Ambiguity in the UMLS® Metathesaurus®

1999 Edition

Alan R. Aronson

March 11, 1999

1. Introduction

It appears that the number of Metathesaurus terms with large degree of explicit ambiguity is increasing. The purpose of this study is to get a rough feel of the extent of the problem (see sections 2 and 3) and to propose short- and long-term solutions to the problem (section 4).

2. The Problem

English Metathesaurus strings with the largest number of ambiguity designators are listed below. Each case is preceded by the number of ways ambiguous the string is.

- 54 'Other <n>'
- 23 'Protocols <n>'
- 18 'Patient Education Plans <n>'
- 16 'Assessment <n>'
- 16 'Limited function/disability <n>'
- 8 'Driver injured in collision with other and unspecified motor vehicles in nontraffic accident <n>'
- 8 'Driver injured in collision with other and unspecified motor vehicles in traffic accident <n>'
- 8 'cde genotype <n>'
- 7 'Passenger injured in collision with other and unspecified motor vehicles in nontraffic accident <n>'
- 7 'Passenger injured in collision with other and unspecified motor vehicles in traffic accident <n>'
- 6 'Other injuries <n>'
- 5 'Cold <n>'
- 2 'FEELING COLD <n>' (just <4> and <5>)
- 5 'Injuries <n>'
- 5 'Premolar tooth <n>'

- 5 'Tourniquet on <n>'
- 5 '[SO] Premolar tooth <n>'
- 5 'adjustment <n>'
- 5 'cd <n>'
- 5 'prostate <n>'
- 3 'sound measurement <n>' (just <3>, <4> and <5>)
- 5 'urethra <n>'
- 4 'Cl <n>'
- 4 'conjunctiva <n>'
- 4 'cornea <n>'
- 2 'Dihydropyridine <n>' (just <3> and <4>)
- 4 'ET - Esotropia <n>'
- 4 'HRF <n>'
- 4 'U <n>'
- 4 'lupus <n>'
- 4 'ms <n>'
- ...

Ignoring case, any concept with a string 'aaa <n>' also has string 'aaa' (without the ambiguity designator). The problem arises in any application which gains access to the Metathesaurus textually, i.e., through the strings comprising Metathesaurus concepts. Any occurrence of *other*, for example, can mean any of the 54 concepts associated with the 'Other <n>' strings. This includes the concepts 'Other' (the *legitimate* one), 'Other location of complaint', 'Other activities involving preparation of a routine non-injectable drug product', 'Other Intervention Target', and 'other Neglected child and/or adult sign or symptom'. The text *other* becomes 54 ways ambiguous. This significantly exacerbates the already difficult problem of disambiguation. There is also a practical consequence of the ambiguity. MetaMap encounters a combinatorial explosion any time the text it is processing refers to multiple concepts and includes the word *other*. For that reason, I have manually eliminated the 'Other <n>' strings for several years.

3. Analysis

The ambiguous strings listed in the previous section are considered in detail here. In each case it is the source of ambiguity that is of interest.

3.1 'Other <n>'

The source distribution of the 54 'Other' strings is:

- 1 SNMI98 (or LNC10J)
- 1 AIR93
- 49 OMS94
- 3 PPAC98

MTH strings are not included here; I am interested in the original source(s) of the strings. For supporting data for these strings, see appendix 5.1.

The SNMI98 string is part of the concept 'Other' which also contains the LNC10J string 'OTHER'. In SNMI98, Other's parent is 'GENERAL ADJECTIVAL MODIFIERS'. This is clearly a legitimate use of the string as a simple modifier.

The four concepts from AIR93 and PPAC98 are called 'Other' within their vocabulary but have expanded MTH strings:

- 'Other location of complaint' (AIR93)
- 'Other activities involving preparation of a routine non-injectable drug product' (PPAC98)
- 'Other activities involving preparation of compounded non-parenteral medications' (PPAC98)
- 'Other activities involving preparation of compounded parenteral medications' (PPAC98)

In each of the PPAC98 cases, the string 'Other' occurs hierarchically below another string giving it context. For example, the part of the PPAC98 hierarch giving rise to the first PPAC98 example above is

       'Prepare a routine non-injectable drug product'
           'Other'

The 49 OMS94 concepts are of the form 'Other aaa'. Examples include

- 'Other Intervention Target'
- 'Other Environmental problem'
- 'Other Psychosocial problem'
- 'other Income deficit'
- 'other Neglected child and/or adult sign or symptom'

