



# CARTOGRAPHIES

## RAPPORT DU PROJET

2020 / 2021

PRÉPARÉE PAR  
HAJAR LACHHEB

## I. Database utilisée

### 1. Mise en contexte

Un **hôpital militaire** est un hôpital sous tutelle d'une armée ou d'un ministère des armées et dont la vocation première est de soigner les militaires et anciens militaires.

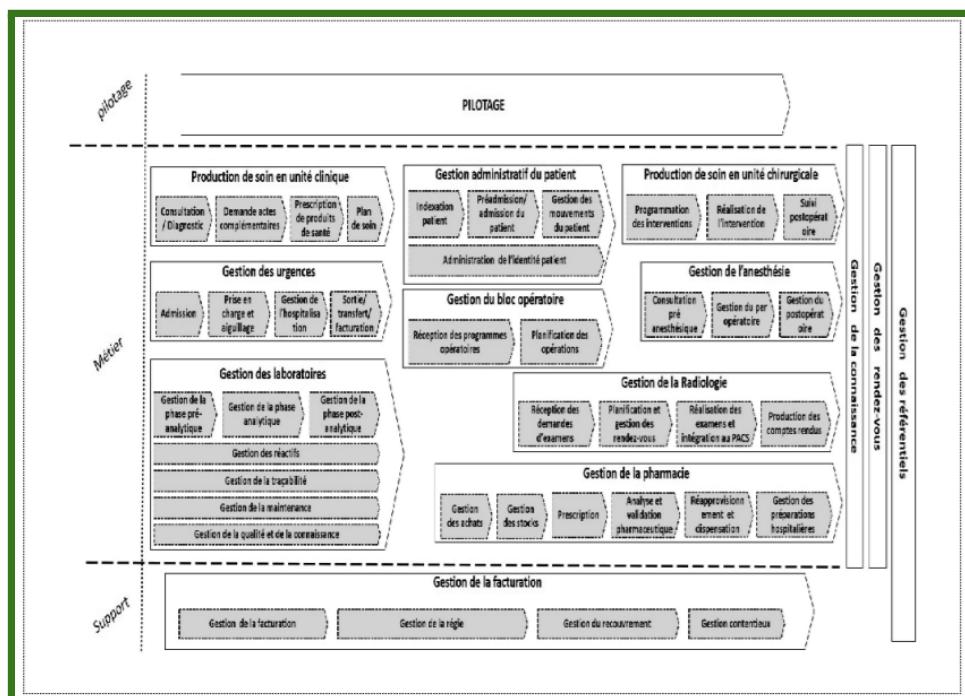
Ils sont établis en ville ou sur une base militaire et aussi à proximité du champ de bataille. Les hôpitaux de marine sont généralement situés dans des villes côtières. Des hôpitaux mobiles sont aménagés sur les navires-hôpitaux en mer.

Durant le Covid19, l'hôpital militaire a fait preuve d'efforts et de compétences hors du commun. Le Maroc était aussi stratégique par rapport au plan qu'il a développé afin de faire face à la pandémie, reste à savoir si ce dernier encourage la production scientifique dans le domaine médical.

La recherche et l'effort scientifique vont sûrement aider le pays à se développer et développer son domaine de santé. Ceci va l'aider aussi à surmonter les futures crises sanitaires et être prêt à faire face à toute future pandémie.

### 2. Motivation du choix

Parmi les raisons qui m'ont motivée à choisir la base de données Hopital Militaire est la cartographie de ce dernier. Il présente une distribution bel et bien faite des différents départements et cellules au sein de l'hôpital. Ceci nous pousse ainsi à déduire que cette distribution pertinente pourrait aider les docteurs au sein de ce dernier à produire plus et plus de papiers scientifiques.



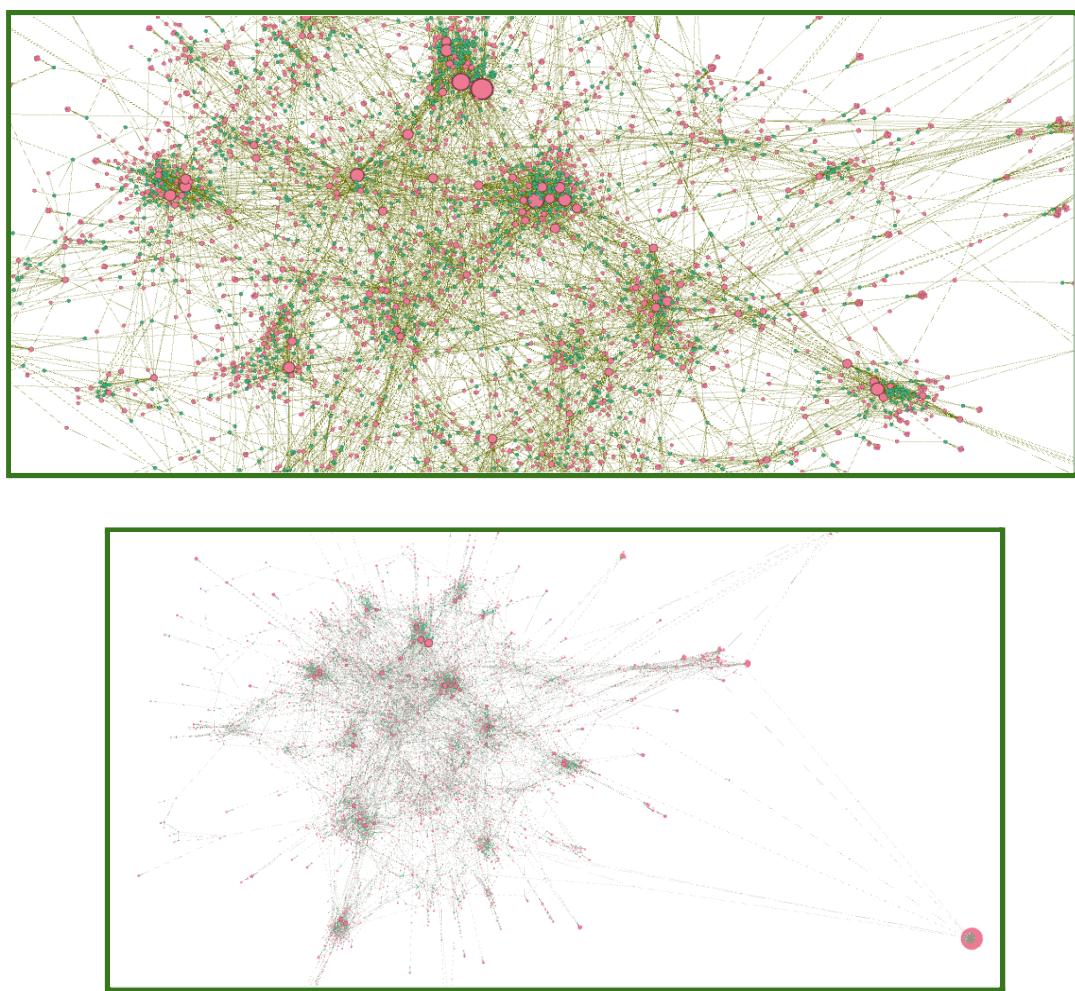
La production scientifique pourrait alors être liée à d'autres indicateurs, non seulement la bonne répartition des cellules et l'organisation interne de l'hôpital. C'est ce qu'on va voir au sein de notre travail.

