

# Creating a Comprehensive Medical Journal Abbreviation List

David Csuka

2022-08-11

The current medical journal abbreviation lists available for bibliography managers such as EndNote do not contain information from PubMed, which is very important for the ease and convenience of academic citations. I created a comprehensive database from both the PubMed terms as well as those already available in EndNote.

```
library(tidyverse)
library(tokenizers)

pubmed_journals <- read_delim("https://ftp.ncbi.nih.gov/pubmed/J_Medline.txt",
                              delim = "\\n", col_names = "Information")
```

```
pubmed_journals
```

```
## # A tibble: 281,456 x 1
##   Information
##   <chr>
## 1 "-----"
## 2 "JrId: 1"
## 3 "JournalTitle: AAE editors' journal"
## 4 "MedAbbr: AAE Ed J"
## 5 "ISSN (Print): 0160-6999"
## 6 "ISSN (Online): "
## 7 "IsoAbbr: AAE Ed J"
## 8 "NlmId: 7708172"
## 9 "-----"
## 10 "JrId: 2"
## # ... with 281,446 more rows
```

```
journal_abbreviations <- pubmed_journals %>%
  filter(str_detect(Information, "[:alnum:]")) %>%
  separate(Information, into = c("field", "info"), sep = ":\s?",
           extra = "merge") %>%
  mutate(id = cumsum(field == "JrId")) %>%
  pivot_wider(names_from = field, values_from = info) %>%
  select(-id, -JrId, -IsoAbbr) #IsoAbbr is always equal to MedAbbr
```

```
write_excel_csv(journal_abbreviations, "Just_PubMed_Raw.csv")
```

```
journal_abbreviations
```

```
## # A tibble: 35,182 x 5
```

```
##      JournalTitle
##      <chr>
## 1 AADE editors' journal
## 2 AANA journal
## 3 AARN news letter
## 4 [Nihon shika hyoron] The Nippon dental review
## 5 Academia Peruana de Cirugia
## 6 Academy review of the California Academy of Periodontology, United States Section, ARPA Internati
## 7 Acarologia
## 8 Acquisitions medicales recentes
## 9 Activitas nervosa superior
## 10 Acta allergologica
## # ... with 35,172 more rows
```

```
new_nlm_terms <- journal_abbreviations %>%
  mutate(split_journal_title = tokenize_words(JournalTitle),
         split_medabbr = tokenize_words(MedAbbr)) %>%
  rowwise() %>%
  mutate(Abbreviation_1 = map_chr(split_medabbr,
                                ~if_else(. %in% split_journal_title,
                                           .,
                                           paste0(., ".")))) %>%
  paste(collapse = " ") %>%
  str_to_title() %>%

  ungroup() %>%
  select(JournalTitle, Abbreviation_1, MedAbbr) %>%
  rename(Journal_Name = JournalTitle,
         Abbreviation_2 = MedAbbr)

write_tsv(new_nlm_terms, "Just_PubMed.txt", col_names = FALSE)

new_nlm_terms
```

```
## # A tibble: 35,182 x 3
##      Journal_Name
##      <chr>
## 1 AADE editors' journal
## 2 AANA journal
## 3 AARN news letter
## 4 [Nihon shika hyoron] The Nippon dental review
## 5 Academia Peruana de Cirugia
## 6 Academy review of the California Academy of Periodontology, United States Section, ARPA Internati
## 7 Acarologia
## 8 Acquisitions medicales recentes
## 9 Activitas nervosa superior
## 10 Acta allergologica
## # ... with 35,172 more rows
```

```
old_endnote_terms <- map(list.files("Endnote_Journal_Files", full.names = TRUE),
                        ~read_tsv(., col_names = c("Journal_Name",
                                                  "Abbreviation_1",
                                                  "Abbreviation_2"))) %>%
  select(!starts_with("X"))
```

```
old_endnote_terms
```

```
## [[1]]
## # A tibble: 6,259 x 3
##   Journal_Name                      Abbreviation_1
##   <chr>                      <chr>
## 1 A Trends Guide                Trends Guide
## 2 AAOHN Journal                AAOHN J.
## 3 Aarsberet. Nor. Fisk.        Aarsberet. Nor. Fisk.
## 4 Aarsberetning. NORCONS-RV    Aarsberet. NORCONS-RV
## 5 Aarsberetning. Norsk Sjoefartsmuseum Aarsberet. Norsk Sjoefartsmus
## 6 Aarsberetning. Sildolje- og Sildemelindustriens Forskningsinstitutt Aarsberet. SSF
## 7 Aarsberetning. Statens Kartverk Aarsberet. Statens Kartverk
## 8 Aarsmelding. Eksportutvalget for Fisk Aarsmelding Eksport. Fisk
## 9 Aarsmelding. Fiskeriforskning Aarsmelding Fiskeriforsk.
## 10 Aarsmelding. Havforskningsinstituttet Aarsmelding. Havforskningsinst
## # ... with 6,249 more rows
##
## [[2]]
## # A tibble: 227 x 3
##   Journal_Name                      Abbreviation_1      Abbreviation_2
##   <chr>                      <chr>          <chr>
## 1 Allergy, Asthma & Immunology Research Allergy Asthma Immunol. Res. Allergy Asthma Immunol
## 2 Anesthesia and Pain Medicine          Anesth. Pain Med.      Anesth Pain Med
## 3 Annals of Dermatology                 Ann. Dermatol.         Ann Dermatol
## 4 Asian Spine Journal                  Asian Spine J.          Asian Spine J
## 5 Basic and Applied pathology           Basic Appl. Pathol.     