Title
PDX-MI: Minimal Information for Patient-Derived Tumor Xenograft Models

Authors
Meehan, Terrence F
Conte, Nathalie
Goldstein, Theodore
et al.

PDX-MI: Minimal Information for Patient-Derived Tumor Xenograft Models

A full list of authors and affiliations appears at the end of the article.

Abstract

Patient-derived tumor xenograft (PDX) mouse models have emerged as an important oncology research platform to study tumor evolution, mechanisms of drug response and resistance, and tailoring chemotherapeutic approaches for individual patients. The lack of robust standards for reporting on PDX models has hampered the ability of researchers to find relevant PDX models and associated data. Here we present the PDX models Minimal Information standard (PDX-MI) for reporting on the generation, quality assurance and use of PDX models. PDX-MI defines the minimal information for describing the clinical attributes of a patient's tumor, the processes of implantation and passaging of tumors in a host mouse strain, quality assurance methods, and the use PDX models in cancer research. Adherence to PDX-MI standards will facilitate accurate search results for oncology models and their associated data across distributed repository databases and promote reproducibility in research studies using these models.

Introduction

Patient-Derived tumor Xenograft (PDX) models are created by implanting tumor cells or fragments from patients with cancer into a transplant-compliant mouse host (sFig 1)(1,2). Human tumors that engraft successfully in host mice are subsequently fragmented and passaged multiple times to generate large cohorts of tumor-bearing mice. PDX models accurately reflect the patient's tumor properties creating a powerful platform to study the molecular mechanisms of tumor growth and drug resistance as well as serving as patient "avatars" for predicting response to anti-cancer therapeutic compounds(3–5). The host strains for PDX model development are typically severely immunodeficient; however, "humanized" immune system mice engrafted with human immune cells are increasingly being used in xenograft studies to explore *in vivo* interactions between the immune system and cancer(6,7).

Corresponding Author: Terrence Meehan, Ph.D., European Molecular Biology Laboratory-European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, UK. Phone: +44(0)1223 492 591, Fax:+ 44 (0) 1223 484 468, tmeehan@ebi.ac.uk.
*= These authors contributed equally to the manuscript

Although many academic and commercial sources of PDX models have emerged in recent years, the size of the resources and the processes for creating and characterizing PDX models is quite variable. Crucial information about tumors, host strains, transplant and quality assurance processes are inconsistently presented in both the scientific literature and in database resources, limiting the ability of researchers to find relevant models and associated data. A standardized data exchange format is needed to foster the ability of researchers to identify appropriate PDX models and share information about them. As developers of NCI-funded informatics resources, we obtained the internal standards developed by four independent PDX model resources (the EurOPDX consortium(5), the IMODI consortium (France), the Patient Derived Models Repository at NCI-Frederick, and The Jackson Laboratory PDX Resource(8). After comparing standards in use across these resources, we generated a draft PDX-Minimal Information standard that was reviewed and modified by the authors of this report. We propose that the standards described here serve as the starting point for community wide adoption.

The PDX-Minimal Information (PDX-MI) Standard

The PDX-MI consists of four modules that reflect the process of generating, validating and using a PDX model: Clinical, Model Creation, Model Quality Assurance, Model Study and an additional associated metadata category (Table 1). Within each module, we define "essential" attributes that are required for accurate description and reporting on PDX models and "desirable" attributes that are frequently recorded by PDX producers and should be available.

The Clinical Module is divided into two sub-modules: "Clinical/Patient" and "Clinical/Tumor". "Clinical/Patient" requires information about the patient from which the engrafted tumor originates including age, sex, ethnicity, and disease diagnosis. To reduce the possibility of patient identification, PDX-MI recommends grouping ages into 5 year groups though more granular groupings maybe used in cases such as pediatric tumors if approved by a contributor's institutional review board. Reporting on patient consent is considered essential as well. Some attributes of patient treatment history are listed as "desirable" as they can impact the characteristics of resulting PDX models but may be challenging to provide due to patient privacy or data inaccessibility. The "Clinical/Tumor" sub-module reports on information about the originating tumor from which the PDX model is derived and includes tumor classification, anatomical location, and tumor histopathology. The presence or absence of specific diagnostic markers is listed as "essential" for tumor types where testing for such marker(s) is considered the clinical standard of care (e.g. FLT3 genotype in AML). Additionally, patient viral infection status has implications for disease biology as well as occupational safety and is included as a desirable field.