## II. La liste des cartes réalisées

### 1. Carte des collaborations entre les institutions

Pour la carte des collaborations entre les institutions, on a utilisé deux couleurs pour représenter la distribution des auteurs par rapport aux affiliations.

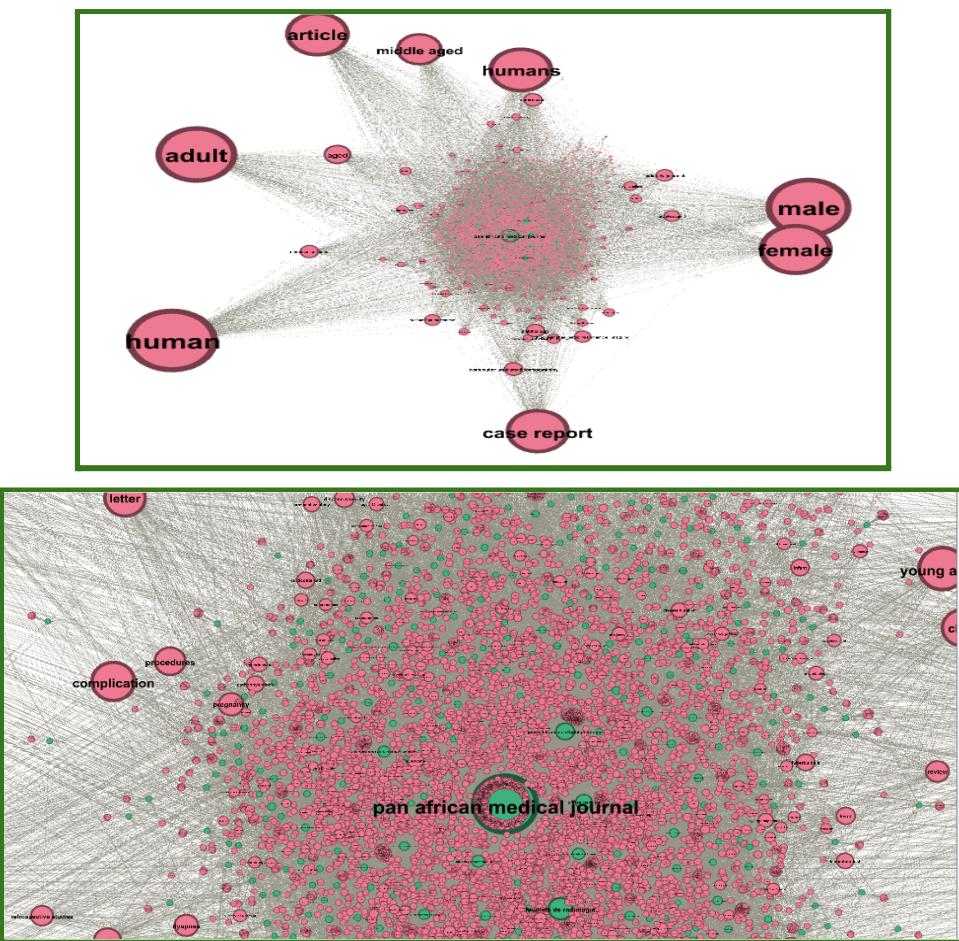
Le rose représente les auteurs alors que le vert représente les affiliations.



### 2. Thématiques de recherche

Pour la carte des thématiques de recherche, on a utilisé deux couleurs pour représenter la distribution des thématiques par rapport aux titres des revues ou les articles sont réalisés.

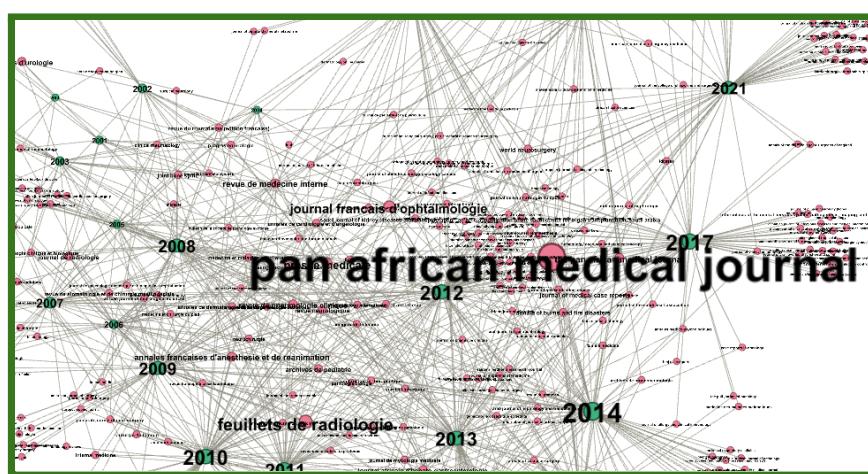
Le rose représente les principales thématiques alors que le vert représente les titres des revues.



