



PERÚ

Ministerio  
de Trabajo  
y Promoción del Empleo

Seguro Social de Salud  
EsSalud

**INSTITUTO DE EVALUACIÓN DE TECNOLOGÍAS EN SALUD E  
INVESTIGACIÓN – IETSI**

**DICTAMEN PRELIMINAR DE EVALUACIÓN DE TECNOLOGÍA  
SANITARIA N.º 056-SDEPFYOTS-DETS-IETSI-2017  
EFICACIA DE ATALUREN EN PACIENTES CON DIAGNÓSTICO DE  
DISTROFIA MUSCULAR DE DUCHENNE PORTADORES DE UNA  
MUTACIÓN SIN SENTIDO EN EL GEN DE DISTROFINA  
- ACTUALIZACIÓN**

**SUBDIRECCIÓN DE EVALUACIÓN DE PRODUCTOS FARMACEÚTICOS Y  
OTRAS TECNOLOGÍAS SANITARIAS-SDEPFYOTS**

**DIRECCIÓN DE EVALUACIÓN DE TECNOLOGÍAS SANITARIAS-DETS**

**INSTITUTO DE EVALUACIÓN DE TECNOLOGÍAS EN SALUD E  
INVESTIGACIÓN-IETSI**

**SEGURO SOCIAL DE SALUD-ESSALUD**

*Octubre, 2017*



**IETSI**  
INSTITUTO  
DE EVALUACIÓN  
DE TECNOLOGÍAS  
EN SALUD E  
INVESTIGACIÓN

## EQUIPO REDACTOR:

1. Fabián Alejandro Fiestas Saldarriaga – Gerente, Dirección de Evaluación de Tecnologías Sanitarias - IETSI-ESSALUD.
2. Maribel Marilú Castro Reyes – Sub Gerente, Subdirección de Evaluación de Productos Farmacéuticos y Otras Tecnologías - IETSI-ESSALUD.
3. Verónica Victoria Peralta Aguilar – Directora, Dirección de Evaluación de Tecnologías Sanitarias- IETSI-ESSALUD.
4. Yuani Miriam Román Morillo - Equipo Técnico Evaluador, Subdirección de Evaluación de Productos Farmacéuticos y Otras Tecnologías Sanitarias - IETSI-ESSALUD.



## CONFLICTO DE INTERÉS

Los miembros del equipo redactor manifiestan no tener conflicto de interés de tipo financiero respecto al producto farmacéutico evaluado.



## FUENTES DE FINANCIAMIENTO

Seguro Social de Salud – ESSALUD.

## CITACIÓN

IETSI-EsSalud. Eficacia de ataluren en pacientes con diagnóstico de distrofia muscular de Duchenne portadores de una mutación sin sentido en el gen de distrofina - Actualizado. Dictamen Preliminar de Evaluación de Tecnología Sanitaria N.° 056-SDEPFyOTS-DETS-IETSI-2017. Lima, Perú. 2017.

## LISTA DE ABREVIATURAS

ADL	Activities of daily living
DMD	Distrofia Muscular de Duchenne
D6CM	Distancia de caminata en 6 minutos
DIGEMID	Dirección General de Medicamentos, Insumos y Drogas – Perú
EA	Evento adverso
EAS	Eventos Adversos Serios
ECA	Ensayo Controlado aleatorizado
EMA	European Medicines Agency
ETS	Evaluación de Tecnología Sanitaria
FDA	(US) Food and Drug Administration
HR	Hazard Ratio
ITT	Intention-to-treat
MA	Meta – Análisis
NICE	The National Institute for Health and Care Excellence
NSAA	North Star Ambulatory Assessment
PICO	Población, Intervención, Comparación, Desenlaces
RS	Revisión Sistemática



## CONTENIDO

I. RESUMEN EJECUTIVO .....	5
II. INTRODUCCIÓN .....	8
A. ANTECEDENTES.....	8
B. TECNOLOGÍA SANITARIA DE INTERÉS .....	10
III. METODOLOGÍA .....	11
A. ESTRATEGIA DE BÚSQUEDA .....	11
B. TÉRMINOS DE BÚSQUEDA.....	11
C. CRITERIOS DE ELEGIBILIDAD .....	12
IV. RESULTADOS .....	13
A. SINOPSIS DE LA EVIDENCIA .....	14
i. EVALUACIONES DE TECNOLOGÍAS SANITARIAS .....	14
ii. ESTUDIOS PRIMARIOS .....	23
V. DISCUSIÓN .....	28
VI. CONCLUSIONES .....	34
VII. RECOMENDACIONES .....	37
VIII. REFERENCIAS BIBLIOGRÁFICAS .....	38
IX. ANEXOS.....	39



