In vitro PDCM Minimal Information Standard - version 1.0

**Contact:** [Zinaida Perova](mailto:zina@ebi.ac.uk), Project Lead, [CancerModels.Org](https://www.cancermodels.org/)

Variability in creation, validation and description of patient-derived cancer models is a hurdle to reproducibility and clinical adoption of these models. [CancerModels.Org](https://www.cancermodels.org/) team aims to collect feedback from as many people working in this field as possible to assemble and publish a Minimal Information standard to describe these models, similar to what has been done for the patient-derived xenograft models ([PDX-MI](https://pubmed.ncbi.nlm.nih.gov/29092942/)).

Below is a collection of attributes to describe 3D patient-derived cancer models assembled in collaboration with the representatives from the following institutes: Baylor College of Medicine, Cornell University, DepMap - Broad Institute of MIT and Harvard, Dana-Farber Cancer Institute, Harvard, Huntsman Cancer Institute, MD Anderson Cancer Center, Sage Bionetworks, The Jackson Laboratory, UConn Health - University of Connecticut, University of California San Francisco, University Health Network - Princess Margaret Cancer Centre, University of Verona, University of Pittsburgh Medical Center, VHIO Vall d’Hebron Instituto de Oncología, Wellcome Sanger Institute

We are grateful for the time and effort you put into this work. If you have any feedback, or further suggestions, you can comment on the document here (**please use the “Add Comment” feature**): <https://tinyurl.com/cancermodelsMIdraft>