Basic Appl Pathol
## 6 Cancer Research and Treatment         Cancer Res. Treat.      Cancer Res Treat
## 7 Chonnam Medical Journal              Chonnam Med. J.        Chonnam Med J
## 8 Chung-Ang Journal of Medicine         Chung-Ang J. Med.       Chung-Ang J Med
## 9 Chungbuk Medical Journal             Chungbuk Med. J.        Chungbuk Med J
## 10 Clinical and Experimental Otorhinolaryngology Clin. Exp. Otorhinolaryngol. Clin Exp Otorhinolaryng
## # ... with 217 more rows
##
## [[3]]
## # A tibble: 14,504 x 3
##   Journal_Name                      Abbreviation_1      Abbreviation_2      Abbreviation_3
##   <chr>                      <chr>          <chr>          <chr>
## 1 A Pediatria                Pediatria (Rio)    Pediatria (Rio)    Pediatria (Rio)
## 2 A.M.A. American Journal of Diseases of Children AMA Am. J. Dis. Child. AMA Am. J. Dis. Child. AMA Am. J. Dis. Child.
## 3 A.M.A. Archives of Dermatology AMA Arch. Derm.    AMA Arch. Derm.    AMA Arch. Derm.
## 4 A.M.A. Archives of Dermatology and Syphilology AMA Arch. Derm. Syphilol. AMA Arch. Derm. Syphilol. AMA Arch. Derm. Syphilol.
## 5 A.M.A. Archives of General Psychiatry AMA Arch. Gen. Psychiatry AMA Arch. Gen. Psychiatry AMA Arch. Gen. Psychiatry
## 6 A.M.A. Archives of Industrial Health AMA Arch. Ind. Health AMA Arch. Ind. Health AMA Arch. Ind. Health
## 7 A.M.A. Archives of Industrial Hygiene and Occupational Medicine AMA Arch. Ind. Hyg. Occup. Med. AMA Arch. Ind. Hyg. Occup. Med. AMA Arch. Ind. Hyg. Occup. Med.
## 8 A.M.A. Archives of Internal Medicine AMA Arch. Intern. Med. AMA Arch. Intern. Med. AMA Arch. Intern. Med.
## 9 A.M.A. Archives of Neurology AMA Arch. Neurol.  AMA Arch. Neurol.  AMA Arch. Neurol.
## 10 A.M.A. Archives of Neurology and Psychiatry AMA Arch. Neurol. Psychiatry AMA Arch. Neurol. Psychiatry AMA Arch. Neurol. Psychiatry
## # ... with 14,494 more rows
```

```
combined_terms <- bind_rows(new_nlm_terms, old_endnote_terms, .id = "id") %>%
  filter(!is.na(Abbreviation_1) & !is.na(Abbreviation_2) &
    Abbreviation_1 != "" & Abbreviation_2 != "")
```

```
## # A tibble: 56,138 x 4
##   id      Journal_Name
##   <chr> <chr>
## 1 1      AADE editors' journal
## 2 1      AANA journal
## 3 1      AARN news letter
## 4 1      [Nihon shika hyoron] The Nippon dental review
## 5 1      Academia Peruana de Cirugia
## 6 1      Academy review of the California Academy of Periodontology, United States Section, ARPA Int
## 7 1      Acarologia
## 8 1      Acquisitions medicales recentes
## 9 1      Activitas nervosa superior
## 10 1     Acta allergologica
## # ... with 56,128 more rows
```

```
without_punct <- filter(combined_terms, str_detect(Journal_Name,
                                                    "\\(\\|\\\\[\\.*(\\\\|\\\\\\\\)|[:punct:]]")) %>%
  mutate(Journal_Name = if_else(str_remove_all(Journal_Name,
                                                  "\\(\\(\\|\\\\[\\.*(\\\\|\\\\\\\\))|[:punct:]]") == "",
                                str_remove_all(Journal_Name,
                                                  "\\(\\|\\\\[\\\\\\\\|\\\\\\\\)|[:punct:]]"),
          str_remove_all(Journal_Name,
                          "\\(\\(\\|\\\\[\\.*(\\\\|\\\\\\\\))|[:punct:]]")) %>%
  str_trim()

without_punct
```

```
combined_terms <- bind_rows(without_punct, combined_terms) %>%
  arrange(Journal_Name, as.numeric(id)) %>%
  select(-id) %>%
  distinct(str_to_lower(Journal_Name), .keep_all = TRUE) %>%
  select(`str to lower(Journal Name)`)
```

```
write_tsv(combined_terms, "Combined_Terms.txt", col_names = FALSE)
```

```
combined_terms
```

```
## # A tibble: 59,128 x 3
```

```
##   Journal_Name
```

```
##   <chr>
```

```
## 1 ... Asian Conference on Pattern Recognition. Asian Conference on Pattern Recognition
```

```
## 2 ... Electronic Components & Technology Conference. Electronic Components & Technology Conference
```

```
## 3 ... Health innovations and point-of-care technologies conference. Health innovations and point-of-
```

```
## 4 ... IEEE Conference on Collaboration and Internet Computing. IEEE Conference on Collaboration and
```

```
## 5 ... IEEE Conference on Communications and Network Security. IEEE Conference on Communications and
```

```
## 6 ... IEEE conference on high performance extreme computing. IEEE Conference on High Performance Ex
```

```
## 7 ... IEEE Global Communications Conference. IEEE Global Communications Conference
```

```
## 8 ... IEEE Global Conference on Signal and Information Processing. IEEE Global Conference on Signal
```

```
## 9 ... IEEE International Conference on Computer Vision workshops. IEEE International Conference on C
```

```
## 10 ... IEEE International Conference on Multimedia and Expo workshops. IEEE International Conference
```

```
## # ... with 59,118 more rows
```