**Rocklin Site Tissue Data Syntax Guide**

**Qualified Tissue Inventory (Active Inventory)**

Columns:

* Block ID #
  + **Format**: “TTT-#SM”.
    - TTT is the tissue type abbreviation (this can be longer than 3 letters, have spaces, and/or multiple dashes)
    - # is the patient number (this should be only integers)
    - S is the sample number, present if a sample of 1 patient was split into multiple blocks
    - M (if present) indicates that the sample from one patient was further sub-split into multiple blocks by the CM team.
  + **Example**: APX-32A1
    - APX – Appendix
    - 32 – patient identifier
    - A – sample from that patient
    - 1 – indicating the first of a further sub-split block
* Negative Antibody Scores
  + **Format**: “UVNegative:0, Antibody(Clone),
    - Different antibodies should be separated by commas.
    - Same antibody but different clones should be separated by semicolons.
  + **Examples**:
    - UVNegative: 0, Adipophilin(polyclonal), CK 19(A53-B/A2.26), Ep-CAM(Ber-EP4): 0/0
    - UVNegative: 0, Calretinin(polyclonal)(SP13): 0/0, HMB-45(HMB-45): 0/0
      * Sometimes there will be multiple parentheses associated with one antibody scoring
        + **This makes it very difficult to split apart data**
        + **Is there another way to format the additional clone info?**
* Positive Antibody Scores
  + **Format**: “Antibody: Score(Clone)”
    - Different antibodies should be separated by commas.
    - Same antibody but different clones should be separated by semicolons.
  + **Examples**:
    - Inhibin: 3.5/0(R1), Synaptophysin: 3.5/0(MRQ-40)
    - Calretinin: 0.5/0(polyclonal); 0.5/0(SP13)
    - Ki-67: 4/0(10%)(SP6), p53: 3/1(DO7); 3/0(SP5), GFAP: 3.5/0(SP78); 4/0(EP672Y), Vimentin: 3.5/0(EP21); 3/0(SP20), 1/0(V9)
      * We see again that some tests will have multiple parentheses around them
        + If it is a % value, it represents the percentage of cells stained
        + **Difficulty arises here. The information contained in parentheses can have a range of meanings, making extraction difficult. How can we efficiently change this?**
    - Ki-67: 3.5/0F(20%)(SP6), p53: 3.5/1.5(F(DO7); 3/0F(SP5), Olig2: 3.5/0F(211F1.1); 3.5/0F(EP112), GFAP: 4/0(SP78); 4/0(EP672Y)
      * We might also see a letter after the score.
        + This is a way of communicating staining pattern, for example “F” meaning focal.
        + The yellow examples are ways this should be done, red shows an example of an incorrect way
* Date Added
  + **Format**: DD-Mmm-YYYY
    - Mmm is the 3-letter abbreviation for each month
  + **Example**: 21-Jan-2022
* Initials
  + **Format**: Capital letters
    - Ampersand “&” to identify multiple names is fine.
  + **Example**:
    - KMC
    - CEJ & KMC
* Pathologist Name
  + **Format**: Pathologist Name
  + **Example**: Dr. Su
* Diagnostic Comments
  + Pathology of the sample. Normal/Diseased and more specification.
  + **Format:** Text entry
  + **Example:** Glioblastoma, involving cerebellar tissue and with prominent inflammatory reaction.
  + This will go directly as text into LIMS and does not need to be changed.
* Other Comments
  + Contains a range of info. Most important to know: “Same Case as Another Block ID”. This is relevant for patient tracing. Or may say “Dx change” or “Changed from…”. This indicates the diagnosis was changed.
  + **Format:** 
    - “Same Case As”
    - “Split From”
    - **Two different information groups must and should be split by semicolons or periods only.**
  + **Example:** 
    - “Same case as CVX-102, CIN I-41, UTS-52”
    - Split from CVX-102F; Same case as CVX-102, CIN I-41, UTS-52
    - “TYM-1/2/19Androgen Receptore from one patient; Taken for Rabbit 4 Phage Display Block-KMC 10/30/13; returned 1/5/16 KMC”
  + **There is lots of extraneous information here. What can we do to understand the other comments?**
  + **We have to split this into “Other Aliases” for LIMS, grabbing those related block IDs. To do this: splitting, same case, or renaming are the information types we need to glean. Where does the other info go?**
* Previous Block ID
  + Only present if there was a Dx change. Same format as Block ID #
  + **Format**: “TTT-#SM”.
    - TTT is the tissue type abbreviation (this can be longer than 3 letters, have spaces, and/or multiple dashes)
    - # is the patient number (this should be only integers)
    - S is the sample number, present if a sample of 1 patient was split into multiple blocks
    - M (if present) indicates that the sample from one patient was further sub-split into multiple blocks by the CM team.
  + **Example**: APX-32A1
    - APX – Appendix
    - 32 – patient identifier
    - A – sample from that patient
    - 1 – indicating the first of a further sub-split block
* Date
  + Only present if there was a Dx change. Same format as Date Added
  + **Format**: DD-Mmm-YYYY
    - Mmm is the 3-letter abbreviation for each month
  + **Example**: 21-Jan-2022
* Rename Initials
  + Only present if there was a Dx change. Same format as Initials
  + **Format**: Initials, capital letters
    - Ampersand “&” to identify multiple names is fine.
  + **Example**:
    - KMC
    - CEJ & KMC