## I. RESUMEN EJECUTIVO

- La Distrofia Muscular de Duchenne (DMD) es la forma más severa de distrofia muscular; es un trastorno recesivo ligado al cromosoma X, por lo que, afecta principalmente a los hombres. La DMD por mutación sin sentido es ocasionada por la alteración de una base del ADN que lleva a la producción incompleta de distrofina en las fibras musculares esqueléticas, lisas y cardíacas. Los síntomas de DMD aparecen a los tres años de edad y la incidencia es de 1 caso en 3600-6000 varones nacidos vivos, y sólo el 13 % de ellos, son ocasionados por mutaciones sin sentido, alcanzando una frecuencia de 1 en 27000-46000 varones nacidos vivos.
- A la fecha no existe un tratamiento específico para esta condición, el manejo habitual es de soporte que apunta a aliviar los síntomas y manejar las complicaciones. Entre ellos se incluyen: corticoides, dispositivos ortopédicos, inhibidores de la enzima convertidora de angiotensina y beta bloqueantes para el manejo de las cardiomiopatías; cirugías, asistencia ambulatoria, asistencia de movilidad (ejemplo: silla de ruedas) y ventilación artificial.
- Ataluren fue evaluado previamente por el IETSI en el Dictamen Preliminar de Evaluación de Tecnología Sanitaria N.º 020-SDEPFyOTS-DETS-IETSI-2016; en tal evaluación se consideró la evidencia derivada del estudio 007 (Bushby 2014). La eficacia de ataluren no mostró ser diferente a placebo en términos de la DC6M definido a priori como desenlace principal en este ensayo, tanto en el análisis por intención a tratar predefinido como en el modificado. Por lo que, se consideró que no existía evidencia sobre algún beneficio clínico de ataluren respecto al tratamiento de soporte.
- El presente dictamen tiene como objetivo reevaluar la eficacia y seguridad de ataluren 40 mg en el tratamiento de pacientes DMD con mutación sin sentido y en fase de declive ambulatorio; después de la publicación de los resultados del estudio 020. Asimismo, se incluyeron dos ETS, realizadas por NICE (NICE 2017) y por el comité de transparencia de la Alta Autoridad de Salud de Francia (Haute Autorité de Santé 2017) y una evaluación metodológica de la FDA.
- El estudio 020 fue enriquecido para incluir a pacientes en la fase de declive ambulatorio, en base a los resultados de un subgrupo post hoc del estudio 007, y tuvo una muestra dos veces más grande que el estudio 007. Sin embargo, el desenlace primario también falló en mostrar significancia estadística; se observó un efecto no significativo de 13 metros en el 6MWD ( $p=0.21$ ). Estos resultados ilustran claramente que los análisis post hoc con datos a la mano y completamente abiertos tienen alto riesgo de producir resultados espurios.



- Los análisis de un subgrupo de pacientes según el valor basal del 6MWD 300-400 metros del estudio 020 mostraron una diferencia estadísticamente significativa en el desenlace primario, y en los desenlaces secundarios en este subgrupo que podrían sugerir un efecto favorable de ataluren. Sin embargo, al mirar dentro de un estudio en subgrupos no pre-especificados incrementa el error tipo I substancialmente; el problema es aún mayor cuando existen muchos subgrupos, como en este caso. Los revisores de la FDA consideraron que el subgrupo con 6MWD 300-400 metros no había sido pre-especificado, ya que no fue identificado como parte del análisis de efectividad planeado y que incluyera también el control del error tipo I.
- Aunque resulten tentadores, los análisis por subgrupos conllevan a múltiples limitaciones. El más importante a ser considerado es que cualquier subgrupo que se forme no produciría grupos comparables por la falta de aleatorización, de lo cual se desprende que cualquier diferencia que se encuentre puede deberse a varios factores (no balanceados) aparte de la intervención. Los análisis por subgrupos solo pueden ser considerados como generadores de hipótesis, y para ello deben ser medianamente creíbles. Esto significa que estos resultados deben provenir de un número limitado y pre-especificado de comparaciones, para los cuales los autores han brindado las explicaciones de su elección y han especificado a priori las direcciones de los resultados. Adicionalmente, deben contar con apoyo de otros hallazgos similares y existir una plausibilidad biológica
- En el caso del estudio 020, no está disponible el protocolo que permita determinar cuántas comparaciones se pre-especificaron y cuáles fueron las razones de su elección. De nueve comparaciones que reportaron los autores, solo uno resultó siendo significativa, y con el agravante de que no fue ajustado por múltiples comparaciones, incrementando el error tipo I. No sabemos si estas nueve comparaciones forman parte de un número mayor de comparaciones, lo cual incrementa aún más la posibilidad de encontrar alguna diferencia significativa solo por azar. Una observación que hicieron los revisores de la FDA fue la dirección opuesta, a favor de placebo en los subgrupos complementarios (6MWD < 300 o  $\geq$  400 metros) en el estudio 020, sugiriendo que el efecto de tratamiento observado en el subgrupo con 6MWD basal de 300-400 metros haya podido ser el resultado de la exclusión de los subgrupos de pacientes con resultados negativos.
- Los resultados en los desenlaces secundarios (funciones motoras cronometradas) que sugirieron ser favorables con ataluren en el subgrupo con 6MWD basal de 300-400 metros provienen de un análisis post hoc. Es bien reconocido que los análisis estadísticos post hoc no pre-especificados de múltiples subgrupos y con múltiples desenlaces resultarían en falsos hallazgos positivos o significativos y no pueden ser tomados como diferencias reales. Por lo



tanto, los análisis post-hoc no apoyan la efectividad de ataluren cuando los análisis pre-especificados del desenlace primario es negativo o no significativo.

