Search NLM

| Q



COVID-19 Information

Get the latest public health information from CDC Get the latest research information from NIH | Español Learn more about COVID-19 and you from HHS ×

Home > MEDLINE/PubMed Resources

MEDLINE®/PubMed® Data Element (Field) Descriptions

This document describes the major elements (or fields) found on the MEDLINE display format for PubMed MEDLINE records. Some elements (e.g., Comment In) are not mandatory and will not appear in every record. Other elements (e.g., Author, MeSH term, Registry Number) may appear multiple times in one record. **Some** of the elements on this list are searchable fields in PubMed. For searching instructions, see the <u>Search Field Tags</u> section of <u>PubMed Help</u>. This document is supplementary information, to be used in conjuction with <u>PubMed Help</u>. For news about changes to MEDLINE/PubMed data, please see the <u>NLM Technical Bulletin</u>.

MEDLINE/PubMed <u>XML data element descriptions</u> are also available. There are additional fields in the XML data.



Important Note: As of October 2016, the publisher of the original article has the capability to edit the citation data in the PubMed record using the PubMed Data Management system, with the exception of MeSH data. PubMed record data, therefore, may vary from what is described, below.

PubMed MEDLINE display elements are presented in this section in alphabetical order, except in the case of certain elements that are closely related to one another.

Field	Abbreviation
Abstract	<u>(AB)</u>
Copyright Information	<u>(CI)</u>
Affiliation	<u>(AD)</u>
Investigator Affiliation	(IRAD)
Article Identifier	<u>(AID)</u>
Author	<u>(AU)</u>
Author Identifier	(AUID)
Full Author	(FAU)
Book Title	<u>(BTI)</u>
Collection Title	<u>(CTI)</u>
Comments/Corrections	
Conflict of Interest Statement	(COIS)
Corporate Author	(<u>CN)</u>
Create Date	(CRDT)
Date Completed	(DCOM)
Date Created	<u>(DA)</u> .
Date Last Revised	<u>(LR)</u>
Date of Electronic Publication	(DEP)
Date of Publication	<u>(DP)</u>
Edition	<u>(EN)</u>
Editor and Full Editor Name	(<u>ED)</u> (<u>FED)</u>
Entrez Date	(EDAT)

Field	Abbreviatio
Gene Symbol	<u>(GS)</u>
General Note	<u>(GN)</u>
Grant Number	<u>(GR)</u>
Investigator Name and Full Investigator Name	<u>(IR) (FIR)</u>
ISBN	(ISBN)
ISSN	<u>(IS)</u>
<u>Issue</u>	<u>(IP)</u>
Journal Title Abbreviation	<u>(TA)</u>
Journal Title	<u>(JT)</u>
<u>Language</u>	<u>(LA)</u>
Location Identifier	<u>(LID)</u>
Manuscript Identifier	<u>(MID)</u>
MeSH Date	(MHDA)
MeSH Terms	<u>(MH)</u>
NLM Unique ID	(JID)
Number of References	<u>(RF)</u>
Other Abstract	<u>(OAB)</u>
Other Copyright Information	<u>(OCI)</u>
Other ID	<u>(OID)</u>
Other Term	<u>(OT)</u>
Other Term Owner	<u>(OTO)</u>
<u>Owner</u>	<u>(OWN)</u>

Field	Abbreviation
Pagination	<u>(PG)</u>
Personal Name as Subject	<u>(PS)</u>
Full Personal Name as Subject	<u>(FPS)</u>
Place of Publication	<u>(PL)</u>
Publication History Status	(PHST)
<u>Publication Status</u>	<u>(PST)</u>
Publication Type	<u>(PT).</u>
Publishing Model	(PUBM)
PubMed Central Identifier	<u>(PMC)</u>
PubMed Central Release	(PMCR)
PubMed Unique Identifier	(PMID)
Registry Number/EC Number	<u>(RN)</u>
Substance Name	<u>(NM)</u>
Secondary Source ID	<u>(SI)</u>
Source	<u>(SO)</u>
Space Flight Mission	<u>(SFM)</u>
Status	(STAT)
Subset	<u>(SB)</u>
<u>Title</u>	<u>(TI)</u> .
<u>Transliterated Title</u>	<u>(TT)</u>
<u>Volume</u>	<u>(VI)</u>
<u>Volume Title</u>	<u>(VTI)</u>

Abstract (AB)

English-language abstracts are as supplied by the publisher or taken directly from the published article. If the article does not have a published abstract, the National Library of Medicine does not create one, thus the record lacks the Abstract element. However, in the absence of a formally labeled abstract in the published article, text from a substantive "summary," "summary and conclusions," or "conclusions and summary" may be used.

All abstracts in the AB field are in English. For abstracts in other languages, see the <OtherAbstract> element in the XML or other display.

Publishers have given the National Library of Medicine permission to use abstracts for which they claim copyright. NLM does not hold copyright on the abstracts in MEDLINE. Users should obtain an opinion from their legal counsel for any use they plan for the abstracts in the database.

Generally, there are no abstracts for records created before 1975. However, starting in April 2007 NLM began to add abstracts from articles in <u>PubMed Central</u> (PMC) to the equivalent MEDLINE/PubMed citation record if that record does not already contain an abstract. The abstracts are derived from the <u>PMC scanning project</u> which is digitizing the back issues of participating PMC journals. As a result, additional records published prior to 1975 will contain abstracts.

If a collaborating data producer has created an abstract, it will appear in the Other Abstract (OAB) field. This field may be used in addition to or instead of Abstract

Because data entry policies at NLM have changed over the years, abstracts in some records may be truncated, in which case one of the following phrases may appear at the end of the text enclosed in parentheses:

ABSTRACT TRUNCATED AT 250 WORDS

ABSTRACT TRUNCATED AT 400 WORDS

ABSTRACT TRUNCATED (This message occurred infrequently once the maximum length was raised to 4,096 characters in 1996.)

Example:

AB - Many disorders may result in delay of language. The reason for suggesting this diagnostic category is to stress that these children do initially behave in a similar way to those who are peripherally deaf. (ABSTRACT TRUNCATED AT 250 WORDS)

Original policy on inclusion of abstracts set a limit of 250 words for acceptance. Effective with the January 1984 data (i.e., NLM's ELHILL legacy system 8401 Entry Month) two changes were made in this policy: 1) the word limit was expanded to 400 words for abstracts from articles ten pages or more in length or from articles in the core journals identified by the National Cancer Institute and 2) abstracts exceeding the 250- or 400-word limit were to be included in truncated form at the end of the sentence closest to the word limit. The maximum length of abstracts for records created between 2000 and 2016 was 10,000 characters. Abstracts are no longer truncated.

The percentage of records with abstracts has increased over the years as more publishers gave permission for NLM to include these data. A chart showing the number of MEDLINE records containing abstracts in various segments of MEDLINE is available at: http://www.nlm.nih.gov/bsd/medline_lang_distr.html.

<u>Structured abstracts</u>, describing key aspects of the purposes, methods, and results in a consistent way, are published in some journals. The key aspects of structured abstracts are capitalized to stand out, e.g., BACKGROUND, OBJECTIVES, METHOD, etc. Structured abstracts were not truncated in the past, even if they surpassed the previous 250 or 400 word limit.

Example:

AB - BACKGROUND: Superantigens produced by Staphylococcus aureus and Streptococcus pyogenes are among the most lethal of toxins. Toxins in this family trigger an excessive cellular immune response leading to toxic shock. OBJECTIVES: To design an antagonist that is effective in vivo against a broad spectrum of superantigen toxins. METHODS: Short peptide antagonists were selected for their ability to inhibit superantigen-induced expression of human genes for cytokines that mediate shock. The ability of these peptides to protect mice against lethal toxin challenge was examined. RESULTS: Antagonist peptide protected mice against lethal challenge with staphylococcal enterotoxin B and toxic shock syndrome toxin-1, superantigens that share only 6% overall amino acid homology. Moreover, it rescued mice undergoing toxic shock. Antagonist peptides show homology to a beta-strand/hinge/alpha-helix domain that is structurally conserved among superantigens, yet remote from known binding sites for the major histocompatibility class II molecule and T cell receptor that function in toxic T cell hyperactivation. CONCLUSIONS: The lethal effect of superantigens can be blocked with a peptide antagonist that inhibits their action at the top of the toxicity cascade before activation of T cells occurs. Superantigenic toxin antagonists may serve not only as countermeasures to biologic warfare but may be useful in the treatment of staphylococcal and streptococcal toxic shock, as well as in some cases of septic shock.

Table of Contents

Copyright Information (CI)

Copyright Information (CI) associated with Abstract was introduced in 1999, and appears on a limited but increasing number of records. This singly-occurring element contains a copyright statement provided by the publisher of the journal and appears only on records supplied electronically to NLM by the publisher. This information is displayed at the end of the abstract.

Example:

AB - ... Copyright 1999 Academic Press.

See also the related Other Copyright Information (OCI) field.

Table of Contents

Affiliation (AD)

The affiliation of the authors, corporate authors and investigators appear in this repeating field. Until 2014, only the affiliation of the first author was included. The data included in this field and control of the data has changed over time, as follows:

- 1988- The address of the first author's affiliation is included. The institution, city, and state including zip code for U.S. addresses, and country for
 countries outside of the United States, are included if provided in the journal; sometimes the street address is also included if provided in the
 journal.
- 1995-2013 The designation USA is added at the end of the address when the first author's affiliation is in the fifty United States or the District of Columbia.
- 1996- The primary author's electronic mail (e-mail) address is included at the end of the Affiliation field, if present in the journal.
- 2003- The complete first author address is entered as it appears in the article with no words omitted.
- October 2013- Quality control of this field ceased in order to accommodate the affiliations for all authors and contributors.
- December 2014- Multiple affiliations for each author or contributor are included.