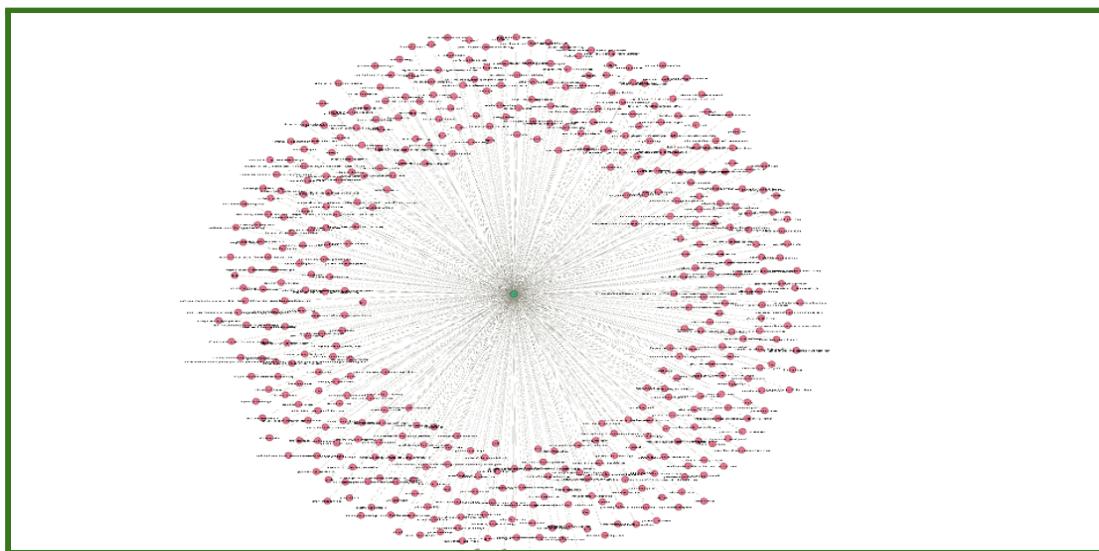
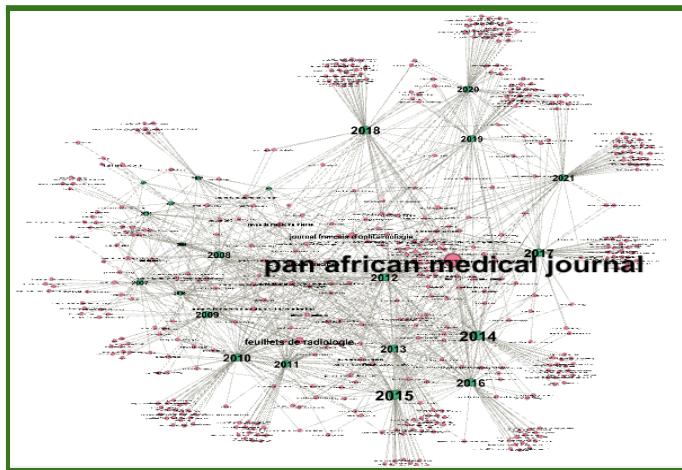
### 3. Revues de publication

Pour la carte des revues de publication, on a utilisé deux couleurs pour représenter la distribution des revues de publication par rapport aux années.

Le rose représente les principales thématiques alors que le vert représente les années de production d'articles.



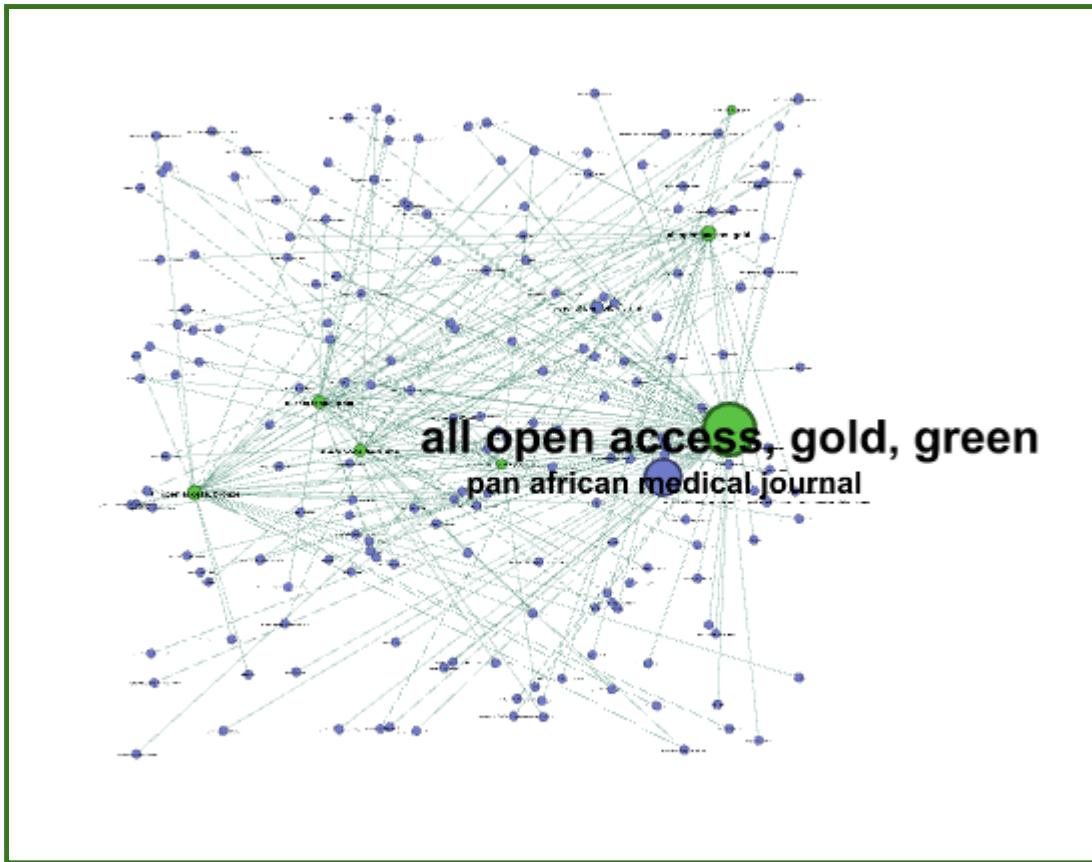
→ Pan African Medical journal est de loin la revue la plus utilisée par les chercheurs afin d'y publier les articles. La publication sur cette revue revient peut être au fait qu'il s'agit d'une revue africaine.



#### 4. Publications en open access

Pour la carte des publications en open access, on a utilisé deux couleurs pour représenter la distribution des publications en Open Access par rapport aux titres des revues.

Le vert représente le type du Open Access alors que le bleu représente les titres des revues ou les articles sont publiés.