The Model Creation module of PDX-MI captures critical attributes in the creation of a PDX model. Host strain is reported using official strain nomenclature (http://www.informatics.jax.org/mgihome/nomen/index.shtml) as well as strain source and any modifications that "humanize" the host strain through engraftment of human immune-progenitor cells(6). Initial engraftment of the tumor describing processing of the tumor (solid or cell suspension) and the anatomical site of implantation (subcutaneous or

orthotopic) is represented. Other model generation characteristics such as engraftment rates and therapeutic response data are considered desirable. A "sub-line" field indicates when a PDX model is derived from an existing model that has changed characteristics (e.g. loss/gain of a biomarker, change in therapy response).

The Model Quality Assurance module captures information about tissue provenance and fidelity of the passaged tumor with respect to key characteristics of the patient tumor. Validation is required to confirm the PDX tumor is of the appropriate patient, and not of murine origin nor consisting primarily of Epstein-Barr virus (EBV) Human B lymphocytic cells as both are frequently observed in PDX model creation(9). Other "desirable" quality assurance methods vary with tumor types and can include histopathology, assessment of human cancer biomarkers by immunohistochemistry, in situ hybridization, and assessment of gene mutations and rearrangements, DNA methylation or gene expression profiling. Some producers evaluate how well a PDX model recapitulates the originating tumor's response by measuring PDX tumor growth response to standard of care treatment and this is included as a desirable attribute. Additional desirable information includes DNA profiling of serial passages to corroborate lineage fidelity and animal health status from standard health surveillance programs. The current PDX-MI requires evidence of quality assurance but does not require every possible technique be performed as methods vary across resources.

Model Study and other associated data. Tumors from PDX often undergo comprehensive genomic characterization and/or treatment in controlled dosing studies to define therapeutic response and resistance. PDX-MI includes "desirable" fields in the reporting of these studies that supplement existing guidelines for reporting on *in vivo* biomedical research(10). Additional optional metadata are accession IDs from data archives and citation IDs (including Digital Object Identifiers) for publications describing the PDX model(s).

Challenges of representing data from PDX models

Diversity of cancer subtypes

PDX models present unique challenges due to the specific approaches needed for the diversity of cancer subtypes. One challenge is that a subset of PDX models will require reporting on diagnostic biomarkers. For example in breast cancer, testing for certain pathological markers (estrogen receptor [ER], progesterone receptor [PR] and human epidermal growth factor 2 [HER2]) is considered the clinical standard of care for prognostic and predictive purposes and should therefore be considered essential for PDX-MI. Another challenge is that tumor grades and disease stages captured in the "Clinical Module", which drive patient diagnosis and treatment, maybe derived from scoring systems with diagnostic and geographic variation(11). PDX-MI will be flexible and allow users to report the system used clinically rather than enforce a particular one.

Terminology and Vocabularies

PDX resources employ a combination of custom and community developed vocabularies. This presents challenges in data integration as it takes expert knowledge to map the divergent systems. For example, cancer diagnoses are represented within different PDX

resources by terms from the NCI-Thesaurus(12), SNO-MED CT(13), MeSH(14), and the Disease Ontology(15). Free text descriptions are used for many PDX model attributes and a mix of generic, commercial, and chemical labels are used for drugs. Ontology resources and community model organism databases have been developing tools to semi-automate mapping of standards that produce unified indices to facilitate data query and discovery. Rather than impose a limited set of terms to describe a given minimal information attribute, PDX-MI will allow the reporting of a resource's internal standards. We will work with the PDX generators and local experts to ensure the quality of the generated mappings and provide feedback to the developers of ontology tools.

Implementation and Future Directions

The current version of PDX-MI describes the minimal information needed to report on a PDX models to facilitate data integration and resource sharing. The authors of this report hope PDX-MI will serve as a guide for authors and journal editors in promoting rigorous yet attainable publication standards and as a template for managers of public molecular archives in the capturing of critical metadata required for submission of PDX model data. PDX-MI standards will also be implemented in an online resource being jointly developed by EMBL-EBI and the Jackson Laboratory called PDX Finder www.pdxfinder.org (see Video 1). This resource currently in the prototype phase will provide a comprehensive global catalogue of PDX models available for researchers and their associated data across distributed repositories when formally launched at the end of 2017. PDX-MI will be used to validate data submissions from producers of PDX models and from data curated from the literature. PDX-MI will also inform scoring algorithms being developed by the Butte group to assess how well PDX models recapitulate hallmarks of human cancers.