The affiliation data is provided as supplied by the publisher. Publishers are requested to include the following data if available, separated by commas: division of the institution, institution name, city, state, postal or zip code, country (USA for the United States) followed by a period, then a space followed by the e-mail address.

Examples:

- AD Department of Anesthesiology, University of Virginia Health Sciences Center Charlottesville 22908, USA. med2p@virginia.edu
- AD Departamento de Farmacologia, Facultad de Medicina, Universidad Complutense de Madrid (UCM), 28040 Madrid, Spain.
- AD Center for Children With Special Needs, Children's Hospital, and the Department of Pediatrics, University of Washington School of Medicine, 4800 Sand Point Way NE, CM:09, Seattle, WA 98105-0371, USA. jneff@chmc.org

Table of Contents

Investigator Affiliation (IRAD)

The Investigator Affiliation (IRAD) It identifies the organization that the researcher was affiliated with at the time the article was written and as published in the journal.

Until 2008, Investigator Affiliation was populated only for data created by one of our former collaborating data producers, the National Aeronautics and Space Administration (NASA). Unlike the Affiliation field associated with the Author field, this affiliation generally did not include detailed address information.

This field is associated with the <u>Investigator Name (IR)</u> field.

Examples:

IRAD - Marquette U, Milwaukee, WI

IRAD - VA Med Ctr, Richmond, VA

Table of Contents

Article Identifier (AID)

This field is populated by the publisher. The article identifier values may include the controlled publisher identifier (PII) or the digital object identifier (DOI):

- A PII, or Publisher Item Identifier, is any internal reference identifier used in the publishing process. This identifier is assigned by the publisher.
- A DOI, or Digital Object Identifier, is a number assigned by an international organization. The DOI System is a system for identifying and exchanging intellectual property in the digital environment. DOIs are issued to registrants by the DOI Registration Agency. More information about this standardized format can be obtained at http://www.doi.org/.

The article identifier is used for LinkOut, and is not considered part of the citation source information.

The Article Identifier (AID) field contains Bookshelf accession numbers with the label: [bookaccession] for <u>citations for books and book chapters</u> from the <u>NCBI Books Database</u>.

Examples:

AID - S0272-7358(05)00023-1 [pii]

AID - 10.1016/j.cpr.2005.02.002 [doi]

AID - NBK7050 [bookaccession]

See also the Location Identifier (LID) field.

Table of Contents

Author (AU)

NLM's author indexing policy is explained in the <u>Authorship in MEDLINE</u>.

Format: last name followed by a space and up to the first two initials followed by a space and a suffix abbreviation, if applicable, all without periods or a comma after the last name.

Examples:

McCrary SV

Smith AB 3rd

See Full Author (FAU) for information on initials.

Limitations to the number of author names included in MEDLINE are as follows:

Time Period	Policy	
1966 - 1983:	For records created during this time period, MEDLINE did not limit the number of authors.	
1984 - 1995:	For records created during this time period, NLM limited the number of authors to 10, with "et al." to indicate the existence of additional authors. This practice began with citations created on October 29, 1983.	
1996 - 1999:	For journal issues published during this time period, NLM increased the number of authors from 10 to 25. If there were more than 25 authors, the first 24 were listed, the last author was used as the 25th, and the 26th and beyond became "et al."	
2000 - Present:	Beginning with journal issues published in 2000, MEDLINE does not limit the number of listed authors.	

Beginning in mid-2005, the various policy restrictions on number of author names entered in past years were lifted so that on an individual basis, a record may be edited to include all author names present in the published article, regardless of the limitation in effect at the time the record was first created.

Effective with 1992 date of publication, letters are indexed individually with authors rather than as an anonymous group.

Notes about transliteration of author names:

- Until 1990, NLM transliterated up to five authors' Cyrillic or Japanese names to the Roman alphabet.
- Between 1990 and 2016, the first ten Cyrillic or Japanese names are transliterated. Chinese ideograms were not transliterated by NLM, but if transliterations of the authors names are available in the journal article or table of contents, they were included in the citation, even if that includes only one author in a multi-author article.
- Beginning in 2016, author names are published in Roman characters in all MEDLINE journals, and NLM no longer transliterates Cyrillic or Japanese names. All author names are included as published.

See also Corporate Author (CN).

Table of Contents

Author Identifier (AUID)

Unique author identifiers may be supplied by the publisher for records added or updated beginning in January 2013. The identifier may be an ORCID, an International Standard Name Identifier (ISNI), or from the Virtual International Authority File (VIAF).

Examples:

AUID- ORCID: 0000000247590453 AUID- ORCID: 0000000253538889 AUID- ORCID: 0000000280166689

Table of Contents

Full Author (FAU)

Beginning with articles with a publication year of 2002, PubMed records may include full names. BIOETHICSLINE citations converted to MEDLINE citations in late 2001 have full names on citations prior to publication year 2002; those data are found in the <u>General Note (GN)</u> field. Citations from the Kennedy Institute of Ethics (KIE) may have full names in the Full Author field for publication dates of prior to 2002.

Full author names are entered when they appear in the author position of an article, usually on the title page of an article. If only the last name and initials appear in the author position, then only the last name and initials will be entered, even if a form of the fuller name appears elsewhere in the article or in the Table of Contents for the journal.

Author Initials:

Initials data are generated from the Data Creation and Maintenance System (DCMS) ForeName data element using an algorithm. Here are the highlights of that algorithm:

- Only two initials are included. Initials are at the beginning of a name string or following a break (a space or hyphen). Only capital letters in the ForeName elements are candidates for initials, except for the letter following a hyphen. The letter following a hyphen is a candidate for an initial, unless the string following the hyphen is "ichi."
- When the ForeName data element consists of only initials, there are spaces between initials.
- An initial includes the following particles: da, de, del, do, dos, du, el, el-, and le. All particles except "el-" are followed by a space and are preceded by a space or are at the beginning of the name string. If found, all particles are converted to lower case when generated as part of the Initials data element.
- If the language of the article is Bulgarian, Russian, Serbo-Croatian (Roman), or Ukranian, then one initial may be a 2- or 4-character transliterated mixed-case initial. Current, mixed-case transliteration values are: Dj, Lj, Nj, Ch, Sh, Iu, Ia, Ie, Zh, Kh, Ts, Dz, Shch.

Examples:

Last Name	Fore Name	Suffix	Initials
Melosh	HJ	3rd	HJ
Abrams	Judith		J
Buncke	Gregory M		GM
Gonzales-loza	Maria del R		Mdel R
Dubuisson	Jean-Bernard		JB
Amara	Mohamed el-Walid		Mel-W
Shan	Yu-fei		YF
Krylov	la K		laK
De Avila	Luiz Francisco Rodriguez		LF

Note that the end result of generating the Initials data is that the two initials are closed up with no space between, even though there may be spaces elsewhere in the Initials string if one or both of the initials has embedded spaces.

There are some author names that have no initials. These are mostly Malaysian names, for which the entire name is entered in the LastName field.

The full name will display on the PubMed MEDLINE display format above the respective name field as in the following example:

Example:

FAU - Foa, Edna B AU - Foa EB

Note: When an author's name has been corrected from a published erratum, the corrected name is placed in the Author field and the incorrect name that was originally published is retained in the last occurrence of the Author field. In this case, there will be an associated commentary linkage.

NLM expended much effort to parse the data converted from the legacy ELHILL format at the end of the 2000 production year accurately. Many citations from the 1966-1974 timeframe were changed to follow data entry conventions established later. For example, particles such as "van der" were moved from the suffix position to the beginning of the last name, and the abbreviations "2d" and "3d" were changed to "2nd" and "3rd". It is possible to have particles associated with initials, such as "Mdel R" for "María del R". It is also possible to have only a last name. Some occurrences of author data in this category are in error and will be corrected manually as time permits.

For <u>OLDMEDLINE</u> records, every published author name is included in the list of authors for citations from the 1951 - 1959 *Current List of Medical Literature* (CLML) and for citations from the 1960 - 1965 *Cumulated Index Medicus* (CIM). For citations from the 1950 CLML, a maximum of three author names were entered. OLDMEDLINE first and last name author elements are in all upper case letters, except in some cases the particle is in lower case letters. A suffix is in upper and lower case letters. OLDMEDLINE records do not contain collective or corporate names. A small percentage of OLDMEDLINE records contain the last name only, because that is the only Author data present in the abstracting and indexing tool used to create the record.

Table of Contents

Book Title (BTI)

Book Title (BTI) is added to <u>citations for books</u> from the <u>NCBI Books Database</u>.

Example:

BTI - Medical Surge Capacity: Workshop Summary

BTI - Drug Class Review on Proton Pump Inhibitors: Final Report

BTI - StemBook

Table of Contents

Collection Title (CTI)

Collection Title (CTI) is added to citations for books from the NCBI Books Database.

Example:

CTI - Drug Class Reviews

Table of Contents

Comments/Corrections (See fields and field tags listed below.)