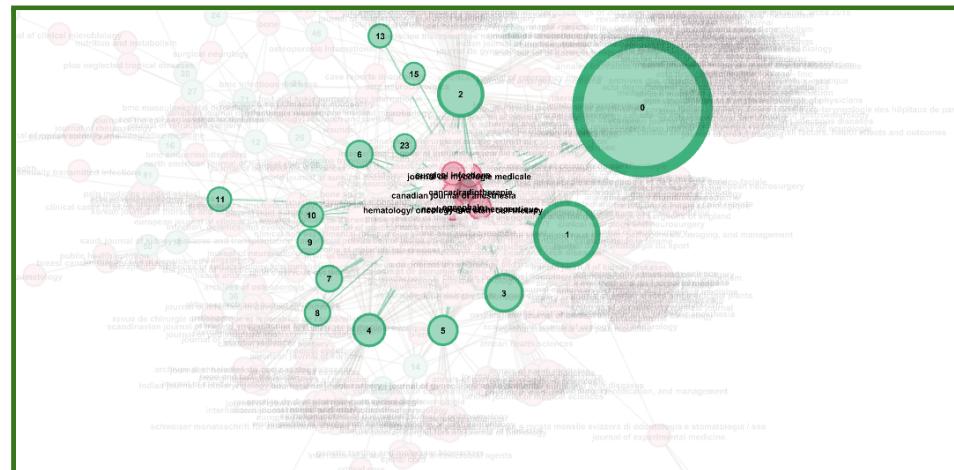
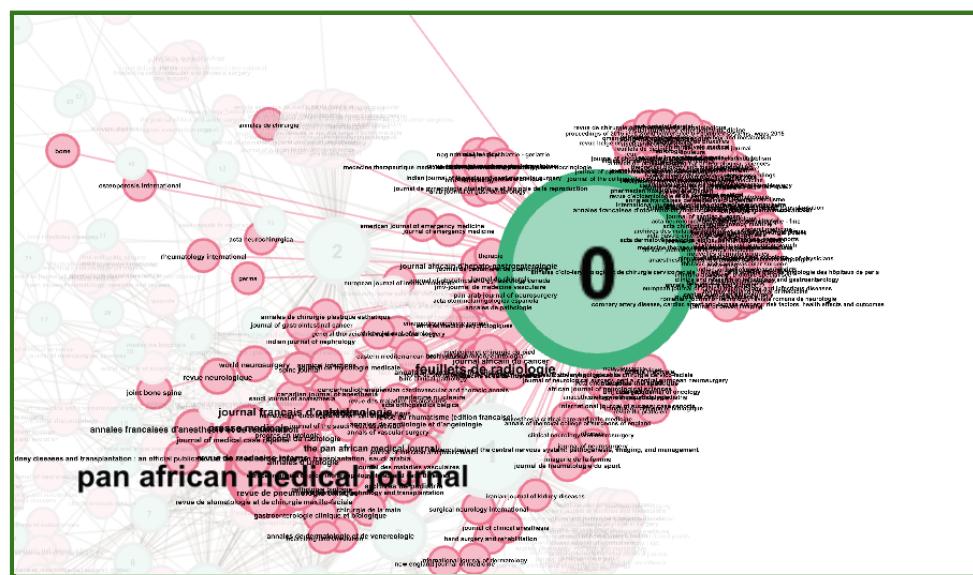
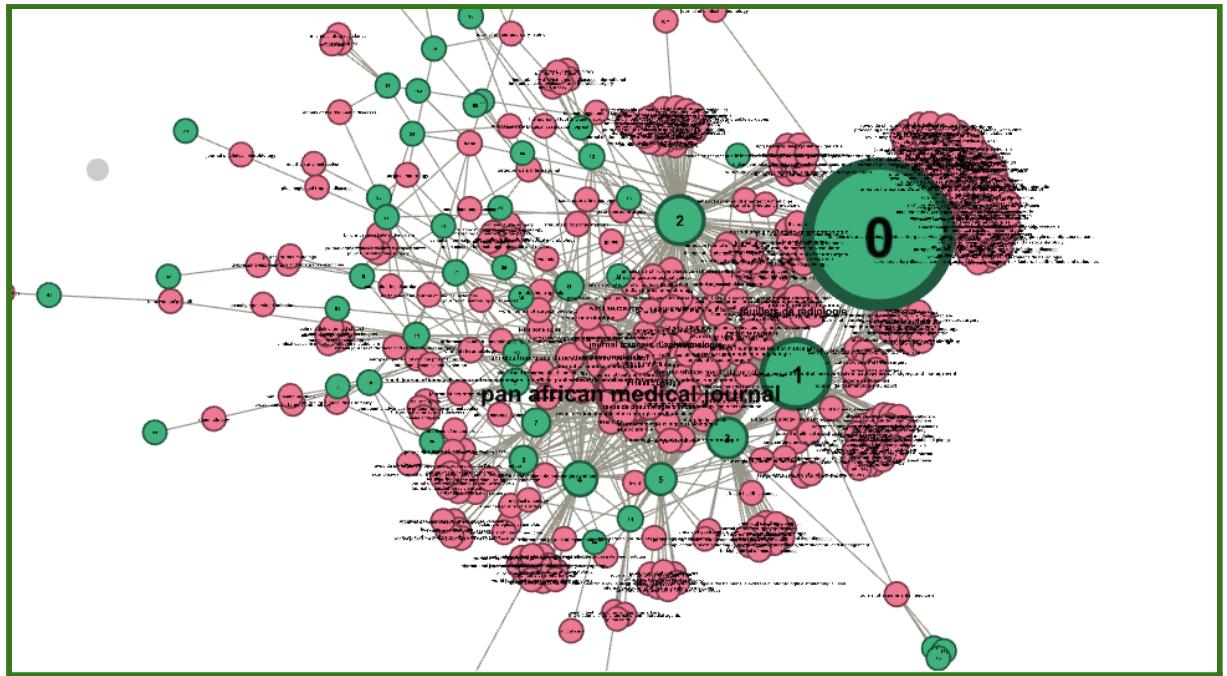


→ [La plupart des publications et articles produits sont des articles qui entrent dans la catégorie du Open Access, gold and green.](#)

## 5. Publications les plus citées

Pour la carte des publications les plus citées, on a utilisé deux couleurs pour représenter la distribution des publications les plus citées par rapport aux titres des revues.

Le vert représente le nombre de citations alors que le rose représente les titres des revues ou les articles sont publiés.

