

EXPLICACIÓN DEL ESQUEMA DE DATOS PARA BIOMODELOS

- **PACIENTE:** cada paciente se localiza en la base de datos mediante su **NHC** y presenta como atributos su **fecha de nacimiento**, su **sexo** y además se puede hacer el cálculo de su edad actualizado en cada momento a partir de su fecha de nacimiento y la fecha actual. Restricciones: el valor almacenado en “sexo” tiene que ser uno de los dos del estándar que se utilice. La edad deberá estar comprendida entre 0 y 150 (años).
Las variables edad del paciente a fecha de la operación, edad del ratón se calculan a partir de las fechas de nacimiento y la actual.

→ Cada paciente presenta N tumores (se presupone que si se registra un paciente es porque lleva asociado al menos un tumor. El registro de la biopsia líquida, si se diera el caso de que no hay tumor, podría quedar registrado en Bio-e-bank/Noray). Atributos de la relación “un paciente presenta un tumor” → se guarda la **fecha de operación**, la **fecha de registro** en el biobanco (son características que no son intrínsecas del tumor, sino que aparecen cuando un paciente presenta un tumor que es extirpado).

- **TUMOR:** cada tumor se localiza mediante el **código asignado por el Biobanco** y recoge datos como la **clasificación** del tumor (CIE-10), el **órgano**, el **grado** (I-IV), el **TNM**, el **estado** (px0/activo/fallido/congelado) y el **código del laboratorio**. Aclaración: si el px0 (proceso en el que se hace crecer el tumor primario para poder empezar con el px1) está creciendo, se marca en estado PX0. Si ha pasado con éxito el pase 0, se marca como **ACTIVO**. Si no ha funcionado, el tumor se considera **FALLIDO**. Restricciones: valores limitados de grado y de estado. Comprobar que la fecha de operación es anterior a la fecha de registro. *Se recoge el id del paciente asociado a ese tumor.*

→ Cada tumor puede tener asociado uno o ningún registro de datos moleculares

→ Cada tumor puede tener asociado uno o ningún registro de secuenciación genómica

- **SECUENCIACIÓN GENÓMICA:** tiene un identificador y datos asociados.
Pendiente: modelado de la secuenciación genómica -> solicitar marcadores de interés
- **DATOS MOLECULARES:** tiene un identificador y datos asociados.
Pendiente: modelado de los datos moleculares -> solicitar marcadores de interés

→ Cada tumor puede presentar 0 ó 1 registro de biopsia líquida (solo interesa almacenar si hay tubos de tipo suero, buffy y plasma o si no los hay)

- **BIOPSIA LÍQUIDA:** cada biopsia será identificada mediante el **código asignado por la unidad de extracciones** y recogerá otros datos como la **fecha** en la que se hizo la biopsia y si se han obtenido o no **buffy**, **plasma** y **suero**. Restricciones: el id tiene que ser un valor numérico y los campos que guardan si tiene o no cada tipo de muestra deberán almacenar uno de los dos valores posibles. *Se recoge el id del tumor asociado a esa biopsia.*

→ Cada tumor (si estado = activo) genera hasta 3 biomodelos (PDX, PDO y LC)

- **BIOMODELO:** cada biomodelo se identifica con un **ID** secuencial -> autogenerado (formato BMPDX-003, BMPDO-021, BMLC-001). Además, contiene campos como **tipo** de biomodelo (PDX, PDO, LC), **fecha de creación**, un **cuadro de texto** descriptivo, un atributo que recoge si se ha **intentado** el biomodelo o aún no, otro que indica si **progresa** o no, otro que indica si se ha utilizado para **ensayos preclínicos** o no y un último que recoge si el **estado** del biomodelo es activo o fallido. El campo que recoge la **viabilidad** podría adoptar valores entre 0 y 1, en función de los pases y ensayos que hayan sido o no exitosos en ese biomodelo específico. Restricciones: la variable “tipo” es única para cada tumor primario, por lo que un tumor solo puede tener asociado un biomodelo de cada tipo. *Se recoge el id del tumor asociado a ese biomodelo y, si es el caso, el id del ensayo del que procede.* **Validar el formato de la codificación → BMPDX-003 (Biomodelo de tipo PDX número 3, es una etiqueta que he creado para identificar los biomodelos)**

→ Cada biomodelo genera de 0 a N pases

- **PASE:** un pase consta de un **ID** (formato BMPDX-003-PX-10), un atributo **número** (número del pase) y otro que refleja el **estado (activo/fallido)**. Además, si el tipo de biomodelo es PDX, el campo que indica el **índice de similitud** pretende hacer un promedio de los índices de similitud de cada ensayo PDX de ese mismo pase con respecto al tumor original. Por otro lado, la **viabilidad** vuelve a ser una medida de éxito, en este caso, de cada pase (se podría calcular recorriendo todos los ensayos de ese pase y calculando cuántos han sido marcados como exitosos sobre el total de ensayos). Restricciones: id, el valor de número es único (solo hay un pase 1, pase 2...). *Se recoge el id del biomodelo asociado a ese pase y también el id del ensayo del que procede el pase.*

→ Un pase genera N ensayos

- **ENSAYO:** cada ensayo tiene un identificador (BMPDX-003-PX-10-EN-07) y almacena: si se ha **enviado** una parte al biobanco o no, si ha tenido **éxito** o no, su **fecha de creación**, y un **cuadro descriptivo**. Si se ha enviado al Biobanco, además aparecerá un campo que guardará la **fecha** de registro. *Se recoge el id del pase asociado a ese ensayo.*

→ Cada ensayo genera de 0 a N registros de uso

→ Cada ensayo genera de 0 a N imágenes

→ Cada ensayo experimenta de 0 a N procesos de crioconservación

→ Cada ensayo puede tener asociado uno o ningún registro de datos moleculares

→ Cada ensayo puede tener asociado uno o ningún registro de secuenciación genómica

- **REGISTRO DE USO:** un registro de uso tiene como propósito documentar y trazar las transacciones de cada ensayo. De esta manera, cada registro almacenará un **identificador**

(BMPDX-003-PX-10-EN-07-RG-05), la **fecha** de la transacción, el **tipo** de uso a realizar y un **campo descriptivo** en el que se especifiquen los detalles de la operación.

Delimitar cuáles son todos los posibles registros de uso: ensayos preclínicos, comercialización, descongelación, cesión, transferencia (Biobanco, IIB-UAM, laboratorio de Julie)

- **IMAGEN:** si se registra una imagen, se asocia un **id** (BMPDX-003-PX-10-EN-07-IM-01), una **fecha** de adquisición y el **tipo** de imagen. También se recogerá si la muestra ha sido **revisada por anatomía patológica** o no (los detalles podrán estar incluidos en la propia descripción del ensayo). **Delimitar posibles tipos de imagen:** IHQ, FACS, cultivo, H&E
- **CRIOPRESERVACIÓN:** cada vez que se congele una muestra, se guarda la **fecha**, el **lugar** y el número de **viales**. **Delimitar posibles lugares de congelación de muestras** ([p.ej:](#) Biobanco/IIB).
- **SECUENCIACIÓN GENÓMICA:** tiene un identificador y datos asociados.
Pendiente: modelado de la secuenciación genómica -> solicitar marcadores de interés
- **DATOS MOLECULARES:** tiene un identificador y datos asociados.
Pendiente: modelado de los datos moleculares -> solicitar marcadores de interés

→ Un ensayo puede dar lugar a 0 ó 1 pases (puede que no se utilice para dar comienzo a un pase o puede que sí).

→ Un ensayo puede generar de 0 a 2 nuevos biomodelos (si el ensayo es de tipo PDX, podrá servir, como máximo, para empezar un PDO y una línea celular)

→ Un ensayo puede ser de tipo LC (si el biomodelo es de tipo LC)

- **ENSAYO LC:** además de heredar las propiedades de ensayo, se guarda el **tipo** de placa utilizada, la **confluencia** (porcentaje de la superficie de cultivo que se encuentra cubierta por células), si aparecen **esferoides** o no, si hay **línea celular** o no y la fecha de digestión.

→ Cada ensayo LC puede tener uno o ningún registros con FACS

- **FACS:** identificador y datos asociados.
Pendiente: marcadores de interés (lista cerrada)

→ Un ensayo puede ser de tipo PDO (si el biomodelo es de tipo PDO)

- **ENSAYO PDO:** además de heredar las propiedades de ensayo, se guarda el **tipo** de placa utilizada, el **número de gotas**, el **número de organoides**, el **día de visualización**, una **valoración** numérica del 1-5 y el **número de organoides congelados**.

→ Un ensayo puede ser de tipo PDX (si el biomodelo es de tipo PDX)

- **ENSAYO PDX:** además de heredar las propiedades de ensayo, se recogen los **aumentos del escáner** utilizados en la muestra (10x/20x/40x), si se ha hecho **bloque de parafina** o no (FFPE), si hay **lámina de hematoxilina eosina** o no, los **datos de inmunohistoquímica**, la **latencia en semanas** y el **índice de similitud** con respecto al

tumor original (para considerarlo, se mezcla el tejido de todos los implantes. Lo mismo con los datos IHQ). [Se recoge el id del ratón asociado a ese ensayo.](#)

→ Cada ensayo tiene asociado un ratón determinado.

- **RATÓN:** además del **id** propio, se tiene en cuenta el estándar **PROEX**, el **animalario** en el que se encuentra (IIB/HURYC), la **cepa**, el **sexo**, la **fecha de nacimiento**, la **fecha de fallecimiento**, la **edad en semanas** y la **causa de fallecimiento**. Restricción: para calcular la edad en semanas, se tiene en cuenta la edad con respecto al momento actual, a no ser que el campo que describe la fecha de fallecimiento no esté vacío porque ya haya fallecido el animal.

→ Cada ensayo genera N implantes en el ratón

- **IMPLANTE:** cada implante tiene un **identificador**, un **lugar** (flanco izquierdo, derecho...**Delimitar todas las posibilidades**) y un **tipo** (subcutáneo/ortotópico). [Se recoge el id del ensayo PDX asociado a ese implante.](#)

→ Cada implante presenta varios registros de tamaño.

- **REGISTRO DE TAMAÑO:** se recoge un **id**, el **número de semana** en la que se hace el registro, el **tamaño inicial** del tumor en mm³, el **tamaño final** y el **tamaño límite** que puede alcanzar en el ratón. [Se recoge el id del implante asociado a ese registro de tamaño.](#)

Se propone el uso de ontologías para utilizar terminología estandarizada en los campos en los que sea posible (sexo, clasificación...)