The data in these fields, listed below, are citations to associated journal publications, e.g., comments, errata, or retractions. These data enable links between the record at hand and its associated citations.

MEDLINE records may contain one or more of the following. See <u>more detailed information about the policy for each type of comment or correction</u>. To search PubMed for records with comments or corrections, <u>see PubMed Help</u>.

Comment or Correction Type	MEDLINE Display Field Tag	Description
Comment in	(CIN)	cites the reference containing a commentary about the article (appears on citation for original article); began use with journal issues published in 1989.
Comment on	(CON)	cites the reference upon which the article comments; began use with journal issues published in 1989.
Erratum in	(EIN)	cites a published erratum to the article (appears on citation for original article); began use in 1987.
Erratum for	(EFR)	cites the original article for which there is a published erratum. As of 2016, partial retractions are considered errata.
Corrected and Republished in	(CRI)	cites the final, correct version of a corrected and republished article (appears on citation for original article). Began use in 1987 as Republished in (RPI); renamed in 2006.
Corrected and Republished from	(CRF)	cites the original article subsequently corrected and republished. Began use in 1987 as Republished from (RPF); renamed in 2006.
Dataset described in	(DDIN)	cites a description of a dataset.

Dataset use reported in	(DRIN)	cites articles reporting results or use of a dataset.
ExpressionOfConcernIn	(ECI)	cites the expression of concern (appears on citation for original article); began use in 2017.
ExpressionOfConcernFor	(ECF)	cites the original article for which there is an expression of concern; began use in 2017.
Republished in	(RPI)	cites the subsequent (and possibly abridged) version of a republished article (appears on citation for original article); began use in 2006.
Republished from	(RPF)	cites the first, originally published article; began use in 2006.
Retraction in	(RIN)	cites the retraction of the article (appears on citation for original article); began use in August 1984 but applied back through 1966.
Retraction of	(ROF)	cites the article(s) being retracted; began use in August 1984 but applied back through 1966.
Update in	(UIN)	cites an updated version of the article (appears on citation for original article); began limited use in 2001.
Update of	(UOF)	cites the article being updated; limited use; began limited use in 2001.
Summary for patients in	(SPIN)	cites a patient summary article; began use in November 2001 (these records contain Publication Type, Patient Education Handout). See the article 'Patient Education Handouts in MEDLINE®/PubMed®' in the <i>NLM Technical Bulletin</i> at http://www.nlm.nih.gov/pubs/techbull/ma02/ma02_new_pt.html for more information.
Original report in	(ORI)	cites a scientific article associated with the patient summary.

The PubMed Identifier (PMID) of the associated record in PubMed is provided (if available) to create a link between an article and its commentary.

Examples:

CON - Dev Cell. 2002 Jul;3(1):85-97. PMID: 12110170

CIN - N Engl J Med. 2003 Jul 17;349(3):211-2. PMID: 12867604

CRI - Orthop Nurs. 2003 May-Jun;22(3):232-9. PMID: 12872752

CRF - Biochemistry. 1994 May 10;33(18):5614-22. PMID: 8180186

EIN - Acta Obstet Gynecol Scand. 2003 Jan;82(1):102

EFR - J Arthroplasty. 2002 Jun;17(4):524-6. PMID: 12066289

RIN - J Biochem Mol Biol. 2002 Nov 30;35(6):642. PMID: 12476908

ROF - Ware FE, Lehrman MA. J Biol Chem. 1996 Jun 14;271(24):13935-8. PMID: 8663248

UIN - Cochrane Database Syst Rev. 2002;(3):CD003688. PMID: 12137706

UOF - Cochrane Database Syst Rev. 2002;(2):CD003680. PMID: 12076500

SPIN- Ann Intern Med. 2003 Jun 3;138(11):160. PMID: 12779314

ORI - Ann Intern Med. 2003 Jun 3;138(11):907-16. PMID: 12779301

Occasionally, a note is added to the Comment or Correction. The note clarifies the data in the Comments or Correction element. It is most often used with **Erratum In** for corrected <u>author names</u>. The following are some possible notes:

- [added] is used when an author name is added to the citation as the result of a published erratum
- [removed] when an author name is removed
- [dosage error in published abstract; MEDLINE/PubMed abstract corrected] is used when there is a dosage error in a citation abstract. Within the abstract, then there is a bracketed phrase that indicates where the correction is: [DOSAGE ERROR CORRECTED]
- [dosage error in article text] when the dosage error is in the text portion of the article
- [abstract no. xxx only] for an erratum of a numbered abstract which is part of an overall citation
- [abstract by author names on page xxx only] when an erratum refers to an unnumbered abstract which is authored and part of an overall citation. The author names and page numbers of the abstract are included.
- [abstract abstract title on page xxx only] when an erratum refers to an unnumbered abstract which is not authored. In this case, the abstract title and page number are included.

As of 2016, partial retractions are considered errata, and data from 2007- (when partial retractions were first referenced in a separate data field) were moved to EIN and EFR notes.

In records from <u>OLDMEDLINE</u>, Comments In (CIN) is the only Comment or Correction found. Other comment or correction elements may be used in the future.

See "Errata, Retractions, Corrected and Republished Articles, Duplicate Publications, Comments (including Author Replies), Updates, Patient Summaries, and Republished (Reprinted) Articles Policy for MEDLINE®" at http://www.nlm.nih.gov/bsd/policy/errata.html for additional information.

Table of Contents

Conflict of Interest Statement (COIS)

This field contains a conflict of interest statement as provided by the publisher. This field was introduced in 2017.

Example:

COIS - [The author] has been paid for developing and delivering educational presentations for [foundation].

COIS - The author acknowledges the donation of [devices] from [company] for the pilot study.

Corporate Author (CN)

This field identifies the corporate authorship of an article. Corporate Author (CN) was introduced in mid-November 2000. NLM's author indexing policy is explained in the <u>Authorship in MEDLINE</u>.

These names enter the database exactly as they appear in the journal (except to delete initial articles such as The, A or An). NLM will not edit the names to standardize them or translate them into English. NLM enters the Roman alphabet words (e.g., German, French) into the Corporate Author field. Transliterated Russian or other cyrillic names are also entered into the Corporate Author field, but for Japanese, Chinese, Hebrew, and Arabic, NLM puts the English translation of the name into this field.

Example:

CN - Centers for Disease Control and Prevention

From mid-November 2000 to April 2006, the corporate author name displayed in PubMed citations as the last occurrence in the author field, as a separate data element after any personal names. Effective May 2006, the collective author is retained in the order of all authors found in the byline of the published article. See the May-June 2006 Technical Bulletin article for details.

Citations prior to 1966, in general, have no indication of collective author unless they were created by NLM's data creation partners. Citations from 1966 to 2000 with collective author field data contain that data in the Title field. These records are generally those created by NLM's data creation partners, and are very few in number and typically in the population or ethics subject areas. As they are encountered, these retrospective records may be individually maintained to move the Corporate Author information from the Title field to the Corporate Author field.

Table of Contents

Create Date (CRDT)

Create Date is the date the record was added to the database. Create Date was implemented with all PubMed records when the 2009 system became available on December 15, 2008. All records added to PubMed prior to implementation received a Create Date equal to Entrez Date (EDAT).

Format: YYYY/MM/DD HH:MM

Example:

CRDT - 2010/07/21 06:00

Table of Contents

Date Completed (DCOM)

Date Completed is the date processing of the record ends; i.e., MeSH[®] Headings have been added, quality assurance validations are completed, and the completed record subsequently is distributed to PubMed and licensees. This is contrasted with <u>Date Created</u>, which is the date processing begins.

Format: YYYYMMDD

Example:

DCOM - 20020207

In Process records lack the Date Completed field. For <u>OLDMEDLINE</u> records, the Date Completed is the approximate date the record entered PubMed, rather than the date processing ends. OLDMEDLINE records are created and processed differently than MEDLINE records.

Table of Contents

Date Created (DA)

Date Created is the date that processing of the record begins.

Format: YYYYMMDD

Example:

DA - 2002051

This is contrasted with <u>Date Completed</u>, which is the date processing ends. Date Created is not the same as NLM's PubMed <u>Entrez Date</u> (EDAT), which is the date the record entered PubMed.

For citations up to about the year 2000, the Date Created (DA) and <u>Date Completed (DCOM)</u> data elements are identical. These dates were derived from NLM's legacy ELHILL system.

For <u>OLDMEDLINE</u> citations converted from the 1964 and 1965 *Index Medicus* (IM), Date Created represents the year and month the citations were printed in the monthly *Index Medicus*, and the day will always be "01". All other OLDMEDLINE records have a year based on the year of the printed index, the month is always "12" for December, and the day is always "01".

Table of Contents

Date Last Revised (LR)

The Date Last Revised indicates the date a change is made to a record. Publisher-supplied and In Process records may be given an LR when <u>Grant information</u> or <u>PMCIDs</u> are added to the record. This is done to facilitate the tracking of compliance for NIH Public Access. Completed (MEDLINE and PubMed-not-MEDLINE) records may be given an LR as a result of individual or global maintenance.

When the 10 million+ MEDLINE records through the 2000 production year were converted to XML from NLM's legacy ELHILL system, all records were assigned a Last Revision Date of 20001218 (December 18, 2000). Subsequently, many of these records have been or will be maintained and given a later LR date.

The nature and content of the revision is not indicated on the record.

Only the latest revision date is displayed.