Future versions of PDX-MI will capture additional details as procedures become more standardized. Input from clinical and translational professional societies will inform evolving requirements for diagnostic markers on a disease-specific basis. Given the recent success of immune checkpoint inhibitors in the treatment of cancer, improving "humanized immune system" PDX models is an area of intense research and PDX-MI will evolve to represent this. Other aspects of PDX models that are rapidly changing include improved surgical techniques and quality assurance methods. As we develop resources to capture and disseminate data related to PDX models, we will continue to improve and version PDX-MI to reflect the state of the art in the field. A web-based form to allow feedback from the community about the standard described here can be accessed at the Mouse Tumor Biology database web site (http://tumor.informatics.jax.org). As has been demonstrated across multiple disciplines, a minimal standard adopted by a research community accelerates the rate of scientific discovery while reducing unnecessary duplication.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Terrence F. Meehan[1,*], Nathalie Conte[1,*], Theodore C. Goldstein[2], Giorgio Inghirami[3], Mark A. Murakami[4], Sebastian Brabetz[5,6], Zhiping Gu[7], Jeffrey A. Wiser[7], Patrick Dunn[7], Dale A. Begley[8], Debra M. Krupke[8], Andrea Bertotti[9], Alejandra Bruna[10], Matthew Brush[11], Annette T. Byrne[12], Carlos Caldas[10], Amanda L. Christie[4], Dominic Clark[1], Heidi Dowst[13], Jonathan R. Dry[14], James H. Doroshow[15], Olivier Duchamp[16], Yvonne A. Evrard[17], Stephane Ferretti[18], Kristopher K. Frese[19], Neal C. Goodwin[20], Danielle M. Greenawalt[21], Melissa A. Haendel[11], Els Hermans[22], Peter J. Houghton[23], Jos Jonkers[24], Kristel Kemper[24], Tin O. Khor[25], Michael T. Lewis[26], KC Kent Lloyd[27], Jeremy C. Mason[1], Enzo Medico[9], Steven B. Neuhauser[8], Jim M. Olson[28], Daniel S. Peeper[24], Oscar M. Rueda[10], Je Kyung Seong[29], Livio Trusolino[9], Emilie Vinolo[30], Robert J. Wechsler-Reya[31], David M. Weinstock[4], Alana Welm[32], S. John Weroha[33], Frédéric Amant[24,34], Stefan M Pfister[5,6,35], Marcel Kool[5], Helen Parkinson[1], Atul J. Butte[2], and Carol J. Bult[8]

Affiliations

[1]European Molecular Biology Laboratory-European Bioinformatics Institute, Hinxton, UK [2]Institute for Computational Health Sciences, University of California, San Francisco, USA [3]Department of Pathology and Laboratory Medicine, Weill Cornell Medical College, New York, USA [4]Dana-Farber Cancer Institute and Harvard Medical School, Boston, USA [5]Division of Pediatric Neuro-oncology, German Cancer Research Center (DKFZ), Heidelberg, Germany [6]German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany [7]Northrop Grumman Information Systems Health IT, Rockville, USA [8]The Jackson Laboratory, Bar Harbor, USA [9]Candiolo Cancer Institute IRCCS and Department of Oncology, University of Torino, Torino, Italy [10]Cancer Research UK Cambridge Institute, Cambridge Cancer Centre, University of Cambridge, United Kingdom [11]Department of Medical Informatics and Clinical Epidemiology and OHSU Library, Oregon Health & Science University, Portland, USA [12]Royal College of Surgeons in Ireland, Ireland [13]Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, USA [14]Oncology Innovative Medicines and Early Development, AstraZeneca R&D Boston, Waltham, MA, USA [15]Center for Cancer Research, National Cancer Institute, USA [16]Oncodesign Biotechnology, and IMODI consortium, France [17]Leidos Biomedical Research, Inc, Frederick National Laboratory for Cancer Research, Frederick, USA [18]Oncology Disease Area, Novartis Institutes for Biomedical Research, Switzerland [19]Cancer Research UK Manchester Institute, The University of Manchester, United Kingdom [20]Champions Oncology, Baltimore, USA [21]Translational Bioinformatics Bristol-Myers Squibb, Pennington, New Jersey USA [22]Katholieke Universiteit Leuven, Leuven, Belgium [23]Greehey Children's Cancer Research Institute, University of Texas Health Science Center at San Antonio, San Antonio, USA [24]The Netherlands Cancer Institute, Amsterdam, Netherlands [25]Institute for Applied Cancer Science, Center for Co-Clinical Trial, The University of Texas MD Anderson Cancer Center, Houston, USA,

