

## Table des matières

Étape initiale à faire une seule fois	1
1. Préparer la procédure de dédoublonnage	2
2. Appliquer les filtres (au besoin)	2
3. Importer les références	3
4. Gérer les doublons	4
5. Exporter les références	4
Annexe 1 : étapes de gestion des doublons	5



### Attention! Cette procédure s'adresse à des techniciens en documentation ou à des bibliothécaires

Cette technique de gestion systématique des doublons se base sur un article de 2016 paru dans la revue *Journal of the Medical Library Association*<sup>1</sup>. Elle consiste à changer subséquemment les paramètres de détection des doublons pour optimiser leur repérage. Sa force consiste en ce qu'elle permet de repérer un maximum de doublons tout en évitant l'élimination de faux doublons. Elle permet aussi de gagner du temps puisque sa structure en entonnoir implique de vérifier manuellement seulement un sous-ensemble de références. Le champ « pages », en combinaison avec d'autres champs, est utilisé à diverses reprises comme identifiant unique pour repérer les doublons. Or, la manière de présenter la pagination diffère entre les bases de données (ex : 1003-1012 vs 1003-12) et il importe donc d'homogénéiser ce champ avant de procéder au repérage des doublons<sup>2</sup>. La procédure qui suit indique comment modifier le champ « pages » pour les bases de données concernées puis comment procéder au retrait des doublons.

## Étape initiale à faire une seule fois

- Aller à <http://bit.ly/emcendnote> et ouvrir le fichier compressé
- Double-cliquer sur \_Correct Pages.ens (le fichier s'ouvrira dans EndNote)
- Dans EndNote, cliquer sur *File* → *Save As*
- Retirer le mot « copy » du nom du fichier et cliquer sur *Save*
- Fermer le style
- Répéter les étapes 3 à 6 pour le fichier \_Import Corrected Pages.enf

<sup>1</sup> Bramer, W. M., Giustini, D., De Jonge, G. B., Holland, L. et Bekhuis, T. (2016). [De-duplication of database search results for systematic reviews in EndNote](https://doi.org/10.5195/JMLA.2016.24). *Journal of the Medical Library Association*, 104(3). doi:10.5195/JMLA.2016.24

<sup>2</sup> Les filtres (import/export) pour l'homogénéisation du champ « pages » ont été importés une première fois lors de l'installation du logiciel EndNote. Advenant le cas où ces filtres devraient être réinstallés, consulter la page 240 de l'article de Bramer et al.



## 1. Préparer la procédure de dédoublonnage

- Créer un **dossier** en le nommant selon la date du jour (AAAAMMJJ) et le nom de famille de l'utilisateur dans le dossier « dedoublonnage » sur le serveur commun<sup>3</sup> (ex : « 20180308-lacourse »)
- **Importer** dans ce dossier les **fichiers (RIS ou txt)** envoyés par l'utilisateur en renommant au besoin les fichiers selon cette convention de nommage :
  - AAAAMMJJ-plateforme-bd-ref-ref (ex : « 20180308-ebSCO-psycinfo-001-386 »)
- Ouvrir EndNote

## 2. Appliquer les filtres (au besoin)



### Appliquer les filtres pour les bases de données PubMed, MEDLINE (Ovid) et Cochrane Library

Si des fichiers envoyés par l'utilisateur proviennent des bases de données PubMed, MEDLINE (Ovid) ou Cochrane Library, appliquer les filtres pour ajuster le format des pages (nous visons la forme complète, i.e. 1003-1012). Advenant que plus d'une de ces bases de données soient concernées, créer des groupes différents dans EndNote pour ne pas mélanger les références.

