

# DICOM Correction Proposal

STATUS	Final Text
Date of Last Update	2017/09/17
Person Assigned	Ulrich Busch ( <a href="mailto:ulrich.busch@varian.com">ulrich.busch@varian.com</a> )
Submitter Name	Ulrich Busch ( <a href="mailto:ulrich.busch@varian.com">ulrich.busch@varian.com</a> )
Submission Date	2016/05/03

Correction Number	CP-1633
Log Summary:	Enhance Specification of Range Matching
Name of Standard	PS3.4 2017c
Rationale for Correction:	<p>The specification of attribute matching does not reflect the level of precision and structural clarity otherwise present in the specification of the DICOM Standard. The specification is therefore re-arranged to organize paragraphs and bullet point systematically along the applicable Value Representation. References to VRs are explicitly spelled out along the VR notation defined in PS 3.4 eliminating the need to interpret narrative text. The changes only enhance the documentation approach while keeping the requirements unchanged. The only statement added is on range matching for VR of TM clarifying that range matching crossing midnight is not supported. It is also clarified that wildcard matching on DT is not supported.</p>
Correction Wording:	

*In Part 4, modify section C.2.2.2 'Attribute Matching' as follows:*

## C.2.2.2 Attribute Matching

The following types of matching may be performed on Key Attributes in the Query/Retrieve Service Class:

- Single Value Matching
- List of UID Matching
- Universal Matching
- Wild Card Matching
- Range Matching
- Sequence Matching

Matching requires special characters (i.e., "\*", "?", "-", "=", "\"), which need not be part of the character repertoire for the VR of the Key Attributes.

### Note

1. For example, the "-" character is not valid for the DA, DT and TM VRs but is used for range matching.
2. When character sets other than the default character repertoire are used, then the rules in PS3.5 apply, such as with respect to the use of the 05/12 "\" (BACKSLASH) (in ISO IR 6) or 05/12 "¥" (YEN SIGN) (in ISO IR 14).

The total length of the Key Attribute may exceed the length as specified in the VR in PS3.5. The Value Multiplicity (VM) may be larger than the VM specified in PS3.6 for the Key Attribute, as defined for particular Matching Type.

The Specific Character Set (0008,0005) Attribute may be present in the Identifier but is never matched. Rather, it specifies how other Attributes are encoded in the Request and Response Identifiers.

It may influence how matching of other Attributes is performed. If Specific Character Set (0008,0005) is absent, then the default character repertoire shall be used. Specific Character Set (0008,0005) shall not have a zero length value.

Specific Character Set (0008,0005) may have multiple values if escape sequences are used to switch between character repertoires within values.

If the SCP does not support the value(s) of Specific Character Set (0008,0005) in the Request Identifier, then the manner in which matching is performed is undefined and shall be specified in the conformance statement.

Note

1. If an SCU sends a Request Identifier with a single byte character set not supported by the SCP, then it is likely, but not required, that the SCP will treat unrecognized characters as wild cards and match only on characters in the default repertoire, and return a response in the default repertoire.
2. Some Specific Character Set values are used with multi-component group person names (e.g., single-byte, ideographic and phonetic and phonetic component groups separated by an "=" (3DH) character), which may also affect the behavior of literal string matching.

The Timezone Offset From UTC (0008,0201) Attribute may be present in the Identifier but is not matched if Timezone query adjustment is negotiated. If Timezone query adjustment is negotiated, it specifies how ~~date and time~~ Attribute values **of VR of DT and TM (including related Attribute values of VR of DA, if present)** are interpreted in the Request and Response Identifiers if those values lack a specific time zone offset specification.

### C.2.2.2.1 Single Value Matching

If the value specified for a Key Attribute in a request is non-zero length and if it is:

- a. not ~~a date or time or datetime, of VR of DA, TM or DT and~~ contains no wild card characters, **or**
- b. ~~a date or time or datetime, of VR of DA, TM or DT and~~ contains a single ~~date or time or datetime value~~ with no "-"

then single value matching shall be performed. Except for Attributes with a PN ~~Value Representation~~VR, only entities with values that match exactly the value specified in the request shall match. This matching is case-sensitive, i.e., sensitive to the exact encoding of the key Attribute value in character sets where a letter may have multiple encodings (e.g., based on its case, its position in a word, or whether it is accented).