- Las recomendaciones a favor del uso de ataluren en las ETS consultadas, exponen claramente las limitaciones de la evidencia disponible, condicionando su aprobación en base a la disminución importante en su inversión. Así, la evaluación de NICE ha condicionado su aprobación a la aplicación de importantes descuentos confidenciales al precio del agente, mientras que la comisión de transparencia de Francia ha propuesto el reembolso solo del 15 % y ha comprometido al fabricante a la realización de un ensayo más grande y de mayor duración.
- La evidencia disponible respecto a ataluren proviene de dos ensayos aleatorizados, doble ciego y controlados con placebo, en los cuales los análisis de eficacia primarios preestablecidos fallaron en mostrar una diferencia estadísticamente significativa respecto a placebo. El último estudio (estudio 020) incluyó una población enriquecida según sugerencia proveniente de un análisis post hoc de un estudio previo. Los autores buscaron basar la eficacia de ataluren en base a los resultados en un subgrupo, proveniente de múltiples subgrupos comparados y sin control del error tipo I por múltiples comparaciones, y también en base a análisis post hoc, los cuales carecen de rigor metodológico. Así, en la actualidad no existe evidencia de que el tratamiento con ataluren ofrezca un beneficio clínico diferente al tratamiento de soporte en los pacientes con DMD con mutación sin sentido.
- Actualmente se ha registrado un nuevo ensayo aleatorizado, doble ciego, controlado con placebo, de mayor duración (72 semanas) y un periodo de seguimiento posterior de 72 semanas adicionales. El desenlace principal será el cambio en la *distancia caminada en 6 minutos* (DC6M) desde el inicio del tratamiento hasta las 72 semanas después. Los desenlaces secundarios evaluados serán los siguientes: la función física, la evaluación con el NSAA y la calidad de vida medida con EQ-5D, tras 72 semanas de tratamiento. Se espera que culmine el estudio en diciembre del 2021.
- Por lo expuesto, el Instituto de Evaluación de Tecnologías en Salud e Investigación- IETSI ratifica lo estipulado en el Dictamen Preliminar de Evaluación de Tecnología Sanitaria N.º 020-SDEPFyOTS-DETS-IETSI-2016, donde no se aprueba el uso de ataluren en el tratamiento de pacientes con DMD con mutación sin sentido en el gen de distrofina.



## II. INTRODUCCIÓN

### A. ANTECEDENTES



Ataluren fue evaluado previamente por el IETSI en el Dictamen Preliminar de Evaluación de Tecnología Sanitaria N.º 020-SDEPFyOTS-DETS-IETSI-2016; en tal evaluación se consideró la evidencia derivada del estudio 007 (Bushby 2014), un estudio de fase 2b, aleatorizado, doble ciego, controlado con placebo, de dos dosis de ataluren en pacientes con DMD con mutación sin sentido,  $\geq 5$  años de edad, una *distancia caminada en 6 minutos basal* (6MWD)  $\geq 75$  metros. En este estudio, no se requería que el paciente estuviera recibiendo concomitantemente terapia con corticoides. Un total de 174 pacientes fueron aleatorizados 1:1:1 para recibir placebo, ataluren 10/10/20 mg/kg, o ataluren 20/20/40 mg/kg por 48 semanas; el desenlace primario fue el cambio en el 6MWD a las 48 semanas.



La eficacia de ataluren no mostró ser diferente a placebo en términos de la DC6M definido a priori como desenlace principal en este ensayo, tanto en el análisis por intención a tratar predefinido como el modificado. En la población por intención a tratar directo, se observó una disminución de la media de la DC6M de 42.6 metros en el grupo placebo, y de 12.9 metros en el grupo de ataluren a 40 mg/kg/día (diferencia del cambio: 29.7 metros,  $p=0.149$ ). En el análisis por intención a tratar corregido, la disminución de la DC6M a las 48 semanas para placebo fue de 44.1 metros y para ataluren a 40 mg/kg/día de 12.8 metros (diferencia del cambio: 31.3 metros,  $p=0.056$ ). La rama ataluren 80 mg/kg no mostró diferencia con la rama placebo, con una media del cambio respecto al placebo a las 48 semanas de -0.1 metros ( $p=0.1$ ). El uso de ataluren a 80 mg/kg/día produjo un mayor deterioro del DC6M respecto al placebo tras 48 semanas de tratamiento, aunque no fue estadísticamente significativo.

En el análisis de los desenlaces secundarios tampoco se observaron diferencias entre el grupo de placebo y alguno de los brazos de ataluren. A las 48 semanas del tratamiento, todos los tiempos empleados para realizar las actividades motoras aumentaron respecto al tiempo empleado al inicio en todos los grupos. Sin embargo, las comparaciones de esos cambios del tiempo no fueron diferentes para ninguna comparación del grupo placebo con alguno de los dos brazos de ataluren.

El fabricante realizó varios análisis post hoc en subgrupos de pacientes en base a tres factores con posible significancia en el pronóstico (edad, uso de glucocorticoides y DC6M basal). El efecto de ataluren 40 mg sugirió mayor beneficio en el subgrupo de pacientes en fase de declive (pacientes de 7 a 16 años, con una DC6M predicha  $\leq 80$  %, tomando corticoides y con una DC6M basal  $\geq 150$  metros), observándose una media de cambio en la DC6M desde el basal hasta las 48 semanas de 49.9 metros a favor del tratamiento con ataluren 40 mg comparado con placebo ( $p=0.0096$ ).