Most of the OMS94 concepts arise from a hierarchical context as with PPAC98. Thus 'Other Environmental problem' arises from

       'DOMAIN I. ENVIRONMENTAL'
           'Other'

the hierarchy for 'other Income deficit' is

       'DOMAIN I. ENVIRONMENTAL'
           'Income'
               'Deficit'
                   'other'

and similarly the hierarchy for 'other Sanitation deficit' is

       'DOMAIN I. ENVIRONMENTAL'
           'Sanitation'
               'Deficit'
                   'other'

The source of the ambiguity here is that within a vocabulary, the term 'Other' appears in a hierarchical context (I could not verify this for the AIR93 case). The Metathesaurus concept for this term contains a context-independent, expanded term but also retains the simple term 'Other'. This phenomenon appears repeatedly. Terms like 'Other' will be referred to as *contextual terms* or *contextual strings* when they acquire their full meaning only through the context(s) in which they appear in their vocabulary.

3.2 'Protocols <n>'

The source distribution of the 23 'Protocols' strings is:

-   1 RCD98 (or CSP98)
- 22 PCDS97

(MTH strings are not included.)

The singleton case is 'Protocols <2>' with strings 'Protocols' in RCD98 and 'protocol' in CSP98.

The 22 cases from PCDS97 have concept names of the form 'Protocols: aaa':

- 'Protocols: Activities'
- 'Protocols: Pre- or Intra- or Post-Procedure'
- 'Protocols: Urinary Elimination'
- 'Protocols: Tissue Perfusion'

They are entirely analogous to the OMS94 'Other <n>' terms. Thus 'Protocols: Activities' arises from the hierarchy

       ACTIVITY COMPONENT
         Protocols

Similar analyses apply to

- 18 'Patient Education Plans <n>' (source PCDS97)
- 16 'Assessment <n>' (one string from MSH99 and 15 from PCDS97)

3.3 'Limited function/disability <n>'

The 16 'Limited function/disability' strings all originate in ICPC93. They are all of the form 'Limited aaa function' with synonym 'Limited aaa function/disability'. In addition they all have MTH strings 'Limited function/disability <n>' and 'Limited function/disability' both of which are suppressible synonyms. Thus the ambiguity in this case can be effectively eliminated.

Similar analyses apply to

-   6 'Other injuries <n>'

3.4 'Driver injured in collision with other and unspecified motor vehicles in nontraffic accident <n>'

The source distribution of the 8 'Driver injured …' strings is:

- 1 RCD98
- 7 ICD10

(MTH strings are not included.)

In all cases RCD98 and ICD10 have string 'Driver injured in collision with other and unspecified motor vehicles in nontraffic accident'. In RCD98 the string means what it says. But in ICD10 the strings are contextual but not as straightforwardly so as in previous examples. For instance, the concept 'Driver pedal cycle injured in collision with other and unspecified motor vehicles in non-traffic accident' is hierarchically represented in ICD10 as

'Pedal cyclist injured in transport accident'
'Pedal cyclist injured in other and unspecified transport accidents'
'Driver injured in collision with other and unspecified motor vehicles in nontraffic accident'

For ICD10 the ambiguous string is a generalization of the concept since *driver* generalizes *pedal cyclist*. But it is also more specific due to *nontraffic accident* which is not mentioned in the hierarchy above the string. Nevertheless, this represents another instance of contextual strings.

Similar analyses apply to

- 8 'Driver injured in collision with other and unspecified motor vehicles in traffic accident <n>'
- 7 'Passenger injured in collision with other and unspecified motor vehicles in nontraffic accident <n>'
- 7 'Passenger injured in collision with other and unspecified motor vehicles in traffic accident <n>'

3.5 'cde genotype <n>'

The 8 'cde genotype' strings are in RCD98. The variation is obtained by combining all upper and lower case combinations of the letters 'c', 'd' and 'e'. The strings are all of type 'Finding' and are children of 'Rh genotype' ('Laboratory or Test Result'). It appears that the 8 strings are legitimate and distinct. This is also a good example where preserving case is vital to distinguishing the concepts. It seems that adding a string (replicating the case) with an ambiguity designator is at least confusing whether a system using the Metathesaurus distinguishes case or not. See 5.2 for all MRCONSO entries.