Format: YYYYMMDD

Example:

LR - 20020320

Table of Contents

Date of Electronic Publication (DEP)

The Electronic Publication Date is the date the publisher made an electronic version of the article available. In January 2003, global maintenance was performed to add this element retrospectively.

Format: YYYYMMDD

Example:

DEP - 20050513

Table of Contents

Date of Publication (DP)

Date of Publication contains the full date on which the issue of the journal was published. The standardized format consists of elements for a 4-digit year, a 3-character abbreviated month, and a 1 or 2-digit day. Every record does not contain all of these elements; the data are taken as they are published in the journal issue, with minor alterations by NLM such as abbreviations.

Example:

DP - 2001 Apr 15

DP - 2001 Apr

DP - 2000 Spring

DP - 2000 Nov-Dec

DP - 2001

Table of Contents

Editor (ED) and Full Editor Name (FED)

Editor and Full Editor Name are included only in citations for books and book chapters from the NCBI Books Database.

Format: See Author (AU) and Full Author (FAU).

Examples:

ED - Altevogt BM

FED - Altevogt, Bruce M

ED - Nadig L

FED - Nadig, Lori

Table of Contents

Edition (EN)

Edition (EN) is included only in citations for books and book chapters from the NCBI Books Database.

Example:

EN - 2nd

Table of Contents

Entrez Date (EDAT)

In most cases, Entrez Date is the date the citation was added to PubMed. However, prior to October 9, 2008, the Entrez Date was set equal to the <u>Publication Date (DP)</u> on records with publication dates before September 1997. Beginning on October 9, 2008, the Entrez Date is set equal to the Publication Date (DP) when the record enters PubMed more than twelve months after the date of publication.

Format: YYYY/MM/DD HH:MM

Example:

EDAT- 2003/01/02 04:00

Table of Contents

Gene Symbol (GS): not currently input

This field contains the "symbol" or abbreviated form of gene names as reported in the literature. This element resides in records processed at NLM from 1991 through 1995. Up to 25 occurrences per record may appear. NLM entered the symbols used by authors; there was no authority list or effort to standardize the data.

Examples:

GS - dyrA

GS - PYRE-F

GS - cpa2

Greek characters, superscripts, and subscripts may appear as part of the gene symbol. The code designations for the Greek characters may be found at http://www.nlm.nih.gov/bsd/licensee/greek_characters.html.

Table of Contents

General Note (GN)

This field contains supplemental or descriptive information related to the document. The data in this field may be preceded by the acronym for the collaborator who provided this information.

The acronyms and names are:

Acronym	Name
HMD	History of Medicine Division, National Library of Medicine
HSR	National Information Center on Health Services Research and Health Care Technology, National Library of Medicine
KIE	Kennedy Institute of Ethics, Georgetown University
NASA	National Aeronautics and Space Administration
PIP	Population Information Program; Johns Hopkins School of Health

Examples:

GN - KIE: Article and commentaries

GN - KIE: KIE BoB Subject Heading: health care/economics

Note: BoB Subject Headings are controlled subject vocabulary terms found in the Kennedy Institute of Ethics' Bioethics Thesaurus under which citations print in their publication, <u>Bibliography of Bioethics</u>. The current format of these data in MEDLINE is reflected in the second example beginning with "KIE Bob" or "KIE Bib".

Table of Contents

Grant Number (GR)

This field was introduced in 1981 and is comprised of four parts (modified from three in December 2007 and reordered in 2009):

- 1. **Number** contains the research grant or contract number (or both) that designates financial support by any agency of the United States Public Health Service, any institute of the National Institutes of Health, or other organization. The data are generally recorded exactly as they appear in the published article*; there is no attempt to standardize the numbers.
- 2. **Grant 2-letter code** contains the 2-letter grant code or acronym.
- 3. **Agency** includes the institute acronym or mnemonic in the case of US PHS institutes, or full organization name. As of 2009 this includes the agency's hierarchical structure from lower to higher entity, when known. For example NCI NIH HHS for National Cancer Institute, National Institutes of Health, Department of Health and Human Services.
- 4. Country contains the home country of the granting agency.

Examples:

GR - LM0577/LM/NLM NIH HHS/United States

GR - M0-1 RR07122/RR/NCRR NIH HHS/United States

GR - Wellcome Trust/United Kingdom

GR - 058423/Wellcome Trust/United Kingdom

GR - 067427/z/02/z/Wellcome Trust/United Kingdom

Grant numbers are added to the record in two ways: 1) by NLM Indexers who have verified the grant information as published and 2) via a submission from PubMed Central to MEDLINE/PubMed when grant data are provided via the NIH Manuscript System (NIHMS) or other manuscript submission system (e.g., Europe PMC Plus).

A list of the possible values for the grant Acronym and Agency is available from: <u>//www.nlm.nih.gov/bsd/grant_acronym.html</u>. Please see links at the bottom to announcements of the addition of various granting agencies.

Be advised that while NLM enters the grant number, acronym and agency values are derived by using a machine algorithm against the grant number string. This may result in some inaccurate derivations.

Through 1999, NLM entered up to three grant numbers for each record. Beginning in 2000, NLM began to transition to an unlimited number of grant numbers or contract numbers. Some MEDLINE citations from 2000 and 2001 may still be limited to three grant numbers or contract numbers, but beginning in 2002, NLM does not limit the number of grant numbers or contract numbers. Some collaborating partners record grant numbers for agencies outside the U.S. Public Health Service in the <u>General Notes</u> field.

*In July 2006, NLM corrected a large number of NIH grant number prefixes which were entered with the letter O (e.g., RO1) rather than the number 0 (e.g., RO1). This practice deviates with NLM's general policy that the data in the online citation match what is in the published article.

Table of Contents

Investigator Name (IR) and Full Investigator Name (FIR)

Beginning in March of the 2008 production year, Investigator Name and Full Investigator Name fields are used to contain personal names of individuals (e.g., collaborators and investigators) who are not authors of a paper but rather are listed in the paper as members of a collective/corporate group that is an author of the paper. The same name listed multiple times will be repeated because NLM cannot make assumptions as to whether those names are the same person.

These fields also reside on MEDLINE citations created or maintained by one of our former collaborating data producers, the National Aeronautics and Space Administration (NASA). They identify the NASA-funded principal investigator(s) who conducted the research discussed in the article cited (but are not necessarily the authors). NASA Investigator Names are associated with the <u>Investigator Affiliation</u> field.

See Author field for details on format. NLM's author indexing policy is explained in the Authorship in MEDLINE.

Examples:

IR - Smith P

FIR - Smith, Paula

IR - Brody BA

FIR - Brody, B A

Table of Contents

ISBN (ISBN)

The ISBN (International Standard Book Number) uniquely identifies a book. This value is found only on citations for book and book chapters from the NCBI <u>Books Database</u>.

Examples:

ISBN - 9780309109475 ISBN - 0309146747

Table of Contents

ISSN (IS)

ISSN (International Standard Serial Number) is an eight-character value that uniquely identifies the cited journal. It is nine characters long in the hyphenated form: XXXX-XXXX. The ISSN field has a qualifier that follows the ISSN data which will state whether it is for the print or the electronic version ISSN of the journal. For journals with multiple ISSNs (e.g., those with separate ISSNs for the print and electronic versions), the ISSN in the MEDLINE citation reflects the version used for MeSH indexing.

Examples:

IS - 0021-5252 (Print)

IS - 1471-2202 (Electronic)

Some records do not contain an ISSN value.

The complete title of the journal can be found in the <u>Journal Title (JT)</u> field. See <u>more information about journals cited in MEDLINE</u>.

See also NLM Unique ID (JID).

Table of Contents

Issue (IP)

Issue identifies the issue, part or supplement of the journal in which the article was published.

NOTE: The publisher of the original article may edit this field in PubMed data at any time.

Examples:

IP - 11

IP - 7 Pt 1

IP - First Half

IP - 3 Suppl 1

For records from <u>OLDMEDLINE</u>, some records contain Issue but lack Volume; some records contain Volume but lack Issue; and some records contain Volume and Issue data in the Volume element.

Table of Contents

Journal Title Abbreviation (TA)

This field contains the standard abbreviation for the title of the journal in which an article appeared. See the NLM Fact Sheet "Construction of National Library of Medicine Title Abbreviations" at https://www.nlm.nih.gov/tsd/cataloging/contructitleabbre.html, which discusses the rules currently used by the National Library of Medicine (NLM) to construct title abbreviations for journals cited in MEDLINE.

NOTE: The publisher of the original article may edit this field in PubMed data at any time.

Examples:

TA - JAMA

TA - J Pediatr

TA - J Comp Physiol B

TA - Ann Biol Clin (Paris)

The complete title of the journal can be found in the <u>Journal Title (JT)</u> field. See <u>more information about journals cited in MEDLINE</u>.

All MEDLINE/PubMed records must be linked to a parent serial record in NLM's <u>online catalog</u>. In <u>OLDMEDLINE</u> records, the journal title abbreviation may differ from that found on the original citation in the printed index.

Table of Contents

Journal Title (JT)

This field contains the full journal title, taken from NLM's cataloging data following NLM rules for how to compile a serial name. The NLM journal title abbreviation is in the (TA) element.

Examples are:

Examples:

JT - Molecular microbiology

JT - American journal of physiology. Cell physiology

Some characters that are not part of NLM's <u>MEDLINE/PubMed Character Set</u> reside in a relatively small number of full journal titles. These characters will display as the string 'inverted question mark.'