DICTAMEN PRELIMINAR DE EVALUACIÓN DE TECNOLOGÍA SANITARIA N.º 056-SDEPFYOTS-DETS-IETSI-2017  
EFICACIA DE ATALUREN EN PACIENTES CON DIAGNÓSTICO DE Distrofia Muscular de Duchenne PORTADORES DE UNA  
MUTACIÓN SIN SENTIDO EN EL GEN DE Distrofina - ACTUALIZACIÓN

A partir de estos resultados exploratorios derivados del análisis no pre-especificado se diseñó un estudio (NCT01826487) cuyo objetivo fue evaluar el efecto del tratamiento con ataluren a 40 mg/kg/día por 48 semanas, en la *distancia caminada en 6 minutos*, la función física, la calidad de vida y la seguridad, de los pacientes con DMD con mutación sin sentido en el gen de distrofina y en fase de declive ambulatoria.



El grupo evaluador del IETSI consideró que no existía suficiente evidencia que apoye la eficacia de ataluren respecto al placebo en la DMD con mutación sin sentido, por lo que no aprobó su uso en EsSalud y recomendó la reevaluación de este dictamen cuando estuvieran disponibles los resultados del estudio NCT01826487.



El presente dictamen tiene como objetivo reevaluar la eficacia y seguridad de ataluren 40 mg en el tratamiento de pacientes DMD con mutación sin sentido y en fase de declive ambulatorio, tras de la publicación de los resultados del estudio NCT01826487 (McDonald 2017) en julio del 2017.



## B. TECNOLOGÍA SANITARIA DE INTERÉS

La PTC124 (3-(5-(2-fluorofenil)-1, 2,4-oxadiazol-3-yl)-ácido benzoico), también conocida como Ataluren (Translarna™) es una molécula pequeña de oxadiazol cuyo mecanismo de acción consiste en continuar la traducción del ARNm sobre los codones de terminación prematuros causados por la mutación sin sentido, permitiendo la síntesis de distrofina completa y funcional.

Actualmente ataluren no ha sido autorizado en la Dirección de Medicamentos, Insumos y Drogas del Perú (DIGEMID). Ataluren obtuvo una "Aprobación condicional" por la Agencia Regulatoria de Medicamentos de Europa (EMA, por sus siglas en inglés *European Medicines Agency*) en Julio 2014 (EMA 2014). Esta aprobación se encuentra ligada a la obligación de la compañía farmacéutica de presentar los resultados de los estudios de fase III para ser revisados por la EMA y reevaluar su condición. Asimismo, la Comisión Europea le ha cedido la autorización para ser comercializada dentro de la Unión Europea de manera condicional, lo que implica que la compañía está obligada a realizar más estudios del medicamento. En Estados Unidos, todavía no ha sido aprobada por el Organismo Regulatorio de Alimentos y Drogas de Estados Unidos (FDA, por sus siglas en inglés *Food and Drug Administration*) y no está disponible para su comercialización, sin embargo, posee una designación de "Fármaco Huérfano" desde el 27 de mayo del 2005 para el tratamiento de DMD y la fibrosis quística.

Ataluren está disponible en una presentación granulada (100, 250 y 1000 mg) la cual puede ser tomada en forma oral luego de ser mezclada con un líquido o un alimento semisólido (como yogurt). Ataluren debe ser tomada tres veces por día, con una recomendación de 10 mg por kilogramo de peso (mg/kg) en la mañana, 10 mg/kg al mediodía y 20 mg/kg por la noche (una dosis total de 40 mg/kg al día).

Actualmente, no se dispone del precio comercial de ataluren en el Perú. Para un niño de 18.5 kilos el costo de tratamiento diario en el mercado internacional es de \$ 833.79, por lo que, el costo anual del tratamiento asciende a \$ 300.164,4 (en soles: 1, 050,575.4) por paciente.



### III. METODOLOGÍA

#### A. ESTRATEGIA DE BÚSQUEDA

Se llevó a cabo una búsqueda sistemática de la literatura con respecto a la eficacia y seguridad de ataluren 40 mg en el tratamiento de pacientes DMD con mutación sin sentido y en fase de declive ambulatorio. La búsqueda se inició revisando la información sobre el uso del medicamento de acuerdo con entidades reguladoras como la *Food and Drug Administration* (FDA), la *European Medicines Agency* (EMA) y la Dirección General de Medicamentos y Drogas (DIGEMID). Posteriormente, se revisaron las bases de datos de PubMed, TRIPDATABASE y [www.clinicaltrials.gov](http://www.clinicaltrials.gov). Adicionalmente, se realizó una búsqueda de evaluaciones de tecnologías sanitarias y guías de práctica clínica en las páginas web de grupos dedicados a la investigación y educación en salud en general como *The National Institute for Health and Care Excellence* (NICE), y especializadas en cardiología como la Fundación del Colegio Americano de Cardiología, la Asociación Americana del Corazón, y la Sociedad Europea de Cardiología, entre otras.

#### B. TÉRMINOS DE BÚSQUEDA

Para la búsqueda de información que pueda ser empleada para responder a la pregunta PICO, se utilizaron los siguientes términos relacionados a la población de interés, la intervención, el fármaco comparador y los tipos de estudio priorizados. Términos de búsqueda:

- DMD: "Muscular Dystrophy, Duchenne"[Mesh], "Muscular Dystrophy Progressive", "Muscular Dystrophy", "Childhood Muscular Dystrophy", "Duchenne Type".
- Ataluren: "Translarna", "Ataluren", "PTC124".

Adicionalmente, se emplearon los filtros correspondientes a meta-análisis, revisiones sistemáticas, ensayos clínicos, y guías de práctica clínica en línea con los criterios de elegibilidad mencionados en la siguiente sub-sección.