3.6 'Cold <n>'

The concepts associated with the five senses of 'Cold' are

- 'cold temperature'
- 'Common Cold'
- 'Cold Therapy'
- 'Chronic Obstructive Airway Disease' (which has strings 'COLD' and 'Chronic Obstructive Lung Disease'
- 'Cold Sensation'

Each of the five senses seems legitimate with the possible exception of 'Chronic Obstructive Airway Disease'. That sense arises out of the acronym COLD which can usually be distinguished in text by its form consisting of all capital letters. But since text is not always mixed case and even when it is, authors sometimes use all lowercase letters for acronyms, it does no great harm to include it as a sense of 'Cold'.

3.7 'Injuries <n>'

The five 'Injuries' concepts are

- 'injuries <1>' (which is the MSH99 subheading 'injuries')
- 'Injury inflicted to the body by an external force' (with string 'Injury' from several sources)

- 'urologic injuries'
- 'Female genital injuries'
- 'Male Genital Injuries'

The 'Injuries <n>' and 'Injuries' strings for the last three concepts are contextual and suppressible leaving only two senses of 'Injuries'.

3.8 'Premolar tooth <n>'

The source distribution of the 5 'Premolar tooth' strings is:

- 1 MSH99 and several other sources including RCD98
- 4 RCD98

(MTH strings are not included.)

The concepts are

- 'Bicuspid'
- 'Permanent upper right first premolar tooth'
- 'Permanent upper left second premolar tooth'
- 'Permanent upper left first premolar tooth'
- 'Permanent lower right second premolar tooth'

Each of the concepts has the string RCD98 '[SO] Premolar tooth' which is the source of the ambiguity. Each of the concepts seems to be a premolar tooth, but that does not generally signal true ambiguity.

Similar analyses apply to

- 5 '[SO] Premolar tooth <n>'

3.9 'Tourniquet on <n>'

The five 'Tourniquet on' concepts are

- 'Application of tourniquet'
- 'Upper tourniquet cuff inflation'
- 'Lower tourniquet cuff inflation'
- 'Manual tourniquet application'
- 'Tourniquet cuff inflation'

RCD98 is the sole source for all but the first concept. There does not seem to be any obvious reason for the ambiguity.

3.10 'adjustment <n>'

The five 'adjustment' concepts are

- 'Clinical Adjustment' (with strings 'Adjustment, NOS' and 'Adjustments')
- 'Individual Adjustment'
- 'Adjustment Action'
- 'Adjustment <4>' (with string 'Adjustment')
- 'Psychological adjustment'

Strings from various vocabularies argue strongly that all five concepts are ambiguous.

3.11 'cd <n>

The five 'cd' concepts are

- 'Cadmium'
- 'CP protocol'
- 'Compact Disks'
- 'Clinical diagnosis'
- 'candela'

It is reasonable but not clear why 'CP protocol' is in this group. (PDQ98 has strings 'CD', 'CDDP/CTX' and 'cisplatin/cyclophosphamide' for the concept.)

3.12 'prostate <n>'

The five 'prostate' concepts are

- 'Prostate'
- 'Prostatic Diseases'
- 'Benign neoplasm of prostate'
- 'Carcinoma in situ of prostate'
- 'Neoplasm of uncertain or unknown behavior of prostate'

CST95 has the following hierarchy for 'Prostatic Diseases'

> Male Genital Disorders
>> prostate

And ICD10 has similar contextual strings for the last three concepts. Thus all concepts except 'Prostate' are contextual. (Actually, the last three cases are marked suppressible.)

Similar analyses apply to

-    5 'urethra <n>'
-    4 'conjunctiva <n>'
-    4 'cornea <n>'

3.13 'sound measurement <n>'

The three 'sound measurement' concepts are

- 'Ultrasonography'
- 'heart sonography'
- 'sound measurement <5>' (with CSP98 string 'skeletal sonography')

All three concepts have CSP98 string 'sound measurement'.

3.14 'Cl <n>'

The four 'cl' concepts are

- 'Chlorine'
- 'Cycloleucine'

- 'centiliter'
- 'Chloride Ion'

Each seems to be a legitimate 'cl' concept.

3.15 'Dihydropyridine <n>'

The two 'Dihydropyridine' concepts are

- 'Dihydropyridines'
- 'dihydropyridine'

Because the concepts are inflectional variants, the ambiguity is harmless. (I believe that there are similar concepts such as 'Chlorides' and 'Chloride' which are not marked with ambiguity designators.)