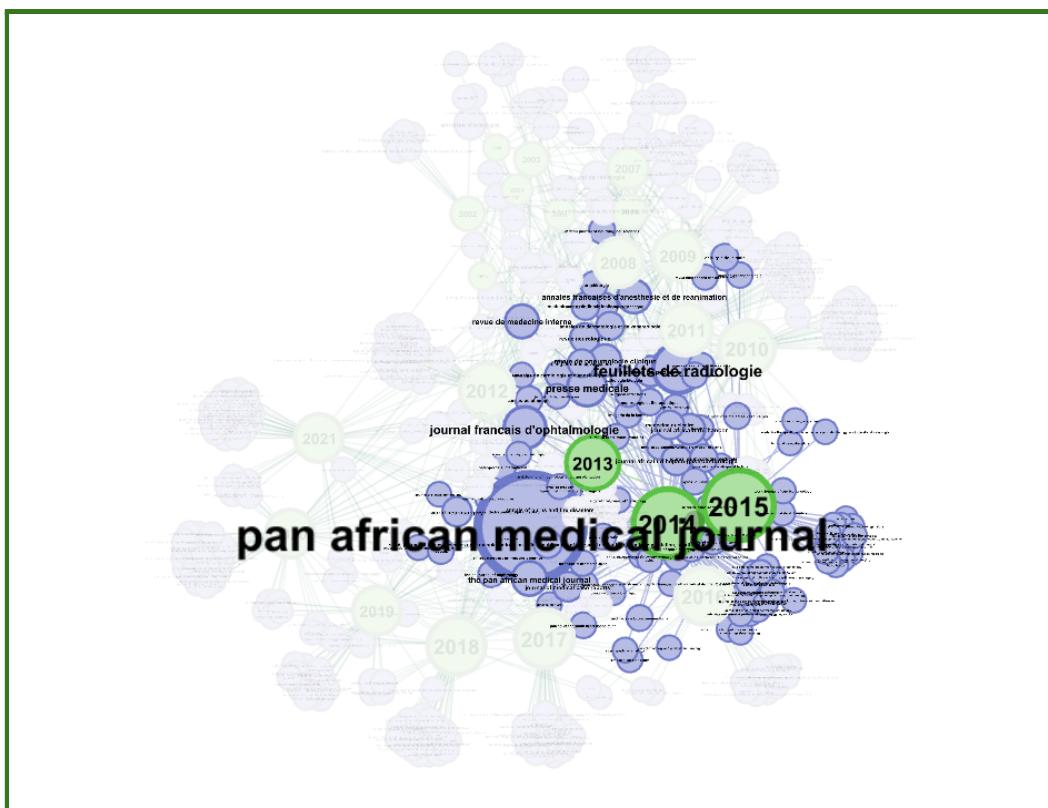


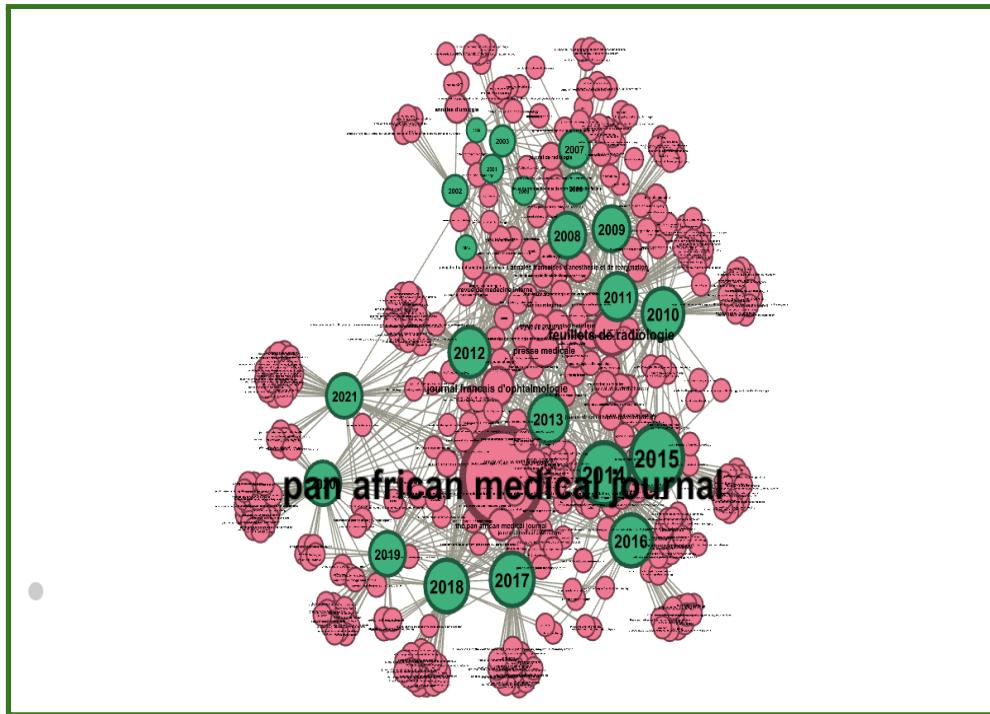
- **Le nombre de citations le plus répandu est un nombre de citations qui atteint le 0. Le nombre de citations dépasse les 200 pour peu d'articles.**

## 6. Évolution de la production scientifique dans le domaine médical par année (évolution dynamique dans gephi)

Pour la carte qui représente l'évolution de la production scientifique par rapport aux années, on a utilisé deux couleurs pour représenter la distribution des publications par rapport aux années de publication.

Le vert représente les années alors que le rose/bleu représente les titres des revues ou les articles sont publiés.