Los artículos producto de esta búsqueda fueron luego elegidos manualmente tal y como se presenta en el flujograma en la sección IV.

### C. CRITERIOS DE ELEGIBILIDAD

En la selección de los estudios se priorizaron las guías de práctica clínica (GPC), evaluaciones de tecnología sanitaria (ETS), revisiones sistemáticas (RS), meta-análisis (MA) y ensayos clínicos aleatorizados (ECA) de fase III de acuerdo con la pregunta PICO. Adicionalmente, se revisaron los ensayos clínicos dentro de las revisiones sistemáticas o meta-análisis encontrados en el marco de la pregunta PICO como una fuente adicional de información y corroboración de los ensayos clínicos incluidos en el presente trabajo.



#### IV. RESULTADOS



## A. SINOPSIS DE LA EVIDENCIA

De acuerdo con la pregunta PICO, se llevó a cabo una búsqueda de evidencia científica relacionada a la eficacia y seguridad de ataluren 40 mg en el tratamiento de pacientes DMD con mutación sin sentido y en fase de declive ambulatorio. En la presente sinopsis se describe la evidencia disponible:

### Evaluaciones de tecnologías sanitarias

#### Publicaciones incluidas en la sección de descripción y evaluación:

- Haute Autorité de Santé. Commission de la Transparence - Opinión. Publicado el 5 de julio del 2017.
- Food and Drug Administration (FDA) - Briefing Document. Peripheral and Central Nervous System Drugs. Advisory Committee Meeting. September 28, 2017.
- NICE 2016- "Ataluren for treating Duchenne muscular dystrophy with a nonsense mutation in the dystrophin gene". Highly specialised technologies Published: 20 July 2016.

#### Publicaciones no incluidas en la sección de descripción y evaluación:

- European Medicines Agency. EMA. Summary of the risk management plan (RMP) for Translarna (ataluren).  
Esta ETS publicada en el 2014, no ha incluido información del reciente ECA 020.

### Estudios primarios

#### Publicaciones incluidas en la sección de descripción y evaluación:

- McDonald et al., 2017- "Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial (estudio 020).

## I. EVALUACIONES DE TECNOLOGÍAS SANITARIAS

### **Haute Autorité de Santé. Commission de la Transparence - Opinión. Publicado el 5 de julio del 2017**

La Comisión reevaluó ataluren con los datos clínicos del estudio de fase III (estudio 020), indicando que deseaba obtener los datos de eficacia sobre la edad de la pérdida de caminar y necesidad de ventilación asistida. A pesar de las limitaciones metodológicas de los análisis de subgrupos y post hoc, se decidió renovar la autorización de comercialización condicional, debido a un potencial efecto de beneficio pequeño y ausencia de tratamientos alternativos. Sin embargo, en este mismo documento se propuso que la subvención fuera del 15 %. Adicionalmente, el fabricante está

comprometido a la realización de un nuevo estudio controlado con placebo para incluir a los pacientes a partir de 5 años con una prueba de caminata de 6 minutos  $\geq 150$  metros, con una duración de 18 meses seguida de una fase de seguimiento de 18 meses. Este estudio se encuentra registrado en la página de [clinicaltrials.gov](http://clinicaltrials.gov) desde el 7 junio del 2017 y se espera el informe final para diciembre del 2021.

El estudio 020 incluyó 230 pacientes de 7 a 14 años con DMD con mutación sin sentido, y aleatorizados en dos grupos para recibir placebo o ataluren (dosis diaria total de 40 mg/kg/día) durante 48 semanas. Los pacientes tenían un promedio de 4 años de edad en el diagnóstico. El promedio de la distancia según el 6WMT fue aproximadamente 363 metros, sabiendo que el 37 % de los pacientes tenían un 6WMT < 350 metros, y casi el 84 % de los pacientes recibió corticosteroides durante más de un año.

Después de 48 semanas de tratamiento, no hubo diferencia en el desenlace primario, la *distancia caminada en 6 minutos* entre el grupo ataluren y el grupo placebo ( $p=0.21$ ). La disminución de la *distancia caminada en 6 minutos* fue de 42 metros en el grupo ataluren y 58 metros en el grupo placebo a las 48 semanas.

En lo que respecta al resultado del análisis primario de eficacia, el estudio 020 no pone en evidencia diferencias entre placebo y ataluren sobre la base del declive de la distancia caminada según el 6WMT.

En esta evaluación, todos los resultados de los desenlaces secundarios y por subgrupos fueron considerados exploratorios debido a que el procedimiento del análisis estadístico era jerárquico; es decir, si el desenlace principal era estadísticamente significativo, se evaluaría el primer desenlace secundario del tiempo hasta el deterioro persistente del 10 % en el 6MWD con un valor  $p$  bilateral de 0.05. Si este primer desenlace secundario resultaba estadísticamente significativo, se aplicaría el procedimiento de Hochberg para el control de múltiples comparaciones con un nivel de 0.05 para las siguientes funciones motoras: subir cuatro escalones, bajar cuatro escalones y correr/caminar 10 metros. Debido a que el desenlace primario de eficacia resultó no significativo, todos los demás análisis no fueron ajustados por multiplicidad de comparaciones y se presentaron de manera descriptiva.