Table of Contents

Language (LA)

The language in which an article was published is recorded in the Language field. All entries are three letter abbreviations stored in lower case, such as eng, fre, ger, jpn, etc. A record may contain more than one language value. Some records provided by collaborating data producers may contain the value "und" to identify articles whose language is undetermined.

Examples:

LA - eng

LA - rus

A table listing all languages found in MEDLINE is available at: http://www.nlm.nih.gov/bsd/language_table.html. A chart showing the number of English language MEDLINE articles in various segments of MEDLINE is available at: http://www.nlm.nih.gov/bsd/language_table.html.

Table of Contents

Location Identifier (LID)

In April 2008, NLM began accepting LocationID data from publishers for journal citations. This data consists of either a Digital Object Identifier (DOI) or another publisher ID that the publisher has determined serves the role of pagination in a citation in terms of locating the article (PII). This data will be submitted by the publisher in the LocationID field as part of the XML citation.

Examples:

LID - 8083

LID - s0212 16112008000100001

See also the Article ID field.

Table of Contents

Manuscript Identifier (MID)

A Manuscript ID is an identifier assigned to an author manuscript submitted to the <u>NIH Manuscript Submission System</u>. This may be in support of the <u>NIH Public Access Policy</u>, or another funding agency's policy. The following four types of MIDs currently exist in PubMed records:

NIH Manuscript System (NIHMS)

<u>United Kingdom Manuscript System</u> (UKMS)

Howard Hughes Medical Institute (HHMIMS)

Hyper Articles En Ligne (HAL) from the Centre pour la communication scientifique directe (CCSD)

Examples:

MID - NIHMS3373

MID - HHMIMS35653

MID - UKMS1522

MID - HALMS108756

Table of Contents

MeSH Date (MHDA)

The date MeSH terms were added to the citation is recorded in the MeSH Date field. The MeSH date is the same as the Entrez date until MeSH are added.

Format: YYYYMMDD HH:MM

Example:

MHDA- 2005/08/03 09:00

Table of Contents

MeSH Terms (MH)

NLM's controlled vocabulary, Medical Subject Headings (MeSH®), is used to characterize the content of the articles represented by MEDLINE citations.

Of the various MeSH headings assigned to a record, those representing the most significant points are identified with an asterisk (*) in the MeSH display. The remaining descriptors are used to identify concepts that have also been discussed in the item, but that are not the primary topics. In Process and publisher-supplied records lack MeSH terms. See the MeSH home page (http://www.nlm.nih.gov/mesh/meshhome.html) for additional information about MeSH.

Subheadings (also known as Qualifiers) are often used with MeSH terms to help describe more completely a particular aspect of a subject. Subheadings are displayed after the MH and a slash (/). A major topic asterisk before a subheading indicates when the combination of that subheading with its associated MeSH term is a central concept of the article.

The presentation of MeSH terms is alphabetical. The Subheadings associated with a MeSH term are also in alphabetical order, without regard to the presence of the major topic asterisk (*).

Example 1:

MH - Adult

MH - Cardiovascular Diseases/etiology/*mortality

MH - Child Development/*physiology

MH - Embryo and Fetal Development/*physiology

MH - English Abstract

MH - Fetal Growth Retardation/complications/*physiopathology

MH - Human

MH - Infant, Newborn

MH - Nutrition

MH - Risk Factors

MH - Survival Rate

In the above example, the mortality aspect of cardiovascular diseases, the physiology of child development, as well as embryo and fetal development, and the physiopathology aspect of fetal growth retardation are the central concepts of the article. Note that the MeSH term English Abstract (also present in above example) means that a substantive English language abstract is present in the journal or was written by one of NLM's collaborating

data producers. The abstract may or may not be present in the MEDLINE citation as the input policy changed over the years. There are many older non-English language citations without abstracts in MEDLINE but with the MeSH term English Abstract; this indicates that an English abstract is present in the journal, even if not a part of the online record.

Example 2:

MH - Animal

MH - Dogs

MH - *Myocardial Contraction

MH - Myocardium/*metabolism

MH - *Oxygen Consumption

MH - Surface Tension

In the above example, myocardial contraction, the metabolism aspect of myocardium, and oxygen consumption are the central concepts of the article.

There is an ongoing project to map original subject headings assigned to <u>OLDMEDLINE</u> records to current MeSH. Most of the terms assigned to <u>OLDMEDLINE</u> citations are identified as the major topic of the article. The original subject headings have been placed in the <u>Other Term (OT)</u> field.

Table of Contents

NLM Unique ID (JID)

This field is the alpha-numeric identifier for the cited journal. The element's value is the accession number for the journal's record assigned in the NLM Catalog, available at //www.ncbi.nlm.nih.gov/nlmcatalog or via LocatorPlus at https://catalog.nlm.nih.gov/. An NLM Unique ID may appear as seven, eight or nine characters.

Examples:

Citations from the New England Journal of Medicine will have the following JID field value:

JID - 0255562

Citations from the Japanese Journal of Infectious Diseases will have the following JID field value:

JID - 100893704

Table of Contents

Number of References (RF)

This field contains the number of bibliographic references listed in the review article.

Effective October 1, 2010 NLM discontinued the practice of including the number of bibliographic references listed in articles cited in MEDLINE. NLM had included the number of references for the following Publication Types: Review; Consensus Development Conference; Consensus Development Conference, NIH; Interactive Tutorial; Meta-Analysis. This change in policy is prospective only; we will not remove number of references data from existing citations.

Example:

RF - 21

When collaborating data partners recorded the number of references for non-review articles, these data are found in the General Notes field.

Table of Contents

Other Abstract (OAB)

The Other Abstract field can contain an abstract that is written by someone other than the authors of the article.

NLM creates MEDLINE records without an Abstract field when the source journal article does not originally contain an abstract. In cases where a collaborating data partner provides an abstract, or an abstract is available following the journal publication, this abstract may appear in the Other Abstract field for that record. This field is associated with the Other Copyright Information (OCI) field, which displays beneath (OAB) if present. The abstract included in this field is not written by the authors of the article. The internal tracking number of the source document used by the collaborating partner resides in the Other ID field.

Acronyms used in the OAB field are:

AAMC - American Association of Medical Colleges; not currently used

AIDS - Special HIV/AIDS publications with abstracts written by someone other than the author

KIE - Kennedy Institute of Ethics, Georgetown University

PIP - Population Information Program, Johns Hopkins School of Public Health

NASA - National Aeronautics and Space Administration

Publisher - Journal editorial staff (typically for older citations that did not contain abstracts when originally published)

Example

OAB – NASA: The purpose of this review is to delineate the ubiquitous and pivotal role of Ca2+ in diverse physiological processes. Emphasis will be given to the role of Ca2+ in stimulus-response coupling. In addition to reviewing the present status of research, our intention is to critically evaluate the existing data and describe the newly developing areas of Ca2+ research in plants.

Note that the element <<u>OtherAbstract></u> in PubMed XML and other displays can contain abstracts in other languages supplied by the publisher. See the PubMed <u>XML documentation</u> for details.

Table of Contents

Other Copyright Information (OCI)

This field identifies the copyright owner. It appears on some records created by a collaborating data producer if that producer has written the abstract. The information for this field displays beneath the Other Abstract field.

Example:

OCI - NASA Edited

Table of Contents

Other ID (OID)

This field may reside on a record owned by a collaborating partner or on an NLM-owned record to which a collaborating partner added additional information. The data identify: a) the organization responsible for the information on the citation or the document where the information originated, and b) a unique number for that citation or document.

The acronyms used in this field are listed below. Some of the values on this list currently are not in use at this time and some may never be used.

-			
Acronym	Organization		
NASA	National Aeronautics and Space Administration; not currently used		
KIE	Kennedy Institute of Ethics, Georgetown University; not currently used		
PIP	Population Information Program, Johns Hopkins School of Public Health; not currently used		
POP	former NLM POPLINE database; not currently used		
ARPL	Annual Review of Population Law; not currently used		
СРС	Carolina Population Center; not currently used		
IND	Population Index; not currently used		
CPFH	Center for Population and Family Health Library/Information Program; not currently used		
[NLM]	National Library of Medicine; As of January 2018, no longer displays in PubMed Web interface but may exist in downloaded records. Was used for PMCID and manuscript identifier data (NIHMS for NIH Manuscript System, UKMS for United Kingdom Manuscript System, HALMS for the French Manuscript System, or HHMIMS for the Howard Hughes Medical Institute Manuscript System) beginning in 2009 and ending in 2018.		
NRCBL	National Reference Center for Biomedical Literature (for the KIE Reference Library shelving location)		
CLML	Current List of Medical Literature; reserved for future use		
IM	Index Medicus; reserved for future use (intended for pre-1966 publications)		
QCICL	Quarterly Cumulative Index to Current Literature; reserved for future use (intended for pre-1966 publications)		
QCIM	Quarterly Cumulated Index Medicus; reserved for future use (intended for pre-1966 publications)		
SGC	Surgeon General's Catalog; reserved for future use		
For OLDME	r OLDMEDLINE records, this field occurs on records from 1950-1959 and is for internal use at NLM. CLML is currently the value for all OLDMEDLINE		

For <u>OLDMEDLINE</u> records, this field occurs on records from 1950-1959 and is for internal use at NLM. CLML is currently the value for all OLDMEDLINE citations containing this field. The value IM may be used on a limited basis. Other values that may be defined for future use with OLDMEDLINE records are QCICL and QCIM.