| **Attribute** | **Description** | **Example** | **Format** | **Required?** |
| --- | --- | --- | --- | --- |
| **patient\_id** | Unique anonymous/de-identified ID for the patient from the provider | A0088 | free alphanumerical and -\_~. | **essential** |
| **sex** | Sex of the patient | female | male,female,other,not collected, or not provided | **essential** |
| **history** | Cancer-relevant comorbidity or environmental exposure | lynch syndrome | free alphanumerical | **desirable** |
| **ethnicity** | Patient Ethnic group. Can be derived from self-assessment or genetic analysis. | Eastern European | [free text](https://www.ebi.ac.uk/ols/ontologies/ncit/terms?iri=http%3A%2F%2Fpurl.obolibrary.org%2Fobo%2FNCIT_C17049) | **desirable** |
| **ethnicity\_assessment\_method** | Patient Ethnic group assessment method | self-assessed/genetic | free alphanumerical. Leave blank if ethnicity was not collected/provided | **desirable** |
| **initial\_diagnosis** | Diagnosis of the patient when first diagnosed at age\_at\_initial\_diagnosis - this can be different than the diagnosis at the time of collection which is collected in the sample section. No acronyms. | colorectal carcinoma | [Use NCIT term when possible. Ctrl-Click here to see terms](https://www.ebi.ac.uk/ols/ontologies/ncit/terms?iri=http%3A%2F%2Fpurl.obolibrary.org%2Fobo%2FNCIT_C3262) | **desirable** |
| **age\_at\_initial\_diagnosis** | The age at first diagnosis. | 70 | Numerical. Can be exact age or binned in 10 year groups (e.g. 10-19) | **desirable** |
| **age\_category** | Age category at time of sampling | Adult | Adult, Pediatric, Fetus, Not collected, Not provided | **desirable** |
| **HLA\_type** | HLA type | HLA-A | free text | **desirable** |
| **smoking\_status** | Patient's smoking history | current smoker, 1 pack per day | Current smoker, ex-smoker, non-smoker, never smoked; free text | **desirable** |
| **alcohol\_status** | Alcohol intake of the patient, self-reported | yes | yes/no/not provided | **desirable** |
| **alcohol\_frequency** | The average number of days per week on which the patient consumes alcohol, self-reported | 3 days/week | free alphanumerical | **desirable** |
| **family\_history\_of\_cancer** | If a first-degree relative of the patient has been diagnosed with a cancer of the same or different type | yes | free text | **desirable** |
| **sample\_id** | Unique ID of the patient tumour sample used to generate the model | CRC0228PRH0000000000D01000 | unique, free alphanumerical and -\_~. | **essential** |
| **collection\_date** | Date of collections - if this is not available please indicate in column E, F collection events and the elapsed times between the collection. Important for understanding the time relationship between models generated for the same patient. | 42278 | MM YYYY or YYYY (depending on the regulations) | **if several collections this is essential otherwise desirable** |
| **collection\_event** | IF DATE ENTERED IN collection\_date DO NOT FILL order collection 1 correspond to the 1st time a patient was sampled to generate a PDX, subsequent collection events are incremented by 1 | collection 1 | "collection " + event number | **if several collections this is essential otherwise desirable** |
| **months\_since\_collection\_1** | IF DATE ENTERED IN collection\_date DO NOT FILL, provide the time difference between the 1st collection event and the current one (in months) - collection event 1 should be 0, collection event 2 should be 6 if 6 months have elapsed between collection 1 and collection 2 and collection event 3 should be 9 if 9 months have elapsed between collection 1 and collection 3 | 0 | numerical | **if several collections this is essential otherwise desirable** |
| **age\_in\_years\_at\_collection** | Patient age at tissue collection. Can be exact age or binned in 10 y group (1-9, 10-19, 20-29, ...) | 70 | numerical | **essential** |
| **diagnosis** | diagnosis at time of collection of the patient tumour used to generate the model | colorectal carcinoma | [Use NCIT ontology terms. Ctrl-Click here to see terms](https://www.ebi.ac.uk/ols/ontologies/ncit/terms?iri=http%3A%2F%2Fpurl.obolibrary.org%2Fobo%2FNCIT_C3262) | **essential** |
| **molecular\_subtype** | Molecular subtype of cancer | ER-positive | free alphanumerical | **desirable** |
| **tumour\_type** | Collected tumour type | primary | primary, metastatic, recurrent, refractory, precancerous | **essential** |
| **primary\_site** | Site of the primary tumor where primary cancer is originating from (may not correspond to the site of the current tissue sample). | colon | [Use NCIT ontology term where possible. Ctrl-click here to see terms.](https://www.ebi.ac.uk/ols/ontologies/ncit/terms?iri=http%3A%2F%2Fpurl.obolibrary.org%2Fobo%2FNCIT_C12219&viewMode=All&siblings=false) | **essential** |
| **collection\_site** | Site of collection of the tissue sample (can be different than the primary site if tumour type is metastatic) | colon | [Use NCIT ontology term where possible. Ctrl-click here to see terms](https://www.ebi.ac.uk/ols/ontologies/ncit/terms?iri=http%3A%2F%2Fpurl.obolibrary.org%2Fobo%2FNCIT_C12219&viewMode=All&siblings=false) | **essential** |
| **collection\_method** | Method of collection of the tissue sample | biopsy | Biopsy, surgical resection, blood draw, free text | **desirable** |
| **stage** | Stage of the patient cancer at the time of collection. Should be in the format of the stage classification | T2,N1,M0 | free alphanumerical | **desirable** |
| **staging\_system** | Stage classification system used to describe the stage. Add the version if available. | TNM staging system | free alphanumerical | **desirable** |
| **grade** | Tumour grade value | Grade 1 | free alphanumerical | **desirable** |
| **grading\_system** | Grade classification corresponding used to describe the stage, or Not Specified Add the version if available. | Nottingham grading system | free alphanumerical | **desirable** |
| **virology\_status** | Positive virology status at the time of collection. Any relevant virology information which can influence cancer like EBV, HIV,HPV status. | HIV | free text | **desirable** |
| **gene\_mutation\_status** | Outcome of mutational status tests for the following genes: BRAF, PIK3CA, PTEN, KRAS | KRAS:G12C | free text | **desirable** |
| **shareable** | Is patient treatment information available and shareable? If yes fill out following treatment columns: treatment\_naive\_at\_collection, treated\_at\_collection, treated\_prior\_to\_collection. | yes | yes/no/not provided | **essential** |