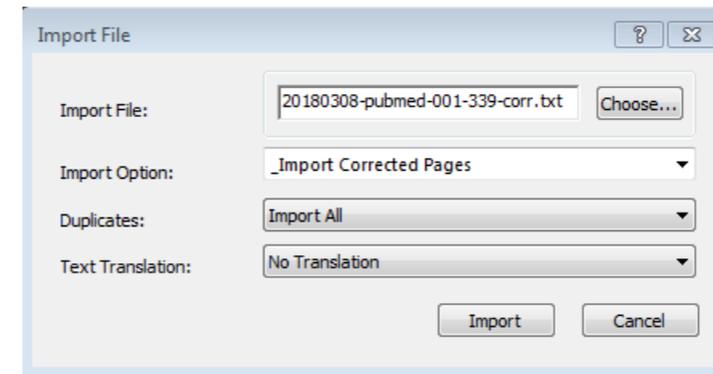
- Créer une **bibliothèque EndNote temporaire** et l'enregistrer dans le dossier de l'utilisateur : *File* → *New...*
  - Reprendre la convention de nommage utilisée pour le dossier de l'utilisateur en ajoutant « -temp » à la fin (ex : « 20180308-lacourse-temp »).
- **Importer** dans EndNote les références de l'une des bases de données concernées (s'il y a plusieurs fichiers pour une même base de données, les importer tous) : *File* → *Import*
  - *Import File* : sélectionner le fichier dans le dossier de l'utilisateur
  - *Import Option* : choisir le filtre d'importation en fonction du type de fichier :
    - Fichier texte (txt) : **Endnote Import**
    - Fichier Research Information System (RIS) : **Reference Manager (RIS)**
  - *Duplicates* : *Import All*
  - *Text Translation* : *No Translation*
- **Exporter** ensuite un fichier avec l'ensemble des références d'une même base de données vers le dossier de l'utilisateur en appliquant le **filtre d'exportation** de correction des pages : *File* → *Export*
  - *Nom du fichier* : ajouter « -corr » à la fin du nom du fichier (ex : « 20180308-ebSCO-psycinfo-001-386-corr ») et choisir comme emplacement le dossier de l'utilisateur
  - *Type* : *Text File (\*.txt)*
  - *Output Style* : « **\_Correct Pages** » : menu déroulant → *Select Another Style* (ce filtre devrait se trouver en tête de liste)
  - Sélectionner *Export Selected References*
- Recommencer au besoin la dernière étape d'exportation pour les autres bases de données dont le champ « pages » doit être ajusté

<sup>3</sup> À l'INSPQ, il s'agit du dossier S:\Communication\documentation\recherche\dedoublonnage

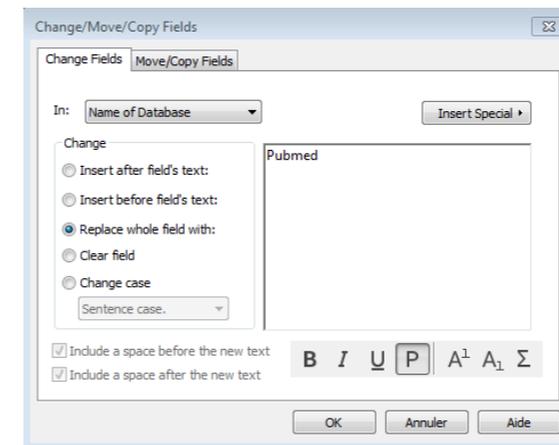


### 3. Importer les références

- Créer une nouvelle bibliothèque EndNote : *File* → *Import*
  - L'enregistrer dans le dossier de l'utilisateur en reprenant la convention de nommage utilisée pour le dossier (ex : « 20180308-lacourse ») ; cette étape créera deux fichiers (avec les extensions .data et .enl)
- Pour chaque base de données, créer un groupe : *My Group* → clic droit → *Create Group*
  - Nommer chaque groupe en fonction de son ordre d'introduction et de son titre :
- Importer les références de la première base de données en fonction de l'ordre précisé par l'utilisateur. Si celui-ci n'a pas défini de préférence, privilégier Medline/PubMed en premier : *File* → *Import*
  - S'il s'agit de l'une des bases de données dont le format de pagination a été modifié, choisir le fichier se terminant par « corr » et appliquer le filtre d'importation « **\_Import Corrected Pages** ». Sinon, choisir le format en fonction du type de fichier :
    - Fichier texte (txt) : *Endnote Import*
    - Fichier Research Information System (RIS) : *Reference Manager (RIS)*