El fabricante proporcionó datos de análisis en subgrupos definidos a posteriori según el 6WMT, estos análisis fueron incorporados al plan de análisis estadístico en 2015. Por lo que estos análisis solo pueden considerarse como información exploratoria que podría sugerir eficacia heterogénea según el 6WMT basal.

El grupo evaluador mencionaron varias limitaciones metodológicas del estudio de fase III 020 como las siguientes:

- Los criterios de elegibilidad no permiten excluir la inclusión de pacientes con distrofia muscular de Becker, la cual tiene una evolución diferente de DMD.
- A pesar de la aleatorización, algunas características de la enfermedad, especialmente la localización de los exones, no estaban equilibrados entre los grupos; este fenómeno supone el cuestionamiento de otras características no evaluadas dentro del estudio y que tienen un impacto sobre el fenotipo y evolución

de la enfermedad, por ejemplo, la representación de la distrofina residual, y la dosis y tipo de corticoide utilizado en tratamientos paralelos al estudio.

- La prueba de *caminata de 6 minutos* elegida como criterio principal de juicio es de pertinencia y de interpretación clínica cuestionable, en razón del aspecto motivacional que implica y de las características propias de la DMD, entre ellas la fatiga muscular variable y la pérdida cognitiva. Su replicabilidad por un mismo infante y entre los centros no es segura.
- La duración del estudio es demasiado corta para evaluar la eficacia del tratamiento y teniendo en cuenta el perfil variable de la evolución de la enfermedad.

Se planeó realizar un meta-análisis con los resultados de los dos estudios (007 y 020). Los evaluadores señalaron que los criterios de inclusión de estos dos estudios no fueron estrictamente semejantes, en particular en términos del umbral de la distancia caminada. Este meta-análisis sugiere que la media de las diferencias de medias de la disminución del 6MWT fueron 20 metros entre los dos grupos, a favor del grupo ataluren. Sin embargo, la falta de control por comparaciones múltiples dificulta la interpretación de estos resultados.

Los resultados del meta-análisis solo pueden ser interpretados de manera exploratoria por la ausencia de medidas de control de la multiplicidad de pruebas. La diferencia promedio constatada entre los grupos en términos de la *distancia caminada en 6 minutos* es débil.

Por otro lado, los umbrales elegidos para definir los subgrupos en el estudio 020 (< 300 metros, 300-400 y  $\geq$  400 metros) diferían del elegido en el de fase 2b (estudio 007) (< 350 metros,  $\geq$  350 metros), con resultados que, aunque no son robustos, parecen ser contradictorios. Con estos resultados es difícil la identificación potencial de pacientes que puedan beneficiarse con ataluren. Los datos proporcionados no permiten aún responder preguntas de criterios de utilización de este medicamento, y de su seguimiento o de su interrupción luego de la pérdida de movilidad.

Teniendo en cuenta la ausencia de demostración de eficacia de ataluren vs placebo en términos de un desenlace poco pertinente clínicamente y de las limitaciones en términos de la extrapolación de los resultados, no se espera que ataluren tenga un impacto sobre la morbimortalidad y la calidad de vida. En consecuencia, ataluren no aporta respuesta a la necesidad de salud médica no cubierta identificada.

A pesar de esta evaluación claramente negativa de la calidad de la evidencia que aporta el estudio 020 respecto a la eficacia de ataluren respecto a placebo, la decisión de aprobar su comercialización condicionada se basó en la falta de una alternativa terapéutica. La aprobación condicionada está establecida a la realización de un ensayo más grande y de mayor duración (18 meses) y con seguimiento posterior de 18 meses adicionales. Además, en acorde al bajo nivel de confianza de los resultados, la comisión opinó favorablemente a la inscripción en la lista de medicamentos reembolsables a los asegurados, pero con una tasa de reembolso del 15 %.



### Food and Drug Administration (FDA) - Briefing Document. Peripheral and Central Nervous System Drugs. Advisory Committee Meeting. September 28, 2017

Este documento contiene información preparada por el panel de miembros del comité consultor de la FDA. Este mismo comité ya había *rechazado la admisión*<sup>1</sup> de ataluren para ser evaluado como nuevo medicamento en el 2011, en el cual se evaluó el estudio 007. En diciembre del 2015 la compañía fabricante de ataluren envió una solicitud a la FDA incluyendo los resultados del estudio 020 y 007, así como un meta-análisis de ambos estudios. La nueva evaluación que hizo la FDA resaltó muchas limitaciones en el estudio 020, por las que nuevamente fue *rechazada la admisión* de la aplicación de ataluren para ser revisado.

El estudio 020 fue un estudio con mayor potencia enriquecido con una población explícita derivada de los hallazgos de subgrupos post hoc del estudio 007. Adicionalmente, solo se incluyó la dosis más baja del estudio 007.

El solicitante había previamente descrito un efecto del tratamiento de 46 metros en el 6MWT a las 48 semanas en la población ambulatoria en fase de declive del estudio 007. Por lo que era particularmente preocupante que, a pesar del enriquecimiento, el análisis pre-especificado de eficacia del estudio 020 fue claramente negativo ( $p=0.21$ ), con menos de la tercera parte del tamaño del efecto observado en el estudio 007 en la semana 48 (13 metros), a pesar de reclutar casi tres veces más el número de pacientes ( $N=228$ ) que fue parte de la población en ADP post hoc del estudio 007 ( $N=63$  para la comparación de la dosis baja con placebo). Adicionalmente, solo uno de los múltiples desenlaces secundarios del estudio fue significativo en la población por intención a tratar (ITT, por sus siglas en inglés intention-to-treat).