USA [26]The Lester and Sue Smith Breast Center, Departments of Molecular and Cellular Biology and Radiology, Baylor College of Medicine, USA [27]Department of Surgery, School of Medicine, and Mouse Biology Program, University of California Davis, Davis, USA [28]Fred Hutchinson Cancer Research Center, Seattle Children's Hospital, USA [29]Research Institute for Veterinary Science and Korea Mouse Phenotyping Center, Seoul, Republic of Korea [30]Seeding Science SAS, Paris, France [31]Tumor Initiation and Maintenance Program, NCI-Designated Cancer Center, La Jolla, USA [32]Huntsman Cancer Institute, University of Utah, Salt Lake City, USA [33]Department of Oncology, Mayo Clinic College of Medicine, Rochester, USA [34]University of Leuven, Leuven, Belgium [35]Department of Pediatric Oncology, Hematology and Immunology, Heidelberg University Hospital, Heidelberg, Germany

Acknowledgments

References

1. Uthamanthil, R., Tinkey, P., De Stanchina, E. Patient Derived Tumor Xenograft Models Promise, Potential and Practice. Academic Press; 2017.

2. Hidalgo M, Amant F, Biankin AV, Budinska E, Byrne AT, Caldas C, et al. Patient-Derived Xenograft Models: An Emerging Platform for Translational Cancer Research. Cancer Discov [Internet]. 2014; 4:998–1013. [cited 2017 Jan 31] Available from: http://www.ncbi.nlm.nih.gov/pubmed/25185190.

3. Tentler JJ, Tan AC, Weekes CD, Jimeno A, Leong S, Pitts TM, et al. Patient-derived tumour xenografts as models for oncology drug development. Nat Rev Clin Oncol [Internet]. 2012; 9:338–50. Available from: http://dx.doi.org/10.1038/nrclinonc.2012.61.

4. Kopetz S, Lemos R, Powis G. The promise of patient-derived xenografts: The best laid plans of mice and men. Clin. Cancer Res. 2012:5160–2. [PubMed: 22912394]

5. Byrne AT, Alférez DG, Amant F, Annibali D, Arribas J, Biankin AV, et al. Interrogating open issues in cancer precision medicine with patient-derived xenografts. Nat Rev Cancer [Internet]. 2017; 17:254–68. [cited 2017 Apr 18] Available from: http://www.ncbi.nlm.nih.gov/pubmed/28104906.

6. Shultz LD, Brehm MA, Garcia-Martinez JV, Greiner DL. Humanized mice for immune system investigation: progress, promise and challenges. Nat Rev Immunol [Internet]. 2012; 12:786–98. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3749872&tool=pmcentrez&rendertype=abstract.

7. Shultz, LD., Goodwin, N., Ishikawa, F., Hosur, V., Lyons, BLGD. [cited 2017 Feb 13] Human cancer growth and therapy in immunodeficient mouse models; Cold Spring Harbor protocols. Cold Spring Harb Protoc 2014 [Internet]. 2014. p. 694-708.Available from: http://www.cshlpress.com/default.tpl?action=full&--eqskudatarq=976

8. Bult CJ, Krupke DM, Begley DA, Richardson JE, Neuhauser SB, Sundberg JP, et al. Mouse Tumor Biology (MTB): A database of mouse models for human cancer. Nucleic Acids Res. 2015; 43:D818–24. [PubMed: 25332399]

9. Bondarenko G, Ugolkov A, Rohan S, Kulesza P, Dubrovskyi O, Gursel D, et al. Patient-Derived Tumor Xenografts Are Susceptible to Formation of Human Lymphocytic Tumors. Neoplasia [Internet]. 2015; 17:735–41. [cited 2016 Oct 19] Available from: http://www.ncbi.nlm.nih.gov/pubmed/26476081.

10. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. PLoS Biol [Internet]. 2010; 8:e1000412. [cited 2016 Aug 9] Available from: http://www.ncbi.nlm.nih.gov/pubmed/20613859.

11. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol [Internet]. 2010; 17:1471–4. [cited 2016 Oct 19] Available from: http://www.ncbi.nlm.nih.gov/pubmed/20180029.

12. Sioutos N, Coronado S de, Haber MW, Hartel FW, Shaiu WL, Wright LW. NCI Thesaurus: A semantic model integrating cancer-related clinical and molecular information. J Biomed Inform. 2007; 40:30–43. [PubMed: 16697710]

13. NIH-NLM. SNOMED Clinical Terms® (SNOMED CT®) [Internet]. NIH-US Natl. Libr. Med. 2015. Available from: http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html

14. National Library of Medicine. MeSH (Medical Subject Headings): is the NLM controlled vocabulary thesaurus used for indexing articles for PubMed [Internet]. Bethesda Natl. Cent. Biotechnol. Inf. 2016. Available from: http://www.ncbi.nlm.nih.gov/mesh

15. Kibbe WA, Arze C, Felix V, Mitraka E, Bolton E, Fu G, et al. Disease Ontology 2015 update: An expanded and updated database of Human diseases for linking biomedical knowledge through disease data. Nucleic Acids Res. 2015; 43:D1071–8. [PubMed: 25348409]

Table 1

The PDX-MI consists of four modules that reflect the process of generating and validating a PDX model: Clinical, Model Creation, Model Quality assurance, and Model Study/associated metadata. The "field" column describes each module attribute, the "recommendation" column defines whether the attribute is essential or desirable. All essential attributes must be submitted to provide an accurate description and reporting of PDX models, desirable attributes should be submitted if available. Finally the" example entry" gives an example(s) of each attribute.