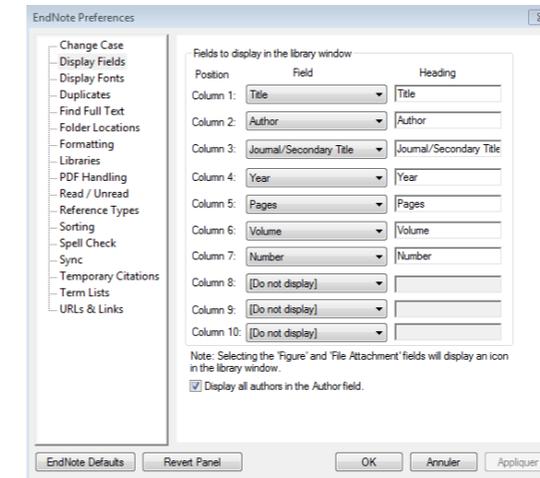
- À chaque importation, déplacer les références dans le groupe correspondant dans EndNote en sélectionnant toutes les références du dossier « Imported References » (ctrl+A) puis en les glissant dans le groupe correspondant.
- Pour chacune des bases de données, s'assurer que le champ « Name of Database »<sup>4</sup> soit rempli. Au besoin, indiquer cette information en procédant ainsi : Ctrl-A pour sélectionner toutes les références du groupe → *Tools* → *Change/Move/Copy Fields*
  - *In* : *Name of Database* (choisir dans le menu déroulant)
  - *Change* : *Replace whole field with*
  - Inscrire le nom de la base de données dans l'encadré
  - OK



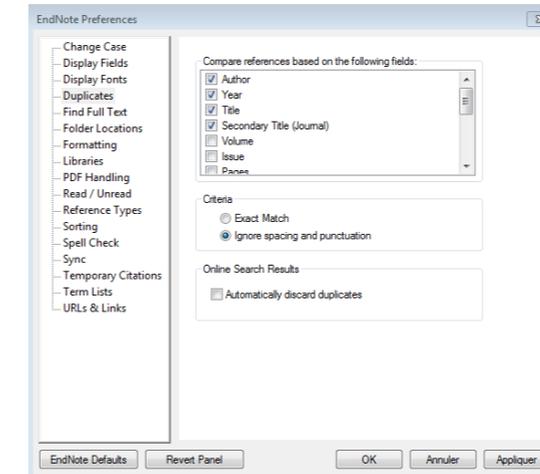
<sup>4</sup> Une fois dans Zotero, le nom de la base de données sera transféré dans le champ « Archive »

## 4. Gérer les doublons

- S'assurer que les champs suivants soient présents et les ajouter au besoin : Title | Author | Secondary Title (Journal) | Year | Pages | Volume | Issue : *Edit* → *Preferences* → *Display Fields*



- Appliquer les préférences d'affichage des champs de la première étape du tableau de l'annexe 1 : *Edit* → *Preferences* → *Duplicates*
- Sélectionner toutes les références de la bibliothèque EndNote (cliquer sur une référence puis Ctrl-A)
  - *References* → *Find duplicates*
  - Cliquer sur *Cancel*
  - Suivre les indications de la colonne « Procédure pour retirer les doublons » du tableau de l'annexe 1 (pour supprimer les doublons : clic droit → *Move References to Trash* ou glisser-déposer vers la corbeille)
- Répéter ces directives pour chacune des étapes du tableau



## 5. Exporter les références

- Dans le dossier de l'utilisateur, créer un sous-dossier et l'intituler « livrable »
- Créer un fichier pour les références uniques en sélectionnant toutes les références restantes dans la bibliothèque : *File* → *Export*
  - *Nom du fichier* : AAAAMMJJ-nomusager-sin-doublons (ex : 20180308-lacourse-sin-doublons)
  - *Type* : *Text File (\*.txt)*