El fabricante también presentó un número de análisis que carecían de interpretabilidad (i.e análisis por subgrupos que no fueron controlados por múltiples comparaciones) o eran completamente análisis post hoc en un intento de explicar la falla del análisis primario pre-especificado.

A continuación, se describen las limitaciones metodológicas que el comité consultor de la FDA identificó en los análisis estadísticos de los desenlaces secundarios y exploratorios; así como en los análisis por subgrupos y análisis post hoc.

<sup>1</sup> El rechazo a la admisión ("refuse to file" en inglés) es una herramienta regulatoria importante de la FDA para ayudar al Centro de Evaluación e Investigación de Medicamentos (CDER) a evitar una revisión innecesaria de postulaciones incompletas o ciertas postulaciones debieron haber sido enviados como una postulación abreviada de medicamento nuevo (FDA Manual of policies and procedure). El rechazo a la admisión es el primer paso del proceso de postulación de un medicamento nuevo, y paso previo a la revisión formal de la postulación. Los comentarios dados por la FDA deben ser resueltos en un período de 30 días desde la postulación inicial. Si no se responden satisfactoriamente los comentarios de la carta de rechazo a la admisión, el proceso termina allí.

### **El tamaño del efecto del desenlace principal era menor del que se sugirió en el análisis post hoc de un subgrupo de pacientes ADP del estudio previo**

Los revisores de la FDA expresaron preocupación por la falta de interpretabilidad de los datos. Notaron que era muy informativo que el tamaño del efecto en el estudio 020 era menor, menos de la tercera parte (13 metros vs 46 metros) del que se observó en la población ADP post hoc del brazo de la dosis baja de ataluren del estudio 007, a pesar de haber enriquecido activamente a esta población y haber reclutado más del 3.5 veces el número de pacientes (N=228 vs N=63 [población ADP del brazo de dosis baja de ataluren y placebo en el estudio 007]). Los revisores del comité consultor de la FDA consideraron que este resultado era una ilustración directa de la naturaleza frecuentemente engañosa que son los análisis tentadores post hoc realizados con datos completamente abiertos, o análisis exploratorios de subgrupos sin ningún control estadístico del error tipo I en el marco de un estudio negativo.

Ante esta observación, los fabricantes del medicamento mencionaron que a pesar de los esfuerzos de enriquecer el estudio con pacientes en la fase de declive ambulatoria; la población del estudio 020 permanecía siendo heterogénea. El rango del 6MWD basal en el estudio 020 era amplio (142 - 521 metros) y el promedio del 6MWD basal era 23 metros mayor en la población ITT del estudio 020 comparado con el subgrupo ADP del estudio 007. Los revisores de FDA evaluaron el impacto de una media más alta del 6MWD en el estudio 020, realizando un análisis en el subgrupo de pacientes con un 6MWD basal < 433 metros. Se eligió como punto de corte a 433 metros porque este valor era el promedio del subgrupo ADP del estudio 007. Los resultados de este análisis mostraron un ligero incremento del efecto, pero sin significancia estadística (15 metros,  $p=0.23$ ), el cual sigue siendo más pequeño que el efecto del tratamiento observado en el subgrupo ADP del estudio 007. Por lo tanto, los evaluadores del comité consultor de la FDA concluyeron que este resultado sugiere que la discrepancia entre los efectos del tratamiento entre ambos estudios no puede ser explicado por las diferencias en la media del 6MWD basal.

### **Los análisis post hoc del desenlace primario**

El fabricante del medicamento también envió a la FDA los resultados de dos análisis post hoc del desenlace primario usando un modelo no lineal de dos partes y un análisis slope-based. El fin de estos análisis era evaluar el efecto del tratamiento en el contexto de pacientes que podían deambular durante el tiempo que duró el estudio y en aquellos que perdieron la habilidad de deambular. Los revisores de la FDA encontraron varias limitaciones en estos análisis, siendo las principales las siguientes:

- El modelo no lineal de dos partes resultó en una diferencia de tratamiento marginalmente significativo a favor de ataluren de 13.9 metros ( $p=0.04$ ). Sin embargo, los revisores notaron que los valores predichos no eran independientes, debido a que los mismos estimados de los parámetros del modelo fueron usados para predecir el cambio en el 6MWD para todos los pacientes en el grupo de ataluren. Tal correlación o dependencia, así como la incertidumbre de los

estimados de los parámetros del modelo no fueron apropiadamente incorporados en el análisis. Se podría haber infra estimado el error estándar, así como las covarianzas de los valores predichos no fueron apropiadamente contabilizadas.

- El modelo slope-based no resultó significativo estadísticamente ( $p=0.1$ ). Además, los revisores estadísticos de la FDA utilizaron un gráfico de residuales para examinar si el modelo lineal se ajustaba bien a los datos. Sin embargo, se evidenció que los residuales del modelo slope-based no estaban distribuidos aleatoriamente alrededor de la línea de cero, sugiriendo que el modelo no se ajustaba apropiadamente, limitando la interpretabilidad de estos análisis post hoc.

