

# L'interrogation « à la Google »

Le moteur de recherche préféré des Français a une influence néfaste sur la recherche bibliographique dans PubMed. L'entrée d'un mot clé du langage courant dans la fenêtre d'interrogation de PubMed déclenche souvent une avalanche de notices sans rapport avec la requête (« bruit »). Seul petit avantage : ce type de recherche explore superficiellement les articles non encore indexés dans MEDLINE (intérêt modeste dans la veille bibliographique).

## La recherche (trop) simple

Dans PubMed, la recherche simple utilise la même syntaxe que Google. Les mots clés du langage courant sont entrés les uns à côté des autres et séparés par un espace. Les mots composés sont mis entre guillemets.

Exemple avec une recherche sur les complications de l'ostéoporose. L'équation de recherche s'écrit : osteoporosis complications. Elle affiche près de 15 000 notices le 5 février 2012.

Pour savoir de quelle façon PubMed interprète l'équation, il suffit de regarder dans le cadre « search details » :

```
("osteoporosis, postmenopausal"[MeSH Terms]
OR ("osteoporosis"[All Fields] AND
"postmenopausal"[All Fields]) OR
"postmenopausal osteoporosis"[All Fields]
OR "osteoporosis"[All Fields] OR
"osteoporosis"[MeSH Terms]) AND
("complications"[Subheading] OR
"complications"[All Fields]).
```

En fait, PubMed a recherché tous les équivalents MeSH de osteoporosis et de complications, les a trouvés (c'est l'Automatic Term Mapping, v. encadré), mais a étendu la requête à tout le contenu de la notice (tous les champs).

Note : la recherche en langage MeSH affine le qualificatif complications au descripteur Osteoporosis (Osteoporosis/complications) et affiche 2 fois moins de notices.

## La page de résultat fait office de tableau de bord de PubMed

Dans la page d'accueil de PubMed, il n'y a qu'une grande fenêtre pour les requêtes. Ce n'est qu'après avoir formulé l'équation de

recherche et cliqué sur « search » que s'affiche le véritable « tableau de bord ».

Sous la « grande fenêtre » est placée une barre d'outils (flux RSS de l'équation, sauvegarde dans le coffre-fort My NCBI, fonction Limits et « usine à gaz » avancée). Sous le logo PubMed (en haut et à gauche), l'intitulé « display settings » propose de modifier l'affichage par défaut des résultats qui est : format summary ; 20 notices par page et ordre anté-chronologique.

Dans la colonne de droite et de haut en bas sont placés :

- les filtres avec, par défaut, Free Full Text et Review ;
- les articles en libre accès de PubMed Central ;
- les cadres « search details » qui contiennent l'interprétation de la requête.

## Modeste avantage et gros inconvénient

L'avantage d'une formulation à la Google est mince. Elle couvre très imparfaitement la partie non indexée de Medline\*. En revanche, la partie indexée n'est pas explorée correctement et, là, c'est un gros inconvénient. Il y a 2 raisons à cela : le langage et la syntaxe. L'interrogation en langage courant n'est pas cohérente quand on sait que les articles sont indexés avec les termes d'un thésaurus (MeSH). Quant à la syntaxe, les qualificatifs et les champs sont laissés pour compte, ce qui est dommageable pour la précision des requêtes. Au total : ce qui n'est pas indexé est survolé et ce qui l'est est négligé. •

\* Cette partie est conséquente, les documentalistes de la National Library of Medicine prenant régulièrement du retard dans l'indexation des articles.

## L'AUTOMATIC TERM MAPPING

L'entrée d'un mot clé dans la fenêtre d'interrogation de PubMed déclenche l'action de l'Automatic Term Mapping. Celui-ci confronte successivement le mot clé à des tables de concordance et à des index. Dans l'ordre : la table de concordance des termes du MeSH, puis celle des périodiques, enfin celle des auteurs (nom et prénom depuis 2002) ou l'index des auteurs (nom seul avant 2002). Quand un mot clé est retrouvé dans une des tables ou un des index, le processus de recherche s'arrête.

Si un mot clé est trouvé dans la table de correspondance des termes du MeSH, le mot clé est recherché dans PubMed en tant que terme du MeSH et en tant que mot clé recherché dans tous les champs. Il en est de même pour les titres de périodiques. Il n'y a pas d'interprétation pour les auteurs.

Exemple avec l'équation densitometry methods, interprétée de la façon suivante :

```
("densitometry"[MeSH Terms] OR
"densitometry"[All Fields]) AND
("methods"[Subheading] OR
"methods"[All Fields] OR
"methods"[MeSH Terms])
```

Les défauts de l'Automatic term Mapping sont encore plus graves quand il y a plusieurs mots clés ou un mot clé composé à interpréter. Dans tous les cas, l'Automatic Term Mapping élargit le champ de la requête et perd en précision ce qu'il gagne en exhaustivité. Il fait du « bruit » et il est rare que ce soit la qualité recherchée.