3.16 'ET - Esotropia <n>'

The four 'ET - Esotropia' concepts are

- 'Esotropia'
- 'Manifest alternating convergent squint'
- 'Intermittent convergent squint'
- 'Incomitant esotropia'

None of the source vocabularies has string 'ET - Esotropia <n>' or 'ET - Esotropia'; they are all MTH strings. There does not seem to be any textual basis for this case of ambiguity.

3.17 'HRF <n>'

The four 'HRF' concepts are

- 'Hypothalamic Releasing Factor'
- 'Gonadorelin'
- 'homologous restriction factor'
- 'HRF <3>' (with RCD98 string 'HRF')

Each concept seems to be legitimate although it is interesting that none of the 75 distinct strings for 'Gonadorelin' (except for 'HRF <1>' and 'HRF', of course) are obvious expansions of the acronym HRF.

3.18 'U <n>'

The four 'U' concepts are

- 'Iodine'
- 'Uranium'
- 'Lower case you'
- 'Unit'

The string 'U' for 'Iodine' occurs in LNC10J as does the string 'I'. Both 'u' for 'Lower case you' and 'U' for 'Unit' occur in RCD98. 'Iodine' may be incorrect and 'Lower case you' may be annoying, but this still appears to be legitimate ambiguity.

3.19 'lupus <n>'

The four 'lupus' concepts are

* 'Lupus Vulgaris'
* 'Lupus Erythematosus, Discoid'
* 'Lupus Erythematosus, Systemic'
* 'Lupus Erythematosus'

Each concept has a string 'Lupus' or 'lupus' (from a wide variety of sources). This seems to be a paradigmatic case of ambiguity (at least from one non-medical person's perspective).

3.20 'ms <n>'

The four 'ms' concepts are

* 'Mitral Valve Stenosis' (with string 'Mitral Stenosis')
* 'Multiple Sclerosis'
* 'Morphine Sulfate'
* 'millisecond'

Each seems to be a legitimate 'ms' concept although this and other examples may signal the beginning of acronym proliferation. Acronyms are useful only in small quantities. As their number grows, they have much the same effect as ambiguity. And since authors of journal articles are fairly conscientious about defining their acronyms, their presence in the Metathesaurus is questionable.

4. Conclusions and Recommendations

The most common source of ambiguity arises from vocabularies which have contextual strings. In many cases, the UMLS staff has already provided an expanded form of the string which does not rely on context for understanding. In addition some of the contextual strings, themselves, have been marked as suppressible synonyms. This combination of steps seems to be a good way to solve the ambiguity problem. As odious as it may be to correct the problems caused by contextual strings, I think it allows for a semantically more correct incorporation of vocabularies into the Metathesaurus, and it provides a useful service to the community of Metathesaurus users.

Another source of ambiguity arises directly from a constituent vocabulary. RCD98, for example as four terms (which I believe are not ambiguous) each of which has the variant '[SO] Premolar tooth'. This may be a way of aggregating terms into more general terms, but it still causes an unnecessary degree of ambiguity. Making these variants suppressible would solve the problem here, too. (This task is probably more offensive than that of contextual string, but it probably occurs less frequently.)

The 'cde genotype <n>' example is interesting because the ambiguity arises only when case is ignored. We may be close to the time when Metathesaurus users can no longer blithely ignore case differences.

Finally, I only reviewed all cases of ambiguity with degree five or more. The less ambiguous cases also merit attention.

5. Appendices

Note that examples are given in this section consist of lines from a file, MRCONSO, joining
MRCON and MRSO so that source information is explicitly shown.

