

Affirmations inappropriées	Affirmations correctes
<i>Treatment A may be useful in controlling cancer-related fatigue in patients who present with severe fatigue but is not useful in patients with mild or moderate fatigue.</i>	<i>Treatment A was not more effective than placebo in controlling cancer-related fatigue.</i>
<i>This study confirmed a previous trial demonstrating improved PFS and response for the treatment A compared with comparator B alone, although this did not result in improved survival.</i>	<i>The treatment A was not more effective than comparator B on overall survival in patients with metastatic breast cancer previously treated with anthracycline and taxanes.</i>

Définition du problème

«Spin» (embellissements, affirmations inappropriées) dans les articles biomédicaux: affirmation qu'un effet positif (efficacité ou sûreté) du traitement étudié est plus grand que celui prouvé dans l'essai.

Le spin :

- cause la perception incorrecte de l'efficacité du traitement par les médecins et les patients
- a de l'influence sur les communiqués de presse et les actualités de santé.

Le spin peut être liée:

- aux **résultats** étudiés:
 - omission du résultat principal (primaire)
 - présentation des résultats non-significatifs comme positifs
 - extrapolation : présentation d'un résultat différent/plus général
- à la **population de patients**:
 - focus sur des sous-groupes
 - extrapolation : présentation d'une population plus large
- aux **caractéristiques d'essai**:
 - présentation incorrecte du type d'essai (randomisé/non-randomisé)
- aux effets négatifs

Objectif du projet

- assister les auteurs d'articles scientifiques et leurs relecteurs à identifier les occurrences probables de spin.

Méthodes

Traitement Automatique des Langues (TAL):

- classification automatique;
- extraction des entités et des relations:
 - les résultats de l'essai;
 - la population de patients;
 - les effets négatifs;
 - la significativité statistique.
- analyse des négations et des modalités;
- analyse des paraphrases

Expériences: extraction des entités

Le corpus

3938 articles sur des essais randomisés contrôlés de PMC (PubMed Central).

Le résultat examiné

Le but:

- retrouver les descriptions des résultats initiaux d'un essai en faisant la distinction entre les résultats primaire(s) et secondaire(s) ;
- identifier les résultats présentés dans le résumé de l'article étudié ;
- identifier les cas où les résultats présentés dans le résumé ne correspondraient pas aux résultats initiaux.

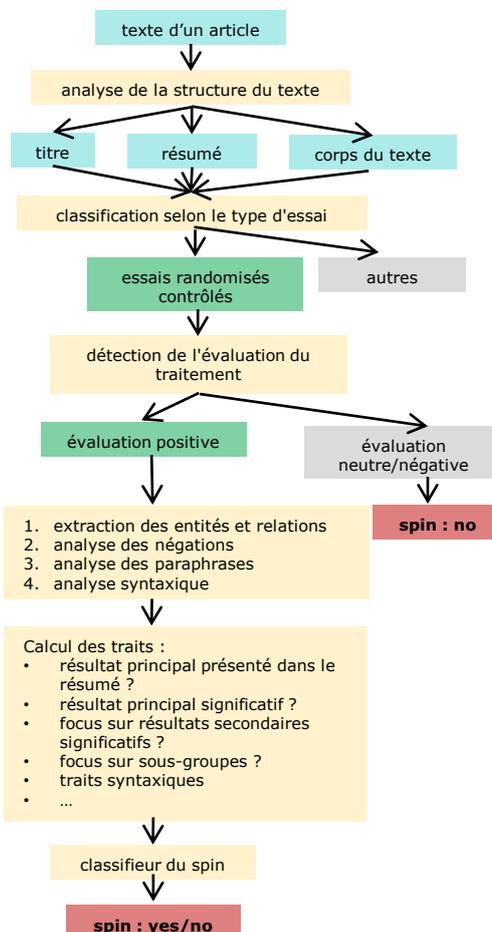


Figure 1: Schéma de l'algorithme

Première étape : évaluer la possibilité d'extraire les résultats des descriptions explicites :

- Expected outcome is...* (« le résultat attendu est ... »)
- ... was our primary outcome.* (« ... était notre résultat principal »)
- Outcomes include...* (« le résultat inclut ... »)

La population de patients

Le but:

- retrouver la description de la population de patients étudiée: âge, sexe, nationalité
- retrouver la description de la population dans les conclusions du résumé ;
- identifier les cas où ces descriptions ne correspondraient pas l'une à l'autre.

Méthodes:

- des règles à base de grammaires locales (la boîte à outils Unitex)
- 9 graphes pour détecter les résultats
- 5 graphes pour détecter les descriptions de la population.

Résultats et perspectives

Les constructions décrivant les résultats et la population couvrent une grande partie de textes :

- 99,9% pour la population ;
- 91,5% et 94% pour les constructions les plus générales concernant le résultats examiné ;
- une mention explicite du résultat principal est trouvée dans 51,8% d'articles ;
- une mention de l'objectif principal (du type « *Our main goal is <...>* », etc.) est présente dans 21,1% des articles.

Nos prochains travaux:

- en l'absence d'une mention explicite du résultat, identification du résultat principal sur la base des mentions de l'objectif principal et des mesures utilisées;
- l'extraction du résultat à partir de références implicites;
- l'association des résultats avec leur significativité statistique;
- la vérification de la correspondance entre les résultats décrits dans le corps d'un article et dans les sections « Résultats » et « Conclusions ».