Module	Field	Recommendation	Example entry or choice
Clinical/Patient	Submitter Patient ID	Essential	PAT-123
	Gender	Essential	female
	Age	Essential	30–35 (binned in 5 year age groups)
	Diagnosis	Essential	invasive breast cancer
	Consent to share data	Essential	yes/no/available to academic centers only
	Ethnicity/Race	Desirable	caucasian
	Current Treatment drug	Desirable	everolimus; CHEMBL83
	Current Treatment protocol (dose; details)	Desirable	afinitor;10 mg/day
	Prior treatment protocol	Desirable	surgery and nolvadex; 40 mg/day
	Response to prior treatment	Desirable	progressive disease (RECIST1.1)
	Virology status	Desirable	HIV–/HBV–/HCV+/HTLV–/EBV+
Clinical/Tumor	Submitter Tumor ID	Essential	TUM-123
	Primary tumor tissue of origin	Essential	breast
	Primary, metastasis, recurrence	Essential	metastasis
	Specimen tumor tissue	Essential	liver
	Tissue histology	Essential	invasive ductal carcinoma
	Tumor Grade; classification	Essential	grade 3; Elston
	Disease Stage; classification	Essential	T3N2M1; TNM or Non applicable (example blood cancer)
	Specific markers (diagnostic linked); platform	Essential	ER+, PR+, HER2+; IHC
	Is tumor from untreated patient?	Essential	yes/no
	Original tumor sample type	Desirable	biopsy, surgical sample, ascites fluid, blood, etc
	Tumor from an existing PDX model? ID? Why sub-line?	Desirable	Yes, PDX#123, lost cisplatin resistance
Model Creation	Submitter PDX ID	Essential	PDX#123
	Mouse strain (and source)	Essential	NOD.Cg-Prkdc<scid>Il2rg<tm1Wj>l/SzJ, The Jackson Laboratory
	Strain immune system humanised?	Essential	yes/no
	Type of humanisation	Essential	CD34+ hematopoietic stem cell-engrafted/ PBMC/ Thymus/ Thymus-fetal liver/ iPSC/ other
	Tumor preparation	Essential	tumor solid, cell suspension, asite
	Injection type and site	Essential	subcutaneous; right flank

Module	Field	Recommendation	Example entry or choice
	Mouse treatment for engraftment	Desirable	estrogen treatement
	Engraftment rate	Desirable	80%
	Engraftment time	Desirable	8 weeks
Model Quality Assurance	Tumor Characterization Technology	Essential	histology and IHC
	Tumor confirmed not to be of mouse/EBV origin	Essential	yes/no; negative for murine CD45
	Response to Standard of Care (Pharmacological positive control)	Desirable	Not assessed / Assessed- complete response, partial response, stable disease, progressive disease
	Animal health status	Desirable	SPF/SOPF, C.Bovis & Pneumocystis negative/positive
	Passage QA performed	Essential	passage P4
Model study	Treatment, passage	Desirable	pertuzumab in combination with trastuzumab; CHEMBL2007641 and CHEMBL1743082; passage P4
	Treatment protocol (dose; details)	Desirable	trastuzumab (30 mg/kg loading dose, 15 mg/kg weekly); pertuzumab (30 mg/kg loading dose, 15 mg/kg weekly)
	Treatment Response	Desirable	complete response, partial response, stable disease, progressive disease
	Tumor OMICS: sample id; sample site; purity (mouse vs human); technology; passage	Desirable	TUMpdx-123; subcutaneous; 90% human; exome sequencing; passage P5
	Development of metastases in strain (Y/N, site); passage	Desirable	Yes; liver; passage P6
	Lag time/doubling time of tumor	Desirable	48h
Associated metadata	PDX model availability ?	Desirable	yes/no; frozen tumour; live mouse
	governance retriction for distribution	Desirable	available to academic centers only
	ID for associated publication, image, archived data (URL, PMID, DOI)	Desirable	http://www.dataset.org/EXA123; PMID: 28025748; DOI: 10.1186/ s13058-015-0523-1