5.1 Some MRCONSO entries for 'Other <n>'

C0205394:L0249050:S0324820|381465|P|PF|Other|SNMI98|PT|G-A609
C0220886:L0249050:S0324820|397632|S|VO|Other|AIR93|HT|U000081 (see MTH string below)
C0237094:L0308671:S0392468|428733|P|PF|Other Intervention Target|OMS94|MT|NOCODE
C0237098:L0299351:S0392464|428740|P|PF|Other Environmental problem|OMS94|MT|NOCODE
C0237111:L0314149:S0392472|428757|P|PF|Other Psychosocial problem|OMS94|MT|NOCODE
C0237120:L0314148:S0392471|428774|P|PF|Other Physiological problem|OMS94|MT|NOCODE
C0237127:L0285311:S0392467|428792|P|PF|Other Health Related Behavior problem|OMS94|MT|NOCODE
C0237133:L0295307:S0423097|428801|P|PF|other Income deficit|OMS94|MT|NOCODE
C0237142:L0295314:S0423113|428812|P|PF|other Sanitation deficit|OMS94|MT|NOCODE
C0237155:L0295313:S0423110|428832|P|PF|other Residence deficit|OMS94|MT|NOCODE
C0237161:L0280305:S0423101|428840|P|PF|other Neighborhood and/or workplace safety defi-
        cit|OMS94|MT|NOCODE
C0237162:L0295303:S0423088|428843|P|PF|other Environmental deficit|OMS94|MT|NOCODE
C0237169:L0292491:S0423083|428854|P|PF|other Communication with community resources impair-
        ment|OMS94|MT|NOCODE
C0237173:L0293062:S0423115|428861|P|PF|other Social contact impairment|OMS94|MT|NOCODE
C0237178:L0290855:S0423112|428870|P|PF|other Role change impairment|OMS94|MT|NOCODE
C0237185:L0307778:S0423099|428879|P|PF|other Interpersonal relationship impairment|OMS94|MT|NOCODE
C0237190:L0314008:S0423117|428887|P|PF|other Spirituality sign or symptom|OMS94|MT|NOCODE
C0237195:L0305512:S0423091|428895|P|PF|other Grief impairment|OMS94|MT|NOCODE
C0237206:L0299026:S0423087|428909|P|PF|other Emotional stability impairment|OMS94|MT|NOCODE
C0237212:L0306917:S0423096|428917|P|PF|other Human sexuality impairment|OMS94|MT|NOCODE
C0237221:L0280288:S0423080|428932|P|PF|other Caretaking and/or parenting impairment|OMS94|MT|NOCODE
C0237228:L0278437:S0423100|428945|P|PF|other Neglected child and/or adult sign or symp-
        tom|OMS94|MT|NOCODE
C0237237:L0276944:S0423077|428960|P|PF|other Abused child and/or adult sign or symp-
        tom|OMS94|MT|NOCODE
C0237242:L0295864:S0423092|428967|P|PF|other Growth and development impairment|OMS94|MT|NOCODE
C0237243:L0307793:S0423109|428970|P|PF|other Psychosocial impairment|OMS94|MT|NOCODE
C0237247:L0306089:S0423095|428976|P|PF|other Hearing impairment|OMS94|MT|NOCODE
C0237255:L0307795:S0423120|428986|P|PF|other Vision impairment|OMS94|MT|NOCODE
C0237262:L0307782:S0423116|428996|P|PF|other Speech and language impairment|OMS94|MT|NOCODE
C0237266:L0295589:S0423085|429001|P|PF|other Dentition impairment|OMS94|MT|NOCODE
C0237276:L0292137:S0423082|429014|P|PF|other Cognition impairment|OMS94|MT|NOCODE
C0237283:L0314007:S0423104|429024|P|PF|other Pain sign or symptom|OMS94|MT|NOCODE
C0237285:L0293015:S0423084|429030|P|PF|other Consciousness impairment|OMS94|MT|NOCODE
C0237289:L0307777:S0423098|429037|P|PF|other Integument impairment|OMS94|MT|NOCODE
C0237300:L0303091:S0423102|429050|P|PF|other Neuro-musculo-skeletal function impair-
        ment|OMS94|MT|NOCODE
C0237305:L0307794:S0423111|429060|P|PF|other Respiration impairment|OMS94|MT|NOCODE
C0237318:L0291550:S0423081|429078|P|PF|other Circulation impairment|OMS94|MT|NOCODE
C0237324:L0296239:S0423086|429086|P|PF|other Digestion-hydration impairment|OMS94|MT|NOCODE
C0237330:L0287310:S0423079|429109|P|PF|other Bowel function impairment|OMS94|MT|NOCODE
C0237338:L0303078:S0423090|429118|P|PF|other Genito-urinary function impairment|OMS94|MT|NOCODE
C0237343:L0280843:S0423078|429126|P|PF|other Antepartum or postpartum impairment|OMS94|MT|NOCODE
C0237344:L0307791:S0423107|429129|P|PF|other Physiological impairment|OMS94|MT|NOCODE