#### C.2.2.2.1.1 Attributes of VR of PN

For Attributes with a PN ~~Value Representation~~VR (e.g., Patient Name (0010,0010)), an application may perform literal matching that is either case-sensitive, or that is insensitive to some or all aspects of case, position, accent, or other character encoding variants.

Note

1. For multi-component names, the component group delimiter "=" (3DH) may be present in the Key Attribute value, but may give unexpected results if the SCP does not support matching on separate components but interprets the entire value literally as a single string. E.g., "Wang^XiaoDong=□^□□" may or may not match "Wang^XiaoDong" or "□^□□"; wild card matching without the component group delimiter, such as "\*"Wang^XiaoDong\*" or "\*□^□□\*" may be necessary.
- ~~2. Using attributes with VR of AE, LO, PN and SH as matching keys will not allow single value matching on values that contain characters "\*" and "?" such queries will always be treated as queries with wildcard matching.~~
- ~~3. Attributes with VR of ST, LT and UT are intended for conveying narrative text and may contain wildcard characters "\*" and "?". Attempts to match on a string explicitly containing "\*" or "?" will be treated as wildcard matching and thus may return multiple results rather than a single one.~~

If extended negotiation of fuzzy semantic matching rather than literal matching of PN Value Representation is successful, not only may matching be insensitive to case, position, accent, and character encoding (**including combining characters**), but in addition other techniques such as phonetic matching may be applied.

Note

1. Matching of PN Attributes may be accent-insensitive, as specified in the conformance statement. Accent-insensitive matching would successfully match, for instance, a query character "SMALL LETTER a" (06/01 in the default ISO-IR 6) with

"SMALL LETTER a WITH GRAVE ACCENT" (14/00 in ISO-IR 100),

"SMALL LETTER a WITH TILDE" (14/03 in ISO-IR 100),

"SMALL LETTER a WITH BREVE" (14/03 in ISO-IR 101), and

"CAPITAL LETTER a WITH ACUTE ACCENT" (12/01 in ISO-IR 100) (if matching is also case-insensitive),

but would not match 14/00 in ISO-IR 101, which is "SMALL LETTER r WITH ACUTE ACCENT". Matching to particular bit-combinations is specific to each supported character set (note the difference in meaning of 14/00), and should be described in the conformance statement.

2. An SCU application may elect to perform additional filtering of the responses by applying the matching rules itself. In the event that both the SCU and SCP are applying the matching rules, this process will be successful as long as literal matching is performed by both, and any additional SCU filtering is insensitive to case, position, accent, or other character encoding variants.

If fuzzy semantic matching of person names has been negotiated, matching by the SCP may result in responses that are not obviously related to the request, hence care should be taken if any additional filtering of responses is performed by the SCU. For example, if phonetic matching is performed, a query for "Swain" might well return "Swayne", or if name component order insensitive matching is performed, a query for "Smith^Mary" might well return "Mary^Smith" or "Mary Smith" or "Smith, Mary". Fuzzy semantic matching may also take into account separate single-byte, ideographic and phonetic name component groups.

#### **C.2.2.2.1.2 Attributes of VR of PN, AE, LO, SH, ST, LT and UT**

The PN, AE, LO, SH, ST, LT, and UT VRs allow the presence of wild card matching characters "\*" and "?". Single value matching against such characters is not supported. See C.2.2.2.4 Wild Card Matching.

#### **C.2.2.2.1.3 Attributes of VR of DA, DT or TM**

If the Timezone Offset From UTC (0008,0201) Attribute is present in the Identifier and Timezone query adjustment was negotiated, it shall be used to adjust values of ~~time~~-Attributes **of VR of TM** (and associated ~~date~~ Attributes **of VR of DA**, if present) from the local timezone to UTC. It shall also adjust values of ~~datetime~~-Attributes **of VR of DT** that do not specify a timezone offset. The encoding and semantics of the Timezone Offset From UTC (0008,0201) Attribute shall be as defined in the SOP Common Module in PS3.3.

The manner in which matching is performed is implementation dependent and shall be specified in the conformance statement.

Note

1. This definition implies that ~~dates or times or datetimes~~ **values of VR of TM, DA and DT** are matched by their meaning, not as literal strings. For example:
  - the DT "19980128103000.0000" matches "19980128103000"
  - the DT "19980128103000" with no timezone offset matches "19980128073000" with timezone offset "-0300"
  - the TM "2230" matches "223000"
2. If an application is concerned about how single value matching of dates and times is performed by another application, it may consider using range matching instead, which is always performed by meaning, with both values in the range the same.
3. Exclusion of the "-" character for single value matching implies that a Key Attribute with DT Value Representation may not contain a negative offset from Universal Coordinated Time (UTC) if single value matching is intended. Use of the "-" character in ~~date, time or datetime~~ **values of VR of TM, DA and DT** indicates range matching.