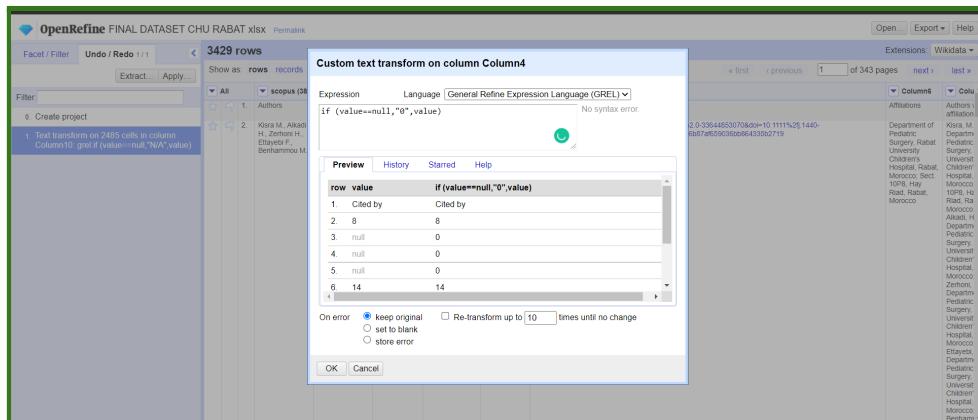


- Le développement de la production scientifique par rapport aux années est très intéressant quoique pour les années 2019, 2020 et 2021 ça a diminué.
- La diminution pourrait être dû à la pandémie de Covid19 qui a réduit la production scientifique vu l'incapacité des docteurs et chercheurs de se focaliser sur leur recherches. Le Covid19 nécessitait les efforts de tout le staff médical.
- La diminution pourrait être aussi dû au fait qu'il n'y a pas eu une bonne collecte de données au niveau de ces années.

### III. La liste des cartes réalisées

#### 1. La relation entre Auteurs et Affiliations :

##### - Procédure détaillée de l'utilisation de Open Refine



On a commencé par utiliser la fonction suivante afin de remplir les valeurs vides du nombre de citations par 0. La fonction est la suivante : **if (value==null,"0",value)**.

The screenshot shows the OpenRefine interface with a project titled "FINAL DATASET CHU RABAT.xlsx". A custom text transform dialog is open for column "Column10". The expression is set to "if (value==null,"N/A",value)". The preview pane shows the first six rows, where row 1 has "Open Access" and rows 2-6 have "N/A". Below the preview, there are options for handling errors: "keep original" (selected), "Re-transform up to [10] times until no change", "set to blank", and "store error".

On a ensuite utilisé la fonction suivante afin de remplir les valeurs non existantes de la partie Open Access par N/A. La fonction est la suivante : **if (value==null,"N/A",value)**.

The screenshot shows the OpenRefine interface with a project titled "FINAL DATASET CHU RABAT.xlsx". A split column dialog is open for column "Column1". The "by separator" option is selected, with a colon ":" entered in the "Separator" field. The "Remove this column" checkbox is checked. The preview pane shows the first few rows of the dataset.

Ensuite, on a utilisé la fonction split afin de diviser les titres en titre en Anglais et en Français. On utilise un séparateur “[“.

The screenshot shows the OpenRefine interface with a project titled "FINAL DATASET CHU RABAT.xlsx". A custom text transform dialog is open for column "Column2". The expression is set to "value.replace("[", "")". The preview pane shows the first three rows, where the "[" character is removed from the titles. Below the preview, there are options for handling errors: "keep original" (selected), "Re-transform up to [10] times until no change", "set to blank", and "store error".

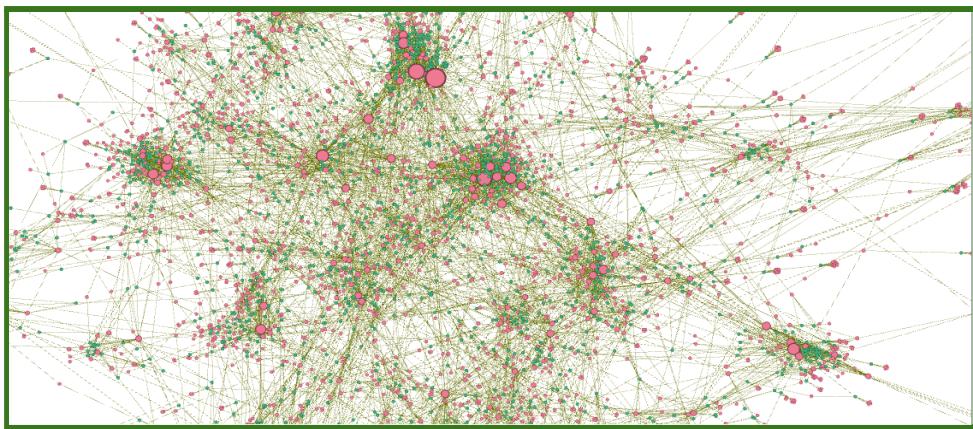
Ensuite, on a utilisé la fonction replace afin de diviser les titres en titre en Anglais et en Français. On remplace le “[“ par ””.

On a ensuite utilisé la fonction suivante afin de remplir les valeurs non existantes de la partie Titres par N/A. La fonction est la suivante : **if (value==null,"N/A",value)**.

### - Procédure détaillée de l'utilisation de Gephi

On a commencé par utiliser Table2Net, par importer la CSV table de l'hôpital militaire. On choisit nos deux nodes et on télécharge le fichier adéquat et homogène avec GEPHI.

On a importé le fichier et utilisé le Layout “**Force Atlas 2**”. On a modifié la couleur et la taille des noeuds ainsi que la taille des textes.

