDOS DE COLOSCOPIA E MANOMETRIA ANORRETAL\***

CHERMAN EA; SPOLAÔR N; LEE HD; COSTA LHD; FAGUNDES JJ; COY CSR; WU FC

Laboratório de Bioinformática/UNIOESTE/PTI e Serviço de Coloproctologia/Faculdade de Ciências Médicas/UNICAMP

Processos computacionais, como o de Mineração de Dados (MD), podem auxiliar na análise de grandes conjuntos de dados. Para que sejam aplicados é necessário que os dados estejam em Bases de Dados digitais (BD).

**OBJETIVO:** Propor uma metodologia de mapeamento computacional de informações para BD aplicada a Laudos Médicos (LM) de Coloscopia (C) e de Manometria Anorretal (MA).

**MATERIAL E MÉTODOS:** Neste trabalho são utilizadas informações descritas em língua natural contidas em 100 LM de C e em 68 LM de MA provindos do Serviço de Coloproctologia da UNICAMP. A metodologia utilizada para mapeamento dessas informações é composta por duas fases: (1) Definição dos atributos da BD e do Arquivo de Padronização (AP) em conjunto com especialistas e (2) Mapeamento automático dos LM. As padronizações do AP têm como intuito transformar termos sinônimos para um único termo adequado à BD e aumentar a uniformidade da descrição dos LM. O preenchimento da BD é realizado na fase (2) com base nos documentos resultantes da fase (1) e em um algoritmo computacional, o qual tem como objetivo a busca de informações dos LM, a padronização conforme registros do AP e o preenchimento da BD com as informações.

**RESULTADOS:** Na fase (1) foram definidas BD com 277 e 16 atributos para os LM de C e MA, respectivamente. O AP foi composto por 274 e 115 padronizações para os LM de C e MA, respectivamente, o que possibilita observar que em relação à quantidade de laudos, os LM de C contêm 1,6 vezes mais padronizações que os LM de MA. Na fase (2) os LM foram processados pelo algoritmo computacional e foi possível mapear automaticamente para as BD 82% e 99% das informações contidas nos LM de C e MA, respectivamente. A maior uniformidade na descrição dos LM de MA, observada pela quantidade de padronizações, contribuiu para estes resultados, de modo que para os LM de MA a metodologia foi 17% mais eficiente. As informações não preenchidas foram mapeadas manualmente, completando 100% de preenchimento.

**CONCLUSÃO:** Com base nos resultados foi possível observar que a metodologia pode ser considerada adequada por permitir o mapeamento completo das informações contidas nos LM de C e MA. As BD construídas serão disponibilizadas para a MD e o conhecimento extraído poderá auxiliar especialistas na tomada de decisão. Como trabalhos futuros propõem-se a aplicação da metodologia em uma casuística maior e a avaliação da redução de tempo provido pelo mapeamento automático das informações.

\*Trabalho apoiado pelo PDTA-FPTI/BR.