4. If an application is in a local time zone that has a negative offset then it cannot perform single value matching using a local time notation. Instead, it can convert the Key Attribute value to UTC and use an explicit suffix of "+0000".

- ~~5. Matching of PN Attributes may be accent insensitive, as specified in the conformance statement. Accent insensitive matching would successfully match, for instance, a query character "SMALL LETTER a" (06/01 in the default ISO-IR 6) with~~

~~"SMALL LETTER a WITH GRAVE ACCENT" (14/00 in ISO-IR 100),~~

~~"SMALL LETTER a WITH TILDE" (14/03 in ISO-IR 100),~~

~~"SMALL LETTER a WITH BREVE" (14/03 in ISO-IR 101), and~~

~~"CAPITAL LETTER a WITH ACUTE ACCENT" (12/01 in ISO-IR 100) (if matching is also case-insensitive),~~

~~but would not match 14/00 in ISO-IR 101, which is "SMALL LETTER r WITH ACUTE ACCENT". Matching to particular bit combinations is specific to each supported character set (note the difference in meaning of 14/00), and should be described in the conformance statement.~~

- ~~6. An SCU application may elect to perform additional filtering of the responses by applying the matching rules itself. In the event that both the SCU and SCP are applying the matching rules, this process will be successful as long as literal matching is performed by both, and any additional SCU filtering is insensitive to case, position, accent, or other character encoding variants.~~

~~However if fuzzy semantic matching of PN Attributes has been negotiated, matching by the SCP may result in responses that are not obviously related to the request, hence care should be taken if any additional filtering of responses is performed by the SCU. For example, if phonetic matching is performed, a query for "Swain" might well return "Swayne", or if name component order insensitive matching is performed, a query for "Smith^Mary" might well return "Mary^Smith" or "Mary Smith" or "Smith, Mary". Fuzzy semantic matching may also take into account separate single byte, ideographic and phonetic name component groups.~~

#### C.2.2.2.2 List of UID Matching

A List of UIDs is encoded by using the value multiplicity operator, backslash ("\"), as a delimiter between UIDs. Each item in the list shall contain a single UID value. Each UID in the list contained in the Identifier of the request may generate a match.

Note

A list of single values is encoded exactly as a VR of UI and a VM of Multiple (see PS3.5).

#### C.2.2.2.3 Universal Matching

If the value specified for a Key Attribute in a request is zero length, then all entities shall match this Attribute. An Attribute that contains a Universal Match specification in a C-FIND request provides a mechanism to request the selected Attribute value be returned in corresponding C-FIND responses.

#### C.2.2.2.4 Wild Card Matching

If the Attribute is not ~~of VR of DA, DT, TM, SL, SS, US, UL, FL, FD, OB, OW, OD, OF, OL, UN, DS, IS, AS, UI-a date, time, signed long, signed short, unsigned short, unsigned long, floating point single, floating point double, other byte string, other word string, unknown, Attribute tag, decimal string, integer string, age string or UID~~ and the value specified in the request contains any occurrence of an "\*" or a "?", then "\*" shall match any sequence of characters (including a zero length value) and "?" shall match any single character. This matching is case sensitive, except for Attributes with an PN ~~Value Representation~~VR (e.g., Patient Name (0010,0010)).

For Attributes with a PN ~~value representation~~VR, including the case of extended negotiation of fuzzy semantic matching, wild card matching is implementation dependent and shall be specified in the conformance statement.

Note

1. Wild card matching on a value of "\*" is equivalent to universal matching.

2. The wild card matching method specified by DICOM might not be supported by some non-DICOM multi-byte character text processors.
3. For multi-component group names, the component group delimiter "=" (3DH) may be present in the Key Attribute value, but may give unexpected results if the SCP does not support matching on separate components but interprets the entire value literally. E.g., "\*"="\*" or "\*=\*="\*" may or may not return all strings, and hence is not equivalent to "\*", nor to universal matching.
4. Using attributes with VR of AE, LO, PN and SH as matching keys will not allow single value matching on values that contain characters "\*" and "?"; such queries will always be treated as queries with wildcard matching.
5. Attributes with VR of ST, LT and UT are intended for conveying narrative text and may contain wildcard characters "\*" and "?". Attempts to match on a string explicitly containing "\*" or "?" will be treated as wildcard matching and thus may return multiple results rather than a single one.