Examples:

OID - KIE: 30206

OID - NRCBL: 18.2

OID - CLML: 5834:20412:395

Table of Contents

Other Term (OT)

This field contains largely non-MeSH subject terms (also referred to as Keywords) that describe the content of the article. Beginning in January 2013, author-supplied keywords are included in Other Term. These are displayed below the abstract in PubMed. Other Terms may also be assigned by a collaborating data producer who is identified in the <u>Other Term Owner</u> field.

The Other Term data may be marked with an asterisk (*) to indicate a major concept. Asterisks are for display only.

Examples:

OT - Legal Approach

OT - Health Care and Public Health

Other Terms appear in addition to MeSH terms, with the exception of some OLDMEDLINE records which do not yet contain MeSH headings.

Table of Contents

Other Term Owner (OTO)

This field contains an acronym that precedes the <u>Other Term</u> field. It identifies the organization that provided the Other Term data. The Other Term Owner acronyms and their respective organizations are:

Acronym	Organization	
KIE	Kennedy Institute of Ethics, Georgetown University	
NASA	National Aeronautics and Space Administration	
PIP	Population Information Program, Johns Hopkins School of Public Health	
NLM	National Library of Medicine (used for the OLDMEDLINE records)	
NOTNLM	(The journal publisher or other data provider. This is used when the OT field includes author-supplied keywords. NLM began using this value in January 2013.)	

Table of Contents

Owner (OWN)

The acronym for the organization that supplied the citation data is recorded in this field. Each citation has only one Owner and there are eight possible values in this field:

Acronym	Organization
NLM	National Library of Medicine, Index Section
NASA	National Aeronautics and Space Administration
PIP	Population Information Program, Johns Hopkins School of Public health (not a current value; only on older citations)
KIE	Kennedy Institute of Ethics, Georgetown University
HSR	National Information Center on Health Services Research and Health Care Technology, National Library of Medicine
HMD	History of Medicine Division, National Library of Medicine
SIS	Specialized Information Services Division, National Library of Medicine (not yet used; reserved for possible future use)
NOTNLM	For licensees' use. Not used by NLM.

Table of Contents

Pagination (PG)

Pagination indicates the inclusive pages for the article cited. The pagination can be entirely non-digit data. Redundant digits are omitted. Document numbers for electronic articles are also found here.

Examples:

PG - 12-9

PG - 304-10

PG - 335-6

PG - 1199-201

PG - 24-32, 64

PG - suppl 111-2

PG - 564

PG - E101-6

PG - 44; discussion 44-8

PG - 925; author reply 925-6

PG - e66

PG - XC-CIII

PG - 10.1-8

NOTE: The publisher of the original article may edit this field in PubMed data at any time.

For letters to the editor published between 2003 and 2012, if a reply is written by one or more of the authors of the original article, "author reply" and the relevant pagination are noted in the pagination field of the citation for the letter. For letters published after 2012, author replies are cited separately and linked to the commenting letter and to the original article.

"Discussion" is used within pagination for other types of articles, such as an article presented at a meeting that is followed by the text of a separate discussion or verbal exchange by a panel or others attending the meeting.

Table of Contents

Personal Name as Subject (PS)

Individuals' names appear in this field for citations that contain a biographical note or obituary, or are entirely about the life or work of an individual or individuals. Data is entered in the same format as author names in the Author field. See <u>Author (AU)</u> field for details of format. See also the associated <u>Full Personal Name as Subject (FPS)</u> field description.

Examples:

PS - Koop CE

PS - Zerhouni EA

Table of Contents

Full Personal Name as Subject (FPS)

This field contains the full name of the subject of the article for citations with a date of publication beginning with 2002. It is associated with the <u>Personal Name as Subject (PS)</u> field. See details on format in the <u>Full Author (FAU)</u> field description.

Examples:

FPS - Koop, C Everett

FPS - Zerhouni, Elias Adam

Table of Contents

Place of Publication (PL)

This field indicates the cited journal's country of publication. Valid values are those country names found in the <u>Geographic Locations</u> within the Medical Subject Headings (MeSH). Values may appear in all upper case or in mixed case. On older records, in cases where the place of publication is unknown, the (PL) value is Unknown.

Examples:

PL - United States

PL - FRANCE

PL - Nigeria

Place of Publication data are not maintained when names may change over time. These data indicate where the journal is published, not where the research was conducted.

Table of Contents

Publication History Status (PHST)

This field contains information from the publisher regarding important events in the publishing process. It contains a date, and one of the following values for each date in the publication history:

Value	Meaning
[received]	date manuscript received for review
[accepted]	accepted for publication
[revised]	article revised by publisher or author
[aheadofprint]	published electronically, to be followed by the print

Examples:

PHST- 2004/06/01 [received]

PHST- 2004/09/01 [revised]

PHST- 2005/02/15 [accepted]

Table of Contents

Publication Status (PST)

This field contains the current status of the publication, as submitted by the publisher. Possible values are:

ppublish - published in print (default value)
epublish - electronically published only, never published in print
aheadofprint - electronically published, but followed by print

Examples:

PST - aheadofprint

PST - ppublish

The date on which the current status took effect is submitted in the Date of Publication field.

Table of Contents

Publication Type (PT)

This field describes the type of material that the article represents; it characterizes the nature of the information or the manner in which it is conveyed (e.g., Review, Letter, Retracted Publication, Clinical Trial). Records may contain more than one Publication Type, which are listed in alphabetical order.

Examples:

PT - Multicenter Study

PT - Review

Almost all citations have one of these four basic, most frequently used Publication Types applied to them: Journal Article, Letter, Editorial, News. One of the above four Publication Types is applied to more than 99% of all citations indexed for MEDLINE.

A <u>list of the Publication Types</u> is available from PubMed Help. A list with definitions is available at <u>//www.nlm.nih.gov/mesh/pubtypes.html</u>. Publication Types are also available in the MeSH Database at <u>//www.ncbi.nlm.nih.gov/mesh</u>.

Some Publication Types have similar main headings equivalents in MeSH. In 2008, the form of similar main headings was modified to include "AS TOPIC". For example, the Publication Type CLINICAL TRIAL has a similar MeSH heading CLINICAL TRIALS AS TOPIC.

Historical Notes:

Five citation types were used to describe various types of articles in MEDLARS before 1991. These were HISTORICAL ARTICLE, HISTORICAL BIOGRAPHY, CURRENT BIOGRAPHY-OBITUARY, MONOGRAPH and REVIEW. MONOGRAPHS were chosen for indexing in Index Medicus only from 1976 through 1981. Indexers were expected to identify each of these citation types during the indexing process.

In 1991 these citation types were replaced with a new MEDLARS data element called Publication Type (PT). The number of publication types was expanded to include types which are indexed for other NLM databases in addition to Index Medicus and MEDLINE.

All of the rubrics used to qualify titles have also become publication types.

At the request of the History of Medicine Division of NLM the list of publication types was expanded in 1997 to include 34 "genre terms" to describe the intellectual or literary type of presentation. These are used chiefly by catalogers and indexers for HISTLINE, rather than by indexers for Index Medicus and MEDLINE.

Every item indexed from 1991 forward is described by one or more of the Publication Types.

Beginning in October 2017, publishers are permitted to include publication types in their records provided to NLM. These publication types are verified during the MEDLINE indexing process. Records in PubMed that are not MEDLINE may have incomplete or inaccurate publication types.

Table of Contents

Publishing Model (PUBM)

This field was removed in the 2009 MEDLINE/PubMed production year.

This field described the medium/media in which the cited article is published. This information was derived from data submitted by the publishers.

Examples:

PUBM - Print

PUBM - Print-Electronic

Table of Contents

PubMed Central Identifer (PMC)

This field contains the unique identifer for the cited article in **PubMed Central**. The identifier begins with the prefix PMC.

Example:

PMC - PMC1463022

PMC - PMC2271135

PubMed Central Release (PMCR)

This field contains the embargo date associated with the availability of the published article in PMC.

Example:

PMCR - 2009/08/01

Table of Contents

PubMed Unique Identifier (PMID)

This field is a 1- to 8-digit accession number with no leading zeros. It is present on all records and is the accession number for managing and disseminating records. PMIDs are not reused after records are deleted.

Beginning in February 2012 PMIDs include extensions following a decimal point to account for article versions (e.g., 21804956.2). All citations are considered version 1 until replaced. The extended PMID is not displayed on the MEDLINE format. View the citation in abstract format in PubMed to access additional versions when available (see the article in the <u>Jan-Feb 2012 NLM Technical Bulletin</u>).

Examples:

PMID - 10097079

PMID - 6012557

Prior to the 2004 version of PubMed (available December 3, 2003), many records contained a MEDLINE Unique Identifier in addition to the PMID. NLM no longer displays the MEDLINE Unique Identifier. The PMID has become the unique identifier for the MEDLINE record.

Table of Contents

Registry Number/EC Number (RN)

This field contains identifiers representing the substances mentioned in the article when such identifiers are included in the MeSH record for the substance. The RN field may contain any of the following:

- the unique 10-digit <u>Unique Ingredient Identifiers (UNIIs)</u> assigned by the Food and Drug Administration (FDA) Substance Registration System (SRS)
- the 5- to 9-digit number in hyphenated format assigned by the Chemical Abstracts Service (CAS)

Following the RN/EC field display is the Substance Name (NM) field value, enclosed in parentheses. Please see <u>Substance Name (NM)</u> field section for an explanation of that field.