| **treatment\_naive\_at\_collection** | Was the patient treatment naive at the time of collection? This includes the patient being treated at the time of tumour sample collection and if the patient was treated prior to the tumour sample collection. The value will be "yes" if either treated\_at\_collection or treated\_prior\_to\_collection are "yes". | Yes | yes/no/not collected/not provided | **essential** |
| **treated\_at\_collection** | Was the patient being treated for cancer at the time of tumour sample collection. | radiotherapy | radiotherapy, chemotherapy, targeted therapy, homorno-therapy, immunotherapy | **desirable** |
| **treated\_prior\_to\_collection** | Was the patient treated prior to the tumour sample collection. | radiotherapy | radiotherapy, chemotherapy, targeted therapy, homorno-therapy, immunotherapy | **desirable** |
| **response\_to\_treatment** | Patient’s response to treatment. | Progressive disease | RECIST criteria: CR, PR, PD, SD | **desirable** |
| **model\_id** | Unique identifier for models derived from the same tissue sample | CRC0014LM | free alphanumerical and and -\_~. | **essential** |
| **model\_name** | Most common name associate with model. Please use the CCLE name if available | LAMA84\_HAEMATOPOIETIC\_AND\_LYMPHOID\_TISSUE | free alphanumerical | **essential** |
| **model\_name\_aliases** | Additional model names, if known | A 549 | free alphanumerical | **desirable** |
| **type** | Type of model | PDO | ~~organoid, cell line, other~~  PDO, PDXO, CDO, cell line, other | **essential** |
| **parent\_id** | Model Id of the original model used to generate this model. If the model was not in this set, please refer to it by external id | CRC022 | free alphanumerical and and -\_~. | **essential if applicable** |
| **origin\_patient\_sample\_id** | Unique ID of the patient tumour sample used to generate the model | CRC0228PRH0000000000D01000 | unique, free alphanumerical and -\_~. | **essential** |
| **growth\_properties** | Observed growth properties of the model | adherent | Adherent, Semi-adherent, suspension, Matrigel/BME domes | **essential** |
| **media\_id** | Unique identifier for each media formulation (Catalog number) | MF-002-001 | free alphanumerical | **essential if applicable** |
| **growth\_media** | Base media formulation the model was grown in | DMEM | free alphanumerical | **essential** |
| **plate\_coating** | Coating on plate model was grown in | None | Laminin, Matrigel, Collagen, None | **essential** |
| **other\_plate\_coating** | Other coating on plate model was grown in (not mentioned above) | None | Peg-based hydrogel, none | **essential** |
| **passage\_number** | Passage number at time of sequencing/screening | 20 | Freetext | **essential** |
| **contamination\_testing** | Has the model been tested for contamination? | yes | Y/N | **essential** |
| **contamination\_details** | What are the details of the contamination | mycoplasma | free alphanumerical | **desirable** |
| **supplements** | Additional supplements the model was grown with | 10% R-spondin, 10% Noggin, 10% FBS | free alphanumerical | **desirable** |
| **drug** | Additional drug/compounds the model was grown with | EGF | free alphanumerical | **desirable** |
| **drug\_concentration** | Concentration of Additional drug/compounds the model was grown with | 50ng/ml | free alphanumerical | **desirable** |
| **publications** | The original (or earliest available) publication outlining how the model was established | PMID: 27771609, PMID: 27771610 | PMID: 8 digit format | **desirable** |
| **supplier** | Model supplier name | ATTC | [ROR term](https://ror.org/03thhhv76) (if available), otherwise free text, not collected or not provided | **essential** |
| **supplier\_type** | Model supplier type - commercial, academic, other | Commercial | Commerical, Academic, Other | **essential** |
| **catalog\_number** | Catalogue number of cell model, if commercial | CCL-185 | free alphanumerical | **essential** |
| **vendor\_link** | Link to purchasable cell model, if commercial | <https://www.atcc.org/products/ccl-185> | URL | **essential** |
| **rrid** | Cellosaurus ID | CVCL\_0023 | free alphanumerical | **essential if applicable** |
| **related\_models** | Model\_ids of any related models linked to the same patient. | A088a | free alphanumerical and \_~. Separated by commas in case of multiple models | **essential if applicable** |
| **external\_ids** | Depmap accession, Cellusaurus accession or other id. Please place in comma separated list | CVCL\_0023 | free alphanumerical | **desirable** |
| **msi\_status** | MSI status of the model | MSS | MSI: Miscosatellite Instable, MSS: Miscrosatellite stable | **desirable** |
| **mutational\_burden** | Mutational rate per Mb | 4.76 | numerical | **desirable** |
| **ploidy** | Ploidy status of the model | 2.03 | numerical | **desirable** |
| **morphological\_features** | Morphological features of the model | elongated | free text | **essential** |
| **histological\_validation** | Immunofluorescent or immunohistochemical validation of the model | High concordance | Free text | **essential** |
| **SNP\_analysis** | Was SNP analysis done on the model? | yes | Yes, no | **essential** |
| **STR\_analysis** | Was STR analysis done on the model? | yes | Yes, no | **essential** |
| **tumour\_status** | Gene expression validation of established model |  | Free text | **essential** |
| **model\_purity** | Presence of tumour vs stroma or normal cells |  | Free text | **essential** |
| **comments** | Comments about the model that cannot be expressed by other fields |  | free text , not collected or not provided | **desirable** |
| **accessibility** | Any limitation of access of the model per type of users, such as academia only, industry and academia, or national limitation (e.g. no specific consent for sequencing) | academia only | academia, industry and academia or free text | **essential** |
| **email** | Contact email for any requests from users about models. If multiple, include as comma separated list. | j.doe@example.com | email format | **essential** |
| **name** | Contact person (should match that included in email column). | Jane Doe | free text | **essential** |
| **form\_url** | Link to a contact form, if available | <http://tumor.informatics.jax.org/mtbwi/pdxDetails.do?modelID=J000106527> | URL | **essential if applicable** |
| **database\_url** | Link to the database/biobank/resource, if available | <http://tumor.informatics.jax.org/mtbwi/pdxDetails.do?modelID=TM01220> | URL | **essential if applicable** |