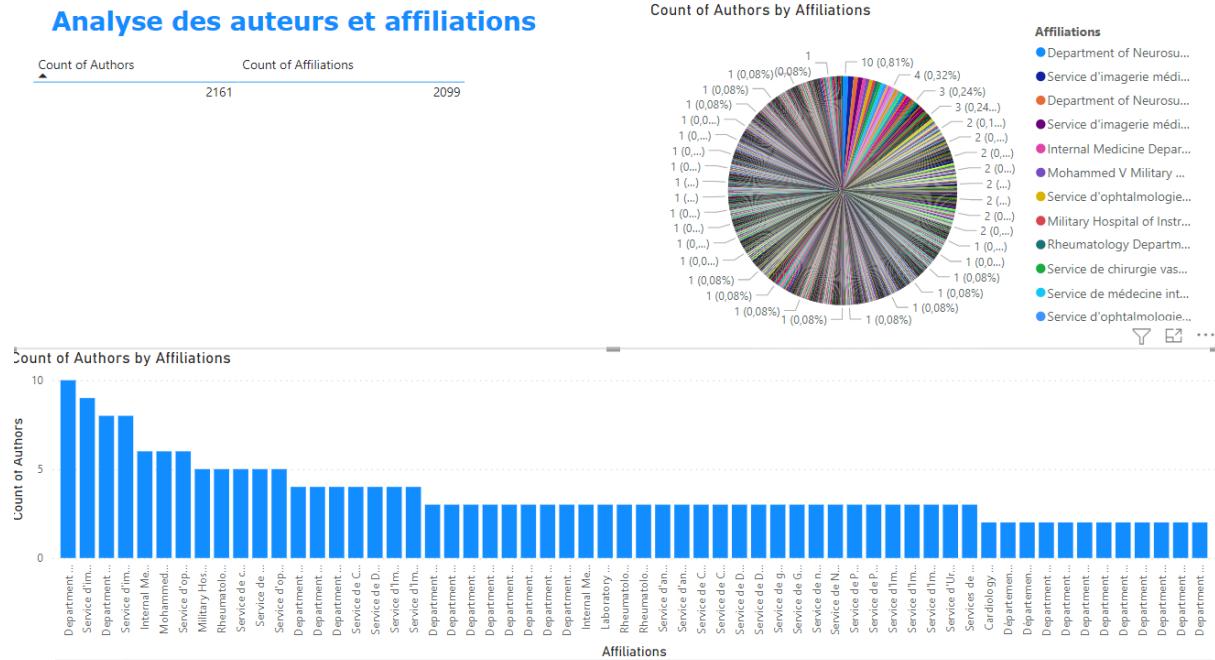


#### - Analyse des données des auteurs et affiliations

L'analyse des données est faite en utilisant Power BI. Microsoft Power BI est une solution d'analyse de données de Microsoft. Il permet de créer des visualisations de données personnalisées et interactives avec une interface suffisamment simple pour que les utilisateurs finaux créent leurs propres rapports et tableaux de bord

On pourrait déduire d'après les graphes ci dessous:

- La majorité des auteurs font partie du département de neurochirurgie. Ensuite les auteurs appartiennent au service d'imagerie médicale.
  - Un grand nombre d'affiliations créent une diversité au niveau de la publication des articles.
  - La plupart des articles les plus citées ont des auteurs appartenant à plus d'une affiliation.



## **2. La relation entre les thématiques et les titres de revues :**

- ## - Procédure détaillée de l'utilisation de Open Refine

Custom text transform on column Column4

Expression: if (value==null,"0",value)

Language: General Refine Expression Language (GREL)

Preview:

row	value
1.	Cited by
2.	8
3.	null
4.	null
5.	null
6.	14

On error:  keep original  Re-transform up to [10] times until no change  
 set to blank  store error

OK Cancel

On a commencé par utiliser la fonction suivante afin de remplir les valeurs vides du nombre de citations par 0. La fonction est la suivante : **if (value==null,"0",value)**.

Custom text transform on column Column10

Expression: if (value==null,"N/A",value)

Language: General Refine Expression Language (GREL)

Preview:

row	value
1.	Open Access
2.	null
3.	null
4.	null
5.	null
6.	null

On error:  keep original  Re-transform up to [10] times until no change  
 set to blank  store error

OK Cancel

On a ensuite utilisé la fonction suivante afin de remplir les valeurs non existantes de la partie Open Access par N/A. La fonction est la suivante : **if (value==null,"N/A",value)**.

Split column Column 1 into several columns

How to Split Column

by separator  
Separator: |  regular expression

by field lengths

After Splitting

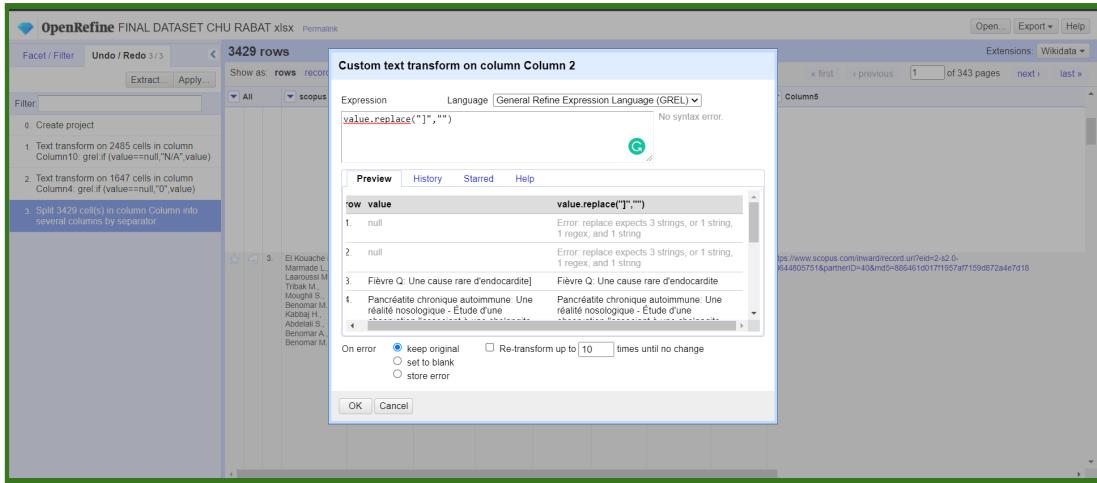
Guess type  Remove this column

Split into [ ] columns at most (leave blank for no limit)

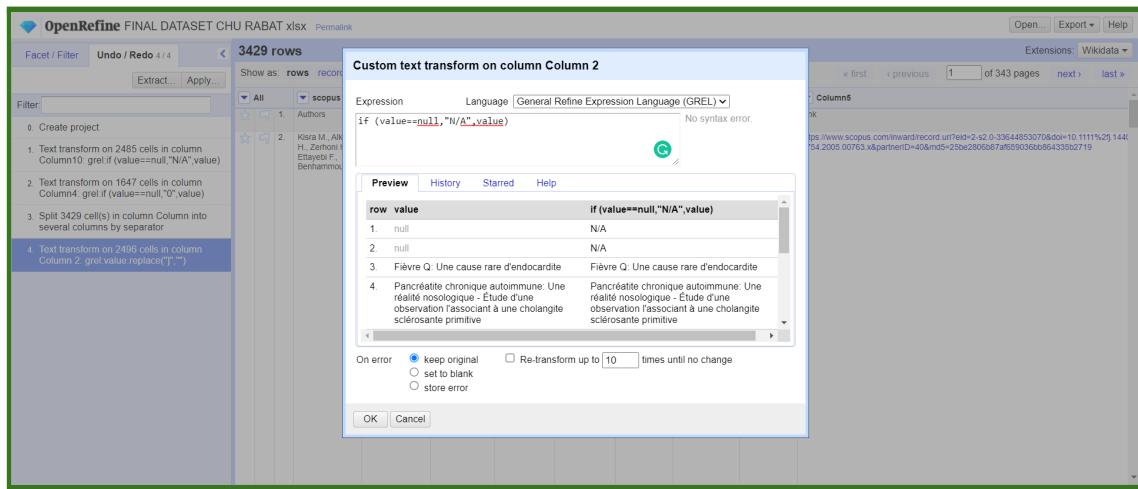
List of integers separated by commas, e.g., 5, 7, 15

OK Cancel

Ensuite, on a utilisé la fonction split afin de diviser les titres en titre en Anglais et en Français. On utilise un séparateur “|“.