A zero (0) is a valid value when an actual number cannot be located or is not yet available.

Note: NLM has not added Chemical Abstracts Service (CAS) registry numbers to MeSH since 1998, therefore only registry numbers that were previously available are displayed in the MEDLINE record.

Examples:

RN - Y920US2H9B (benphothiamine)

RN - 69-93-2 (Uric Acid)

RN - EC 3.1.1.34 (Lipoprotein Lipase)

Table of Contents

Substance Name (NM)

This field may contain any of 4 types of supplementary concept record (SCR) data: 1) MeSH SCR chemical and drug terms (Class 1); 2) protocol terms (Class 2); 3) non-MeSH rare disease terms (Class 3) from the NIH Office of Rare Diseases; and 4) organisms (Class 4) including viruses conforming to the International Committee on Taxonomy of Viruses (ICTV) nomenclature.

The MeSH Database and MeSH Browser contain all of these terms.

The Class 1 SCR chemical and drug terms contain the name of the substance that the registry number or the EC number identifies. These Class 1 records are of two types: 1) Supplementary Concept Records, which can be found in the MeSH Browser with a record type of C, or 2) MeSH Category D descriptors, identified in the MeSH Browser with a tree number that begins with D. The Substance Name follows the RN/EC field display, enclosed within parentheses for Class 1 SCR data.

Examples:

RN - 69-93-2 (Uric Acid)

RN - 6964-20-1 (tiadenol)

RN - MOPP protocol

RN - KBG syndrome

RN - Lagos bat lyssavirus

Secondary Source ID (SI)

The SI field contains information pertaining to many types of data discussed in MEDLINE articles, including: a) molecular sequence data; b) gene expression/molecular abundance data (beginning February 2006); c) clinical trial numbers (beginning summer 2005); d) PubChem identifiers (beginning in January 2007) e) BioProject identifiers (beginning in Summer 2014) and f) identifiers for other scientific data repositories (beginning in Summer 2014). The field is composed of the source followed by a slash followed by an accession number. See a <u>complete list of databank sources</u> included in the SI field.

Examples:

SI - GENBANK/AF306859

SI - SWISSPROT/P13209

SI - ClinicalTrials.gov/NCT00000419

SI - GEO/GDS275

SI - PubChem-Substance/17424970

Molecular Sequence Data

NLM cooperates with international efforts to collect molecular sequence data. There are numerous databanks that register molecular sequences deposited with them by researchers. In 1988, NLM began with 7 databanks and added 8 more in 2014. In the journal literature, a reference to the databank and the accession number assigned to the sequence may accompany, or substitute for, a lengthy graphic representation of the sequence itself. The Secondary Source field is populated if this information appears in the journal article. There is no attempt to edit or verify the databank accession numbers that appear in the journal. Since sequences may be deposited with more than one databank, there may be multiple occurrences of SI associated with a single article. This information may appear on the article title page, in a footnote or in a statement such as: Sequence data from this article have been deposited with the EMBL, GenBank and DDBJ Data Libraries under Accession No. M16978.

NLM first began to include molecular sequence data with the 1988 indexing year. Prior to 2000, the NLM policy was to enter up to 30 databank accession numbers for each record. Some global maintenance was done over the years to add databank names/accession numbers whether or not the article itself contained those references. From 2000 forward, NLM enters all databank accession numbers published in the journal.

Clinical Trial Numbers

Beginning in summer 2005, NLM includes the <u>ClinicalTrials.gov</u> identifier number in the SI field when the article is devoted solely and entirely to announcing or reporting the results of the clinical trial. The <u>ICMJE Web site</u> contains an editorial and updates on the topic of registering clinical trials before publication of the results.

Beginning mid-2006, MEDLINE citations also carry the International Standard Randomised Controlled Trial Number (ISRCTN) when the article is devoted solely and entirely to announcing or reporting the results of the clinical trial or other study that the Identifier Number represents. The ISRCTN Register is a clinical trials deposit site based in the UK that meets the criteria set forth by the ICMJE (International Committee of Medical Journal Editors) for responsible disclosure of information to the public. The letters ISRCTN are a part of the trial number. Retrospective maintenance was taken on existing citations in MedlineCitation status = MEDLINE to add ISRCTN numbers if an existing citation's article title or abstract contained that data.

Beginning in 2014, MEDLINE citations also carry identifier numbers for many of the primary registries in the World Health Organization (WHO) Registry Network when the article is devoted solely and entirely to announcing or reporting the results of the clinical trial or other study that the identifier number represents.

Gene Expression Data

Beginning in February 2006, accession numbers for data deposited in NLM's <u>Gene Expression Omnibus</u> (GEO) database are included in SI field. GEO is a gene expression/molecular abundance repository supporting data submissions, and a curated, online resource for gene expression data browsing, query and retrieval.

The databank abbreviation is GEO and the accession number is any one of four prefixes followed by a numeric string:

GDSxxxx (GEO Data Set)

GSExxxx (GEO SEries)

GPLxxxx (GEO PLatform)

GSMxxxx (GEO SaMple)

PubChem Identifiers

Beginning in January 2007, identifiers for records in the <u>PubChem</u> Substance database may be included in the SI field when the data are included in the citation XML feeds from the publishers. The <u>PubChem</u> project provides information on the biological activities of small molecules and its overall goal is to identify new and safer drug therapies. There are three PubChem databases: PubChem Substance, PubChem Compound, and PubChem-BioAssay. Each record in each database has a unique identifier, consisting of one or more numerical characters. The abbreviation is PubChem-Substance and the accession number is a numeric string, e.g. 10318689. In 2014, identifiers for PubChem Compound and PubChem BioAssay records may also be added to MEDLINE/PubMed records.

BioProject Identifiers

Beginning in Summer 2014, identifiers for projects profiled in the NCBI <u>BioProject</u> database may be included in the SI field. A BioProject is a collection of biological data related to a single initiative, originating from a single organization or from a consortium. A BioProject record provides users a single place to find links to the diverse data types generated for that project.

Other Scientific Data Repository Identifiers

Beginning in Summer 2014, dataset accession numbers in the <u>Dryad Digital Repository</u> and <u>figshare</u> may be included in the SI field. The Dryad and figshare repositories support wide varieties of datatypes and are not limited to specific scientific disciplines. The Dryad and figshare repositories are added to encourage authors and publishers to include data availability information in publications for inclusion in MEDLINE/PubMed.

Table of Contents

Source (SO)

NLM concatenates the following elements to create the journal source:

Journal Title Abbreviation (TA)

Publication Date (DP) or Date of Electronic Publication (DEP)

Volume (VI)

Issue (IP)

Pagination (PG)

Whether the source field contains the DP or the DEP date depends on the PubModel values, which are derived from data supplied by the publisher. If the DEP field is used the month value is displayed as a 3-character abbreviation. The PubModel can be found on the XML display.

Depending on the PubModel, the source can be followed by a note showing another date. The date is preceded by one of these labels:

Epub

Print

eCollection

ELocation ID data may also be added in addition to pagination or, if normal pagination data is not included, in place of pagination.

Examples:

- SO Am J Med. 2005 May;118(5):567.
- SO Hepatology. 2004 Apr;39(4):915-23.
- SO Health Care Finance Rev. 2003 Winter; 25(2):77-90.
- SO Eur Spine J. 2005 Nov;14(9):887-94. Epub 2005 Sep 8.
- SO Euro Surveill. 2008 Apr 10;13(15). pii: 18832.
- SO Br J Pharmacol. 2012 May;166(2):554-6. doi: 10.1111/j.1476-5381.2011.01818.x.
- SO Euro Surveill. 2010 May 13;15(19):pii/19567.
- SO Nucleic Acids Res. 2004 Jan 15;32(1):e14.
- SO Nucleic Acids Res. 2004 Jan 16;32(1):380-5. Print 2004.
- SO Front Genet. 2011 Apr 25;2:17. doi: 10.3389/fgene.2011.00017. eCollection 2011.

Table of Contents

Space Flight Mission (SFM)

This field resides on citations created by one of our collaborating data producers, the National Aeronautics and Space Administration (NASA). It is the space flight/mission name and/or number when results of research conducted in space are covered in the cited publication. View a complete list of Space Flight Missions and their associated values (manned/unmanned; short/long duration) at https://www.nlm.nih.gov/bsd/space_flight.html.

Examples:

SFM - Flight Experiment

SFM - Mir Project

SFM - long duration

SFM - manned

Table of Contents

Status (STAT)

The status field indicates the status of the record. There are seven possible values for this field:

- In-Data-Review: Records in this status have been submitted to NLM electronically by the publisher, and the journal title, date of publication and volume/issue (source data) are being checked. The source data is either:
 - 1. matched to the print copy in NLM's collection and are correct;
 - 2. matched to the online version of the journal (when NLM assigns MeSH headings from the online version) and are correct; or
 - 3. compared to previously checked-in issues and appear to match the pattern or have been changed to match the established pattern. In these cases, the physical item has not yet been received for NLM's collection and the data have not been positively verified and may still change during NLM's processing cycle.

While all three reviews are at the issue level, most citations fit into the last condition above. It is possible that the source information may be changed at a later point in the NLM quality assurance cycle once the hard copy issue is available for exact comparison.