### **C.2.2.2.5 Range Matching**

#### **C.2.2.2.5.1 Range Matching of Attributes of VR of DA**

In the absence of extended negotiation, ~~if the Attribute is a date,~~ then:

- a. A string of the form "<date1> - <date2>", where <date1> is less or equal to <date2>, shall match all occurrences of dates that fall between <date1> and <date2> inclusive
- b. A string of the form "- <date1>" shall match all occurrences of dates prior to and including <date1>
- c. A string of the form "<date1> -" shall match all occurrences of <date1> and subsequent dates

#### **C.2.2.2.5.2 Range Matching of Attributes of VR of TM**

**All comparison specified in the following shall be based on a direct comparison of times within a day. "Prior" includes all times starting from midnight of the same day to the specified time. "Subsequent" includes all times starting with the specified time until any time prior to midnight of the following day. Range matching crossing midnight is not supported.**

**No offset from Universal Coordinated Time is permitted in the TM VR values. If Timezone Offset From UTC (0008,0201) is present in the query identifier, the specified time values and the definition of midnight are in the specified timezone.**

In the absence of extended negotiation, ~~if the Attribute is a time,~~ then:

- a. A string of the form "<time1> - <time2>", where <time1> is less or equal to <time2>, shall match all occurrences of times that fall between <time1> and <time2> inclusive
- b. A string of the form "- <time1>" shall match all occurrences of times prior to and including <time1>
- c. A string of the form "<time1> -" shall match all occurrences of <time1> and subsequent times

#### **C.2.2.2.5.3 Range Matching of Attributes of VR of DT**

**If the Attribute is a datetime, then:**

- a. A string of the form "<datetime1> - <datetime2>", where <datetime1> is less or equal to <datetime2>, shall match all moments in time that fall between <datetime1> and <datetime2> inclusive
- b. A string of the form "- <datetime1>" shall match all moments in time prior to and including <datetime1>
- c. A string of the form "<datetime1> -" shall match all moments in time subsequent to and including <datetime1>
- d. The offset from Universal Coordinated Time, if present in the Value of the Attribute, shall be taken into account for the purposes of the match.

#### **C.2.2.2.5.4 Range Matching General Rules**

If extended negotiation of combined date-time matching is successful, then a pair of Attributes that are of **VR of DA** ~~date~~ and **TM** ~~time~~, both of which specify the same form of range matching, shall have the concatenated string values of each range matching component matched as if they were a single ~~datetime~~ Attribute **of VR of DT**.

#### Note

For example, a Study Date of "20060705-20060707" and a Study Time of "1000-1800" will match the time period of July 5, 10am until July 7, 6pm, rather than the three time periods of 10am until 6pm on each of July 5, July 6 and July 7, as would be the case without extended negotiation.

Regardless of other extended negotiation, an application may use the value of Timezone Offset From UTC (0008,0201) to adjust values of ~~time and date~~ time and date Attributes of VR TM and DT from the local timezone to UTC for matching. See Section C.2.2.2.1.

#### Note

If extended negotiation of combined date-time matching is successful, the timezone offset may effect a change in date if the local time and UTC are on different sides of midnight.

Range matching is not defined for types of Attributes other than dates and times.

### **C.2.2.2.6 Sequence Matching**

If a Key Attribute in the Identifier of a C-FIND request needs to be matched against an Attribute structured as a Sequence of Items (~~Value Representation~~ VR of ~~Type~~ SQ), the Key Attribute shall be structured as a Sequence of Items with a single Item. This Item may contain zero or more Item Key Attributes. Each Item Key Attribute matching shall be performed on an Item by Item basis. The types of matching defined in Section C.2.2.2 shall be used: Single Value Matching, List of UID Matching, Universal Matching, Wild Card Matching, Range Matching and Sequence Matching (recursive Sequence matching).

If all the Item Key Attributes match, for at least one of the Items of the Attribute against which the match is performed, a successful match is generated. A sequence of matching Items containing only the requested Attributes is returned in the corresponding C-FIND responses.

If the Key Attribute in the Identifier of a C-FIND request contains no Key Item Attribute (zero-length Item Tag), then all entities shall match this Attribute. This provides a universal matching like mechanism to request that the selected Key Attribute value (the entire Sequence of Items) be returned in corresponding C-FIND responses.