### Pruebas funcionales cronometradas

En el contexto de un resultado no significativo (negativo) del desenlace primario, los análisis de estos desenlaces secundarios solo pueden ser considerados como exploratorios (no hubo tampoco un plan del control estadístico del error tipo I). Si el tiempo que le tomaba al paciente realizar la prueba excedía los 30 segundos, o si el paciente no podía realizar la prueba debido a la progresión de la enfermedad, se usaba el valor de 30 segundos. De las tres pruebas funcionales, solo la prueba de subir cuatro escalones fue estadísticamente significativa a favor de ataluren (sin ajuste). Se consideraron que estos resultados eran consistentes con los resultados negativos del análisis primario.

Se realizaron dos análisis de sensibilidad definidos en el protocolo para estas pruebas funcionales cronometradas: (1) usando como límite superior 45 segundos, y (2) usando el valor más alto observado para la prueba determinada. Los evaluadores de la FDA concluyeron que los resultados de los análisis de las pruebas funcionales cronometradas eran sensibles al manejo de máximo valor si el paciente no podía realizar la prueba debido a progresión de la enfermedad. El valor nominal de  $p$  de la prueba de subir cuatro escalones aumentó a 0.04 cuando se usó un tiempo máximo de 45 segundos, y se perdió la significancia estadística de la prueba cuando se usó el mayor valor observado.

### Análisis por subgrupos

El estudio 020 evaluó la eficacia de ataluren en nueve subgrupos, cinco de los cuales se basaron en el valor basal del 6MWD únicamente. En el contexto de resultados no significativos en el análisis de eficacia primario, estos resultados solo pueden ser vistos como exploratorios. No hubo un ordenamiento pre especificado de la importancia relativa de ninguno de estos subgrupos en el plan de análisis.

De los nueve subgrupos que fueron evaluados; solo uno, el de pacientes con 6MWD basal de 300-400 metros, mostraron un valor de  $p$  nominal (es decir no ajustado por múltiples comparaciones)  $=0.007$ , resultado a favor de ataluren en el cambio respecto al basal en el 6MWD. Sin embargo, los evaluadores de la FDA notaron que los resultados en los subgrupos complementarios  $< 300$  metros y  $\geq 400$  metros favorecieron numéricamente a placebo. Estas observaciones sugirieron que los resultados obtenidos



en el subgrupo de 300 a 400 metros hayan sido solo por azar, o el resultado de excluir a pacientes con resultados negativos o no favorables para ataluren.

Tampoco se pre-especificó un ajuste por múltiples comparaciones para el análisis del subgrupo 300 a 400 metros (o ninguno de los nueve subgrupos). El fabricante realizó una prueba de permutación post hoc y el resultado fue un valor de p ajustado en el subgrupo de  $\geq 300$  a  $< 400$  metros de 0.036. Ante lo cual, los revisores de la FDA manifestaron que, sin una pre-especificación, se pueden realizar muchos posibles análisis post hoc para ajuste por múltiples comparaciones. Por lo tanto, el valor de p era difícil de interpretar.

#### **Cambio en las pruebas funcionales cronometradas (PFC) a las 48 semanas por subgrupos del estudio 020**

Se observaron valores de p nominales significativas a favor de ataluren en la prueba del tiempo para subir cuatro escalones y en la de bajar cuatro escalones las PFC en el subgrupo de pacientes con 6MWD basal 300-400 metros. Sin embargo, aunque no significativo estadísticamente, se observó que el tiempo en correr/caminar 10 metros y el de tiempo para subir cuatro escalones, favorecieron numéricamente a placebo en el subgrupo complementario con 6MWD  $\geq 400$  metros, el cual es consistente con los resultados no significativos del análisis del desenlace principal.

Los revisores de la comisión consultora de la FDA resaltaron que aún no se han establecido los puntos de corte específicos del 6MWD basales que puedan predecir un declive del DMD durante un ensayo clínico. Otros factores conocidos y desconocidos también pueden impactar en la progresión y existe importante variabilidad en el declive incluso en el subgrupo de 300-400 metros. Este subgrupo representa el 43 % de los pacientes incluidos en el estudio 020. Sin embargo, el solicitante no remarcó ningún énfasis en los análisis en este subgrupo de pacientes sino hasta después de que la base de datos del estudio 020 fuera revelada (es decir levantar el doble ciego), debilitando aún más cualquier conclusión que pueda ser formulada en base a estos análisis exploratorios.

#### **Subgrupo con 6MWD 300-400 metros del estudio 007**

El solicitante sugirió que los resultados en el subgrupo exploratorio con 6MWD entre 300-400 metros del estudio 020 habían sido retrospectivamente replicados en el estudio 007 (diferencia en el Least Square (LS) de 51 metros en el estudio 007 y 47 metros en el estudio 020). En principio, los evaluadores señalaron que esta suerte de intento retrospectivo para validar un hallazgo exploratorio de un ensayo no significativo en el análisis de eficacia primario, con análisis post hoc de datos develados (sin cegamiento de los evaluadores) de un estudio previo negativo carecía de rigor científico. Asimismo, ninguno de los resultados del 6MWT o de las pruebas funcionales cronometradas fueron estadísticamente significativas a favor de ataluren para el subgrupo con 6MWD de 300-400 m en el estudio 007 para ambas dosis de ataluren.

Es importante resaltar que estos subgrupos fueron seleccionados de diferentes poblaciones respecto a la edad o uso de corticosteroides, limitando su comparabilidad. Los evaluadores de la FDA realizaron unos análisis, en los que se observaron varias tendