

Patient Matching: Integración y estandarización de nomenclatura genómica y su relevancia para la identificación de pacientes candidatos a ensayos clínicos oncológicos en la era de la medicina de precisión

Ivana Tonic¹, Luis García Palacios¹, Eva Díez², Natalia Leronés², Nicolás Palomo², Raquel Galán³

¹PHC Medical Expert, Roche Farma S.A. (Spain), ²Senior Clinical Operations Lead, Roche Farma S.A. (Spain), ³Tumor Agnostic and Rare Cancers Medical Expert, Roche Farma S.A. (Spain).

Introducción y objetivos

- La generación de evidencia en la era de la medicina de precisión tiene el reto de seleccionar pacientes en base a la alteración de variantes genómicas en cáncer. A día de hoy, no todos los tumores de pacientes diagnosticados de cáncer son secuenciados, o se secuencian con una amplia variedad de paneles y los resultados se comparten en formatos digitales diversos.
- Las bases de datos que recopilan la información de ensayos clínicos, incluida la más usada, *ClinicalTrials.gov*, carecen de la precisión y el estado de actualización suficientes para identificar pacientes que cumplan los criterios de elegibilidad
- En un centro sanitario suelen convivir distintas técnicas de secuenciación, y las anotaciones genómicas pueden diferir de una a otra, siendo compleja la unificación en la nomenclatura, especialmente de las alteraciones estructurales, lo que puede llegar a dificultar la identificación del paciente correcto.
- La medicina de precisión se encuentra con el problema del acceso al tratamiento de los pacientes, y los ensayos clínicos son, en ocasiones, la única vía de acceso a tratamientos innovadores cuando se han agotado todas las opciones terapéuticas disponibles.
- El objetivo principal del presente trabajo es indagar sobre la potencial utilidad de la aplicación de un algoritmo que detecta y alerta de alteraciones genómicas a partir de informes terciarios de NGS en formatos electrónicos estructurados en base a las alteraciones genómicas recogidas en los criterios de inclusión/exclusión de ensayos clínicos y determinar las barreras técnicas o logísticas para su aplicación en hospitales españoles.
- Aprovechar los resultados de secuenciación genómica obtenidos para un paciente tanto para tomar una decisión clínica inmediata como para valorar la inclusión del paciente en otro ensayo clínico en el mismo centro, o incluso derivar a otro centro.
- Hipótesis del presente proyecto:** Basar la búsqueda de un ensayo adecuado para un paciente en criterios exactos de biomarcadores, edad e histología tumoral, es suficiente para hacerla más precisa y reducir el tiempo invertido por el profesional sanitario.

Material y métodos

•Se partirá del caso de uso en ensayos clínicos adaptativos (Plataforma y Paraguas) de Roche:

En base a los criterios de inclusión y exclusión de la edad, tipo tumoral y alteraciones genómicas se codifica un sencillo algoritmo para aplicarlo a los informes terciarios de plataformas de secuenciación masiva, capaz de alertar de un potencial candidato a alguno de los dos ensayos, objetos del estudio, y se compara con la búsqueda manual de ensayos clínicos en base a los mismos criterios de inclusión y exclusión publicados en *ClinicalTrial.gov*.

Se analiza mediante entrevistas personales la metodología de hospitales públicos universitarios de Madrid para hacer accesible a todo el departamento de Oncología, Anatomía Patológica y Monitores de Ensayos clínicos, los resultados genómicos de pacientes con cáncer procedentes de resultados de secuenciación genómica del propio hospital, de las plataformas comerciales y alteraciones genómicas analizadas mediante análisis de gen único en el laboratorio de Anatomía Patológica.

Se analiza la homogeneidad en la nomenclatura de alteraciones genómicas de distintas plataformas.

Resultados

- Durante aproximadamente un año, se analizan 860 test de secuenciación masiva tanto a partir de tejido como de Sangre con un algoritmo basado en Biomarker Memo de forma automática para un ensayo plataforma y un ensayo umbrellita con el resultado de 44 informes que cumple criterios de inclusión genómicos, tipo tumoral y edad cuyas cohortes estaban abiertas en España en esos instantes. 17 informes (39%), fueron positivos para el ensayo Plataforma y 26 (59%) para el ensayo Umbrella, mientras que según los criterios de *Clinicaltrials.gov* tan solo identifican 11 para el plataforma y 8 para el umbrellita.
- Es interesante observar que basado tan solo en alteraciones genómicas de *ClinicalTrial.gov*, se obtienen un gran número de coincidencias que en realidad no son pacientes elegibles para alguno de los dos ensayos por los criterios de inclusión/exclusión de biomarcadores o por el estado de reclutamiento de la cohorte concreta.
- Las bases de ensayos clínicos públicas consultadas incluida el Registro Español de Estudios Clínicos (REEC), carecen de la estandarización de la información de biomarcadores y el estado de reclutamiento (Abierto/Cerrado) rigurosamente actualizado según cada cohorte, lo que resulta en baja precisión en la identificación de pacientes.
- En la Imagen 1 se muestra un ejemplo de la información publicada en *Clinicaltrial.gov* vs el detalle extraído a partir de un Biomarker Memo y programado en un algoritmo para la identificación de pacientes en un *Clinical Report* de NGS
- El análisis de flujo de integración de datos genómicos en varios centros hospitalarios universitarios demostró un proceso heterogéneo y manual de los resultados genómicos en cuanto a su incorporación a la historia del paciente y la accesibilidad de los mismos a distintos especialistas. En ocasiones las alteraciones que se reportan son tan solo las Tier I y por consiguiente las Tier II y III se pierden y de esta forma, la oportunidad de inclusión en ensayos adaptativos y el acceso a tratamientos innovadores en investigación.
- Existe una falta de estandarización universalmente aceptada, tanto por los promotores de estudios como por los proveedores de secuenciación, en la nomenclatura genómica. Así, una misma alteración puede ser reportada de formas diferentes según el pipeline bioinformático con el que es analizado. Las mayores diferencias se encuentran en alteraciones estructurales como Indels de más de 10 aminoácidos y reordenamientos que generan proteínas de fusión (Imagen 2).
- Los paneles pequeños (HotSpot y single-gene) carecen de la amplitud suficiente para cubrir las firmas genómicas (MSI, TMB y HRD) y ciertas alteraciones que son criterios de inclusión/exclusión para estos dos ensayos. Este punto resulta crítico si tenemos en cuenta que cada año se incorporan al arsenal terapéutico en investigación hasta 15 nuevas dianas cada año hasta el 2028.

Clinicaltrials.gov rule PIK3CA	Biomarker rule PIK3CA																																																																										
Experimental: Cohort H: PIK3CA multiple mutant-positive tumors Participants with metastatic or advanced solid tumors will receive GDC-0077 QD at a starting dose of 9 mg by mouth (PO) in repeated 28-day cycles.	<p>Table 8. Biomarker definition for Cohort H: PIK3CA alterations</p> <table border="1"><thead><tr><th>Gene (Transcript ID)</th><th>Variant Type</th><th>Biomarker Rules</th></tr></thead><tbody><tr><td rowspan="2">PIK3CA (NM_006218)</td><td rowspan="2">Short Variant</td><td>Samples with at least two alterations:<ul style="list-style-type: none">One alteration must be one of the following:<ul style="list-style-type: none">E542A/K/Q/YE545A/D/K/GH1047R/L/YThe second alteration must be listed in Table A2</td></tr><tr><td>Copy Number</td><td>Not applicable</td></tr><tr><td></td><td>Rearrangement</td><td>Not applicable</td></tr></tbody></table> <p>Table A2. Eligible PIK3CA alterations</p> <table border="1"><tbody><tr><td>E81K</td><td>G106S</td><td>N345H</td><td>C420R</td><td>E542K</td><td>E545Q</td><td>E726K</td><td>H1047P</td><td>G1049R</td></tr><tr><td>R88Q</td><td>G106V</td><td>N345I</td><td>E453A</td><td>E542L</td><td>E545R</td><td>M1004I</td><td>H1047Q</td><td>G1049S</td></tr><tr><td>R88P</td><td>R108H</td><td>N345K</td><td>E453D</td><td>E542Q</td><td>E545V</td><td>M1043I</td><td>H1047R</td><td></td></tr><tr><td>R93Q</td><td>R108L</td><td>N345S</td><td>E453G</td><td>E542R</td><td>Q546E</td><td>M1043L</td><td>H1047T</td><td></td></tr><tr><td>R93W</td><td>R108P</td><td>N345T</td><td>E453K</td><td>E542V</td><td>Q546H</td><td>M1043T</td><td>H1047Y</td><td></td></tr><tr><td>P104L</td><td>K111N</td><td>N345Y</td><td>E453Q</td><td>E545A</td><td>Q546K</td><td>M1043V</td><td>H1048L</td><td></td></tr><tr><td>G106A</td><td>K111R</td><td>R357Q</td><td>E453V</td><td>E545D</td><td>Q546L</td><td>H1047D</td><td>H1048R</td><td></td></tr></tbody></table>	Gene (Transcript ID)	Variant Type	Biomarker Rules	PIK3CA (NM_006218)	Short Variant	Samples with at least two alterations: <ul style="list-style-type: none">One alteration must be one of the following:<ul style="list-style-type: none">E542A/K/Q/YE545A/D/K/GH1047R/L/YThe second alteration must be listed in Table A2	Copy Number	Not applicable		Rearrangement	Not applicable	E81K	G106S	N345H	C420R	E542K	E545Q	E726K	H1047P	G1049R	R88Q	G106V	N345I	E453A	E542L	E545R	M1004I	H1047Q	G1049S	R88P	R108H	N345K	E453D	E542Q	E545V	M1043I	H1047R		R93Q	R108L	N345S	E453G	E542R	Q546E	M1043L	H1047T		R93W	R108P	N345T	E453K	E542V	Q546H	M1043T	H1047Y		P104L	K111N	N345Y	E453Q	E545A	Q546K	M1043V	H1048L		G106A	K111R	R357Q	E453V	E545D	Q546L	H1047D	H1048R	
Gene (Transcript ID)	Variant Type	Biomarker Rules																																																																									
PIK3CA (NM_006218)	Short Variant	Samples with at least two alterations: <ul style="list-style-type: none">One alteration must be one of the following:<ul style="list-style-type: none">E542A/K/Q/YE545A/D/K/GH1047R/L/YThe second alteration must be listed in Table A2																																																																									
		Copy Number	Not applicable																																																																								
	Rearrangement	Not applicable																																																																									
E81K	G106S	N345H	C420R	E542K	E545Q	E726K	H1047P	G1049R																																																																			
R88Q	G106V	N345I	E453A	E542L	E545R	M1004I	H1047Q	G1049S																																																																			
R88P	R108H	N345K	E453D	E542Q	E545V	M1043I	H1047R																																																																				
R93Q	R108L	N345S	E453G	E542R	Q546E	M1043L	H1047T																																																																				
R93W	R108P	N345T	E453K	E542V	Q546H	M1043T	H1047Y																																																																				
P104L	K111N	N345Y	E453Q	E545A	Q546K	M1043V	H1048L																																																																				
G106A	K111R	R357Q	E453V	E545D	Q546L	H1047D	H1048R																																																																				

Imagen 1. Diferencias entre la información proporcionada en el protocolo del ensayo vs. información en *ClinicalTrials.gov*

Una deleción canónica de EGFR en el exón 19 puede ser nombrada de múltiples formas, la mayoría aceptadas en HGVS.

EGFR Del 19
EGFR exon 19 deletion (L747_S752>Q)
EGFR L747_S752delinsQ
EGFR DEL L747-S752INSQ
EGFR LEU747_SER752DELINSGLN
NC_000007.14(EGFR):g.55174776_55174793delinsCAA
NC_000007.13(EGFR):g.55242469_55242486delinsCAA
ENST00000275493.7(EGFR):c.2239_2256delinsCAA
ENSP00000275493.2(EGFR):p.Leu747_Ser752delinsGln
NM_005228.5(EGFR):c.2239_2256delinsCAA
NP_005219.2(EGFR):p.Leu747_Ser752delinsGln

Imagen 2. Diversidad de nomenclaturas que representan una deleción en el Exón 19 del gen EGFR

Conclusión

El presente proyecto carece de la rigurosidad metodológica suficiente ni el muestreo estadístico adecuado, tanto de centros como de ensayos como del número de informes de NGS analizados, como para sacar conclusiones categóricas, pero si que sugiere, en opinión de los autores, las siguientes reflexiones:

- La selección de pacientes potenciales candidatos a ser incluidos en los dos ensayos clínicos adaptativos usados como ejemplo en este proyecto, ha mostrado mayor precisión cuando se hace la selección en base a un Biomarker Memo, que cuando se aplican los criterios de inclusión/Exclusión publicados en *Clinicaltrial.gov*, con el consiguiente beneficio para pacientes, profesionales sanitarios y los promotores de ensayos clínicos.
- La automatización en la detección de pacientes potencialmente candidatos a ensayos clínicos basados tan sólo en criterios moleculares, edad y tipo tumoral podría ser una herramienta muy útil para reducir el tiempo que los profesionales sanitarios emplean en la búsqueda del ensayo adecuado
- La reutilización de la información genómica en la era de la medicina de precisión es crítica, dado el creciente número de terapias dirigidas que se encuentran en estudio. La información generada durante un ensayo clínico y durante la rutina clínica, puede ser aprovechada para considerar al paciente para otro ensayo clínico con terapias dirigidas si esta información se incorpora de manera sistemática a la historia clínica del paciente o existen motores de búsqueda adecuados dentro de los centros sanitarios.
- La heterogeneidad en la nomenclatura de alteraciones genómicas es una barrera para la construcción de bases de datos de vida real (RWD), inclusive en un mismo centro y podría ser también una barrera para la búsqueda automatizada de pacientes con cáncer con ciertas alteraciones genómicas.
- Paneles amplios y validados (CGP) serán cada vez más necesarios en el diagnóstico oncológico, no solo para encontrar tratamientos molecularmente dirigidos aprobados sino para la utilización de los ensayos clínicos como una forma de acceso a tratamientos innovadores molecularmente guiados.
- Esfuerzos unificados en la industria farmacéutica y de sociedades científicas en proyectos similares, son necesarios para ampliar las posibilidades de acceso a terapias dirigidas innovadoras y aumentar la investigación clínica en España.