In-Data-Review records lack a <u>Date Completed</u> field. They are not yet MEDLINE records because they have not undergone complete quality

review and MeSH indexing. The issue level review for In-Data-Review status is the first step in quality control. The records will typically either be reissued as In-process status records or go to PubMed-not-MEDLINE final record status.

• In-Process: Records in this status are undergoing a citation level review; i.e., the author names, article title, and pagination are being checked.

In-process records do not have a <u>Date Completed</u> field, however, they do contain the <u>Subset</u> field. In-process records are not yet MEDLINE records because they have not undergone complete quality review and MeSH indexing.

Most in-process records are eventually indexed with MeSH terms and are elevated to completed MEDLINE status. However, some are determined to be out of scope (e.g., articles on plate tectonics or astrophysics from certain MEDLINE journals, primarily general science and chemistry journals, for which the life sciences articles are indexed for MEDLINE) and are not elevated to MEDLINE status; instead they become PubMed-not-Medline final status records. In rare cases the records are deleted and do not become PubMed-not-MEDLINE records.

MEDLINE: In-process records undergo rigorous quality assurance routines before they are elevated to MEDLINE status or to PubMed-not-MEDLINE status.

Records in MEDLINE status are the only 'true' MEDLINE records in PubMed. They contain <u>Date Completed</u> and <u>Subset</u> fields and, in most cases, contain MeSH terms. MEDLINE records that are Retractions of Publications (see <u>Publication Type</u> element) are exceptions and do not contain MeSH terms.

- **OLDMEDLINE:** Citations from the <u>OLDMEDLINE</u> subset begin in OLDMEDLINE status. Once *all* of the original subject terms from the printed index have been mapped to current MeSH, the citation status is changed to MEDLINE. See <u>more information</u> about OLDMEDLINE.
- PubMed-not-MEDLINE: Records in this status are from journals included in MEDLINE and have undergone quality review but are not assigned
 MeSH headings because the cited item is not in scope for MEDLINE either by topic or by date of publication, or is from a non-MEDLINE journal.
 The categories of non-MEDLINE records in this status are:
 - citations to articles that precede the date a journal was selected for MEDLINE indexing and are submitted for inclusion in PubMed after July 2003;
 - 2. out of scope citations to articles in journals covered by MEDLINE;
 - 3. analytical summaries of articles published elsewhere (see the article, "<u>Linking MEDLINE Citations to Evidence-Based Medicine Assessments and Summaries</u>", in the May-Jun 2002 *NLM Technical Bulletin*, page e2); and
 - 4. (starting in the Summer of 2005) prospective citations to articles from non-MEDLINE journals that submit full text to PubMed Central[®], and are thus cited in PubMed.

PubMed-not-MEDLINE records contain a <u>Date Completed</u> field and lack <u>Subset</u> and <u>MeSH term</u> fields.

- **Publisher:** The non-MEDLINE records in Publisher Status in PubMed contain the notation [PubMed as supplied by publisher] in the PubMed display. At this time the records in Publisher Status are:
 - 1. the retrospective records (prior to July 2005) for the relatively few non-MEDLINE journals in PubMed;
 - 2. the retrospective records for MEDLINE journals prior to date of selection for MEDLINE and that were submitted electronically by the publishers before late July 2003;
 - 3. the prospective records for currently indexed journals when the publisher has submitted an issue's citation data electronically and NLM still awaits its print copy or access to the electronic copy to use for issue level review (i.e., the journal title, date of publication and volume/issue elements) AND the publisher-supplied record contains a validation error of some kind that prevents it from being exported from NLM's Data Creation and Maintenance System (DCMS) along with the records not containing errors from the same issue. If there were no errors, the record would move to In-Data-Review status right away and be exported. In these cases, however, NLM staff must take corrective action before the record can be elevated to In-Data-Review status for export; and
 - 4. citations electronically submitted for articles that appear on the Web in advance of the journal issue's release (i.e., ahead of print citations). Following publication of the completed issue, the item will be queued for issue level review and released in In-Data-Review status.

Table of Contents

Subset (SB)

This field identifies the subset for MEDLINE records from certain journal lists or records on specialized topics. Some of these values are found on extremely small numbers of records. Citations may contain more than one occurrence of this field. The value is true at the time the record was created. If the status of a journal changes, the value on the record does not change.

Subset field values and their definitions are as follows. Note that several are closed subsets no longer being assigned.

Value	Definition
AIM	citations from <u>Abridged Index Medicus</u> journals, a list of about 120 core clinical, English language journals.
В	citations from non-Index Medicus journals in the field of biotechnology (not currently used).
С	citations from non-Index Medicus journals in the field of communication disorders (not currently used).
D	citations from dental journals. See the current list under <u>Dentistry</u> and <u>Orthodontics</u> .
E	citations in the field of bioethics. (includes records from the former BIOETHICSLINE database)
F	older citations from one journal prior to its selection for <i>Index Medicus</i> . Used to augment the database for NLM's International MEDLARS Centers (not currently used).
Н	citations from non-Index Medicus journals in the field of health administration. (includes records from the former HealthSTAR database)

IM	citations from <i>Index Medicus</i> journals.	
J	citations in the field of population information. (not currently used; on records from the former POPLINE® database)	
K	citations from non-Index Medicus journals relating to consumer health.	
N	citations from nursing journals. See the current list under <u>Nursing</u> .	
ОМ	citations from the OLDMEDLINE project that originated from the Cumulated Index Medicus and the Current List of Medical Literature (in 2008 this includes citations from the 1949-1965 print indexes). The ways they differ from other MEDLINE records are documented under the applicable element descriptions. The original MeSH Headings assigned at the time the citation was created in print reside in the Other Term field. Records in the OLDMEDLINE subset have a status of OLDMEDLINE until all of their original subject terms are mapped to current MeSH; then their status changes to MEDLINE. NLM makes available both new and revised OLDMEDLINE records on an irregular and infrequent basis.	
Q	citations in the field of the history of medicine. (includes records from the former HISTLINE® database)	
QIS	citations from non-Index Medicus journals in the field of the history of medicine. (For NLM use effective in late 2006 because they require special handling at NLM; not a subset of Q; some journals previously designated as Q are now QIS.)	
QO	is subset of Q - indicates older history of medicine journal citations that were created before the former HISTLINE file was converted to a MEDLINE-like format. (For NLM use because they require special handling at NLM).	
R	citations from non-Index Medicus journals in the field of population and reproduction (not currently used).	
s	citations in the field of space life sciences. (includes records from the former SPACELINE™ database)	
т	citations from non-Index Medicus journals in the field of health technology assessment. (includes records from the former HealthSTAR database)	
Х	citations in the field of AIDS/HIV. (includes records from the former AIDSLINE® database)	
Examples: SB - AIM SB - IM SB - X		

Table of Contents

Title (TI)

The title field contains the entire title of the journal article, sometimes translated to English when the original article title is in another language. Those titles originally published in a non-English language and translated to English are enclosed in square brackets. All titles end with a period unless another punctuation mark such as a question mark or bracket is present. Explanatory information about the title itself is enclosed in parentheses, e.g.: (author's transl). Corporate/collective authors may appear at the end of this field for citations up to about the year 2000.

NOTE: The publisher of the original article may edit this field in PubMed data at any time.

Examples:

- TI The Kleine-Levin syndrome as a neuropsychiatric disorder: a case report.
- TI Why is xenon not more widely used for anaesthesia?
- TI [Biological rhythms and human disease]
- TI [The effect of anti-arrhythmic drugs on myocardial function (author's transl)]
- TI Prevalence of Helicobacter pylori resistance to antibiotics in Northeast Italy: a multicentre study. GISU. Interdisciplinary Group for the Study of Ulcer.

Additional information/background:

For records in the OLDMEDLINE subset: For citations from the 1964 and 1965 *Cumulated Index Medicus (CIM)*, the article title is in all upper case letters. Some citations contain the value "Not Available." OLDMEDLINE records do not contain Corporate/Collective authors in the title field.

Table of Contents

Transliterated Title (TT)

This field contains the title of each item originally published in a non-English language, in that language. Transliterations of article titles in some Cyrillic languages (Greek, Bulgarian, Russian, Serbian and Ukrainian) were added to this field through 2004.

Examples:

- TT Temoignages et lettres.
- TT Wplyw przebiegu rozwoju plodu i noworodka na ujawnienie sie niektórych chorób okresu doroslego.

For <u>OLDMEDLINE</u> records, the TT field for citations from the 1964 and 1965 *Cumulated Index Medicus* (CIM) is in all uppercase letters. Some OLDMEDLINE citations to articles originally published in a non-English language lack the TT field.

Volume (VI)

The volume number of the journal in which the article was published is recorded here.

Examples:

VI - 7

VI - 5 Spec No

VI - 49 Suppl 20

Some records (especially records from <u>OLDMEDLINE</u>) contain the Issue field but lack the Volume field; some contain the Volume field but lack the Issue field; and some records contain Volume and Issue data in the Volume element.

Table of Contents

Volume Title (VTI)

Volume Title (VTI) is added to citations for books from the NCBI Books Database

Table of Contents

Last Reviewed: September 3, 2019

Connect with NLM



National Library of Medicine 8600 Rockville Pike Bethesda, MD 20894 Web Policies FOIA Vulnerability Disclosure NLM Support Center Accessibility Careers

NLM | NIH | HHS | USA.gov