Ensuite, on a utilisé la fonction replace afin de diviser les titres en titre en Anglais et en Français. On remplace le "[" par " ".



On a ensuite utilisé la fonction suivante afin de remplir les valeurs non existantes de la partie Titres par N/A. La fonction est la suivante : **if (value==null,"N/A",value)**.

### - Procédure détaillée de l'utilisation de Gephi

On a commencé par utiliser Table2Net, par importer la CSV table de l'hôpital militaire. On choisit nos deux nodes et on télécharge le fichier adéquat et homogène avec GEPHI.

Table 2 Net



# Table 2 Net

Extract a network from a table. Set a column for nodes and a column for edges. It deals with multiple items per cell.

**Load your CSV table**

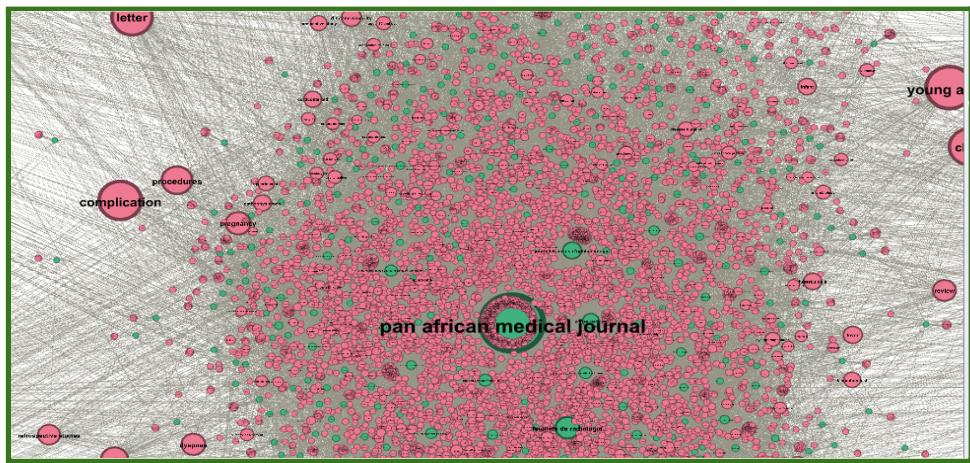
It has to be **comma-separated** and the first row must be dedicated to **column names**.

Parsing successful. 13 columns and 2334 rows.

**Table preview**

Row number	Authors	Title	Year	Source title	Cited by	Link
1	Tagajid M.R., El Annaz H., Belafquih B., Dobail T., Casalegno J.S., Mekki Y., Mrani S.	Factors influencing uptake of influenza vaccine amongst healthcare workers in a regional center after the A(H1N1) 2009 pandemic:	2011	International Journal of Risk and Safety in Medicine	5	<a href="https://www.scopus.com/inward/record.uri?eid=2-s2.0-84255182404&amp;dol=10.3233%2fIJS-2011-0544&amp;partnerID=40&amp;md5=e63433ba3912e298cf5d747eab47c1c">https://www.scopus.com/inward/record.uri?eid=2-s2.0-84255182404&amp;dol=10.3233%2fIJS-2011-0544&amp;partnerID=40&amp;md5=e63433ba3912e298cf5d747eab47c1c</a>

On a importé le fichier et utilisé le Layout “**Force Atlas 2**”. On a modifié la couleur et la taille des noeuds ainsi que la taille des textes.



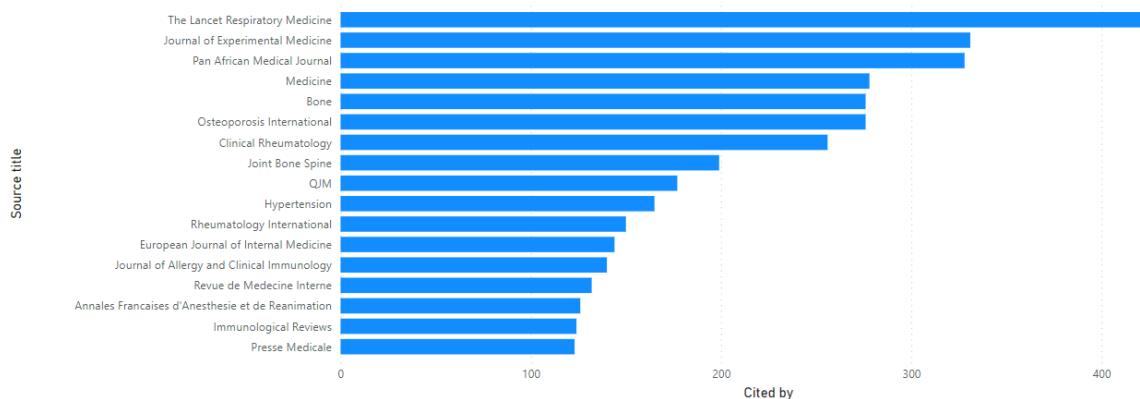
### - Analyse des données des auteurs et affiliations

L'analyse des données est faite en utilisant Power BI. Microsoft Power BI est une solution d'analyse de données de Microsoft. Il permet de créer des visualisations de données personnalisées et interactives avec une interface suffisamment simple pour que les utilisateurs finaux créent leurs propres rapports et tableaux de bord

On pourrait déduire d'après les graphes ci dessous:

- Les mots clés sont non significatifs par rapport au contenu des articles produits. En analysant les mots clés, on n'arrive pas à connaître le vrai sujet ou thématique du sujet en question.
- On finit par analyser le nombre de revues ou on produit le plus d'articles, ça donne plus ou moins une idée sur la thématique principale de chaque article scientifique. Si on compare par l'abondance des revues, on remarque que les plus grandes revues ou on publie le plus d'articles sont des revues dont la thématique principale est : Respiratoire, Maladie respiratoire et Médecine Expérimentale.

#### Cited by by Source title



#### Cited by by Source title