**Customer Blocks**

Columns:

* Receiving Date
  + **Format**: DD MMM YYYY
    - Where MMM is the 3-letter month abbreviation
  + **Example**: 31 AUG 2021
* Sales Rep
  + **Format**: Name
  + **Example**:
    - Ourhay
    - Liz
* Expectation
  + Text column
  + **Format:**
    - “$ per block”
    - “Block Credit”
    - “Double Credit”
  + **Example:**
    - $25 per block
    - “Block Credit”
    - “Double Credit”
* # of Blocks
  + **Format**: Number
  + **Example**: 7
* Received by
  + **Format**: Initials, capital letters
    - Ampersand “&” to identify multiple names is fine.
  + **Example**:
    - KMC
    - CEJ & KMC
* Tissue type
  + **Format**: Tissue type
  + **Example**: Squamous Cell Carcinoma
* Keepers: CM ID#
  + **Format**: TTT-#F-L
    - TTT is the tissue type abbreviation.
    - # is the patient number.
    - B is the first sample number that was qualified
    - E is the last sample number that was qualified
    - Multiple sets should be separated via commas
  + **Example**:
    - MEL-194A-E
    - GB-27, BWL-37, FL-60A-C, LN-98
    - CSC-10/11
      * Use - instead of /
    - CIN II-24&25
      * Use - instead of &
    - AA-1A, C-G
      * Ambiguous where C-G belong. Instead: AA-1A, AA-1C-G
* Return, Discard, or Credit date & Initial
  + **Format**: What happened with how many blocks on which day by whom
    - “Awarded”, “Discarded”, or “Returned” is the language that should be used.
    - Can include who was notified
    - Separate a different outcome of blocks (return/discard/credit) via semicolon
    - Dates should always use slash “/” not dash “-”
  + **Example**:
    - Awarded 10 block credits 10/24/16 CEJ
    - 3 blocks discarded 9/18/15; Awarded 7 credits 10/30/15 and notified Faris
    - Kept All Notified Ourhay and Nicole 3/29/12
      * It is not clear what happened to these blocks. We need to use the keyword “Awarded” or “Credit” when awarding block credit.
* Comments/Customer Block ID #
  + **Format**: Not always consistent.
    - Try to only use dashes when indicating a list of blocks
    - Every block ID must be separated by a comma or dash
    - Try to avoid parentheses
  + **Example**:
    - Melanoma 1-5
    - 04S 9192A1, A2, A4, A5, A6
      * This is not ideal because we don’t know that “A2” is not a full-length ID. This should be: 04S 9192A1-6
    - Liposarcoma (1-23)
      * Parentheses add an unnecessary level of complexity. This should be: Liposarcoma 1-23
* Demographic Info
  + **Format**: Age Sex
  + **Example**: 28 F
* Location
  + **Format**: Location
  + **Example**:
    - H&E sheet H
    - Blocks + H&Es - Sheet C

**Closet Blocks**

* Surgical Number
  + **Format:** Dependent on where it was received from
  + **Example:** 
    - P06-6950 (Princeton Baptist)
    - 01-LV-0896 (Lakeview Hospital)
    - 90-ML-103-3 (NatPkMedCtr AMI HotSprings)
* Specimen
  + Tissue type/pathology
  + **Format:** Text
  + **Example:** Kidney
* Procedure
  + Surgical procedure the tissue came from
  + **Format:** Name of procedure
  + **Example:** Colectomy
* Total Number of Blocks
  + How many received in total
  + **Format:** #
  + **Example:** 27
* Number of Blocks used
  + How many were qualified
  + **Format:** #
  + **Example:** 5
* CM Number
  + **Format**: TTT-#F-L
    - TTT is the tissue type abbreviation.
    - # is the patient number.
    - F is the first sample number that was qualified
    - L is the last sample number that was qualified
    - Multiple sets should be separated via commas
  + **Example**:
    - UTC-174A-C
    - PHEO-15A-B, ADC-38C-E
    - CSC-10/11
      * Use - instead of /
    - CIN II-24&25
      * Use - instead of &
    - AA-1A, C-G
      * Ambiguous where C-G belong. Instead: AA-1A, AA-1C-G
* Initials/H&E Date
  + **Format:** Initials Date
    - Separate multiple assays by commas or semicolons
    - Always put the initials next to the data, even if it is the same person
    - Dates should always use slash “/” not dash “-”
  + **Example:** AY 6/15/15, AY 6/22/15
  + **Note:** This column tends to have a high variance in the way that data is input. It is difficult to separate this data out for LIMS integration. For this reason, we might have to refer to the legacy spreadsheet for this.
* Comment
  + Any additional info
  + It appears that most text in this column is to note discarding, which could be done in the next column. However, this column retains date and initials information.
  + Dates should always use slash “/” not dash “-”
  + **Format:** Any text
  + **Example:** Discarded 2 blocks 10/12/15 CEJ
* # Blocks Discarded
  + **Format:** #
  + **Example:** 2
* Location
  + Storage location
  + **Format:** Location or how many blocks in location
  + **Example:** 1 block in box 33