- *Output style : RefMan (RIS) Export*
- Sélectionner *Export Selected References*
- Enregistrer le fichier dans le dossier « livrable »
- Créer un fichier pour les doublons en sélectionnant toutes les références contenues dans la corbeille : *File → Export*
  - Nom du fichier : AAAAMMJJ-nomusager-doublons (ex : 20180308-lacourse-doublons)
  - Suivre les directives de l'étape précédente pour exporter le fichier

## Annexe 1 : étapes de gestion des doublons

Étapes	Paramètres de dédoublonnage	Procédures pour retirer les doublons	Notes
1	Author   Year   Title   Secondary Title (Journal)	Lancer le repérage des doublons. Supprimer les doublons sélectionnés.	Cette étape est sécuritaire et ne requiert pas de contre-vérification.
2	Author   Year   Title   Pages	Lancer le repérage des doublons. Supprimer les doublons sélectionnés.	Cette étape est sécuritaire et ne requiert pas de contre-vérification.
3	Title   Volume   Pages	Lancer le repérage des doublons. Classifier les références par <b>titre</b> et vérifier manuellement les références sans titre à la recherche de faux-doublons ; supprimer les doublons restants.	Comparer à l'aide du champ « auteur » : si les auteurs diffèrent, c'est qu'il s'agit de faux-doublons; désélectionner alors ces références.
		Classifier les références par <b>pages</b> et vérifier manuellement les références <b>sans pagination ou ayant la valeur 1</b> à la recherche de faux-doublons ; supprimer les doublons restants.	Comparer à l'aide du champ « auteur » : si les auteurs diffèrent, c'est qu'il s'agit de faux-doublons; désélectionner alors ces références.
4	Author   Volume   Pages	Lancer le repérage des doublons. Classifier les références par <b>auteur</b> et vérifier manuellement les références sans auteur à la recherche de faux-doublons ; supprimer les doublons restants.	Comparer à l'aide du champ « titre » : si les titres diffèrent, c'est qu'il s'agit de faux-doublons; désélectionner alors ces références.
		Classifier les références par <b>page</b> et vérifier manuellement les références <b>sans pagination ou ayant la valeur 1</b> à la recherche de faux-doublons ; supprimer les doublons restants.	Comparer à l'aide du champ « titre » : si les titres diffèrent, c'est qu'il s'agit de faux-doublons; désélectionner alors ces références.
5	Year   Volume   Issue   Pages	Créer un nouveau groupe ( <i>My Groups → Create Group → Enter</i> ). Lancer le repérage des doublons.	
		Classifier les références par <b>page</b> puis sélectionner toutes celles dont le champ « pages » est rempli puis les déplacer dans le nouveau groupe.  Cliquer sur ce nouveau groupe ( <i>New Group</i> ) et vérifier manuellement les références dont le champ « <b>pages</b> » correspond à 1, commence par 1 ou par une lettre. Sélectionner ces faux-doublons puis <b>les supprimer du groupe</b> . Lancer le repérage des doublons <b>dans ce groupe</b> puis supprimer les doublons sélectionnés automatiquement.	Cliquer sur la première référence contenant une pagination → cliquer en même temps sur Shift et sur la dernière référence contenant une pagination
6	Title	Lancer le repérage des doublons. Vérifier manuellement les références à la recherche de faux-doublons ; supprimer les doublons restants.	Comparer à l'aide du champ « pages » : si la valeur de ces champs diffère, équivaut à 1 ou est absente, examiner les auteurs et le titre de la revue.
7	Author   Year	Lancer le repérage des doublons. Ne pas vérifier les références sans auteur ou dont la valeur du champ auteur est « anonymous ». Si un <b>vrai doublon</b> est repéré, cliquer sur ce dernier de manière à désélectionner les autres références. Comparer les références suivantes à la recherche de vrais doublons. Supprimer les doublons ainsi trouvés. Si aucun doublon n'est repéré, ne jeter aucune référence à la corbeille.	Comparer à l'aide du champ « pages » : S'il est vide ou si sa valeur est 1, examiner plutôt le titre de la revue.