C0237353:L0307787:S0423103|429144|P|PF|other Nutrition impairment|OMS94|MT|NOCODE
C0237357:L0307790:S0423114|429152|P|PF|other Sleep and rest patterns impairment|OMS94|MT|NOCODE
C0237361:L0277863:S0423106|429158|P|PF|other Physical activity impairment|OMS94|MT|NOCODE
C0237366:L0307087:S0423105|429165|P|PF|other Personal hygiene impairment|OMS94|MT|NOCODE
C0237369:L0314009:S0423118|429171|P|PF|other Substance use sign or symptom|OMS94|MT|NOCODE
C0237373:L0301028:S0423089|429178|P|PF|other Family planning impairment|OMS94|MT|NOCODE
C0237380:L0289427:S0423094|429190|P|PF|other Health care supervision impairment|OMS94|MT|NOCODE
C0237387:L0307783:S0423108|429203|P|PF|other Prescribed medication regime impairment|OMS94|MT|NOCODE
C0237394:L0307792:S0423119|429217|P|PF|other Technical procedure impairment|OMS94|MT|NOCODE
C0237395:L0285310:S0423093|429220|P|PF|other Health Related Behavior impairment|OMS94|MT|NOCODE
C0700060:L0249050:S0324820|1162406|S|VO|Other|PPAC98|ST|158 (see MTH string below)
C0700062:L0249050:S0324820|1162412|S|VO|Other|PPAC98|ST|166 (see MTH string below)
C0700063:L0249050:S0324820|1162415|S|VO|Other|PPAC98|ST|178 (see MTH string below)

MTH strings for those concepts whose only non-Metathesaurus string is 'Other':

C0220886:L0292537:S0392491|397630|P|PF|Other location of complaint|MTH|PN|U000029
C0700060:L1221829:S1465080|1162404|P|PF|Other activities involving preparation of a routine non-injectable drug product|MTH|PN|NOCODE
C0700062:L1221830:S1465081|1162410|P|PF|Other activities involving preparation of compounded non-parenteral medications|MTH|PN|NOCODE
C0700063:L1221828:S1465079|1162413|P|PF|Other activities involving preparation of compounded parenteral medications|MTH|PN|NOCODE

5.2 All MRCONSO entries for 'cde genotype <n>'

C0427651:L0776216:S0826898|752292|P|PF|CDE genotype|MTH|PN|NOCODE
C0427651:L0776216:S0826898|752292|P|PF|CDE genotype|RCD98|PT|X76xB
C0427651:L0776216:S0826899|752293|P|VO|CDE genotype <1>|MTH|MM|U002573
C0427652:L0776216:S0826941|752294|P|PF|CDe genotype|MTH|PN|NOCODE
C0427652:L0776216:S0826941|752294|P|PF|CDe genotype|RCD98|PT|X76xC
C0427652:L0776216:S0826942|752295|P|VO|CDe genotype <2>|MTH|MM|U002574
C0427653:L0776216:S0834239|752296|P|PF|CdE genotype|MTH|PN|NOCODE
C0427653:L0776216:S0834239|752296|P|PF|CdE genotype|RCD98|PT|X76xD
C0427653:L0776216:S0834240|752297|P|VO|CdE genotype <3>|MTH|MM|U002575
C0427654:L0776216:S0834250|752298|P|PF|Cde genotype|MTH|PN|NOCODE
C0427654:L0776216:S0834250|752298|P|PF|Cde genotype|RCD98|PT|X76xE
C0427654:L0776216:S0834251|752299|P|VO|Cde genotype <4>|MTH|MM|U002576
C0427655:L0776216:S1084641|752300|P|PF|cDE genotype|MTH|PN|NOCODE
C0427655:L0776216:S1084641|752300|P|PF|cDE genotype|RCD98|PT|X76xF
C0427655:L0776216:S1084642|752301|P|VO|cDE genotype <5>|MTH|MM|U002577
C0427656:L0776216:S1084643|752302|P|PF|cDe genotype|MTH|PN|NOCODE
C0427656:L0776216:S1084643|752302|P|PF|cDe genotype|RCD98|PT|X76xG
C0427656:L0776216:S1084644|752303|P|VO|cDe genotype <6>|MTH|MM|U002578
C0427657:L0776216:S1084797|752304|P|PF|cdE genotype|MTH|PN|NOCODE
C0427657:L0776216:S1084797|752304|P|PF|cdE genotype|RCD98|PT|X76xH
C0427657:L0776216:S1084798|752305|P|VO|cdE genotype <7>|MTH|MM|U002579
C0427658:L0776216:S1084807|752306|P|PF|cde genotype|MTH|PN|NOCODE
C0427658:L0776216:S1084807|752306|P|PF|cde genotype|RCD98|PT|X76xI
C0427658:L0776216:S1084808|752307|P|VO|cde genotype <8>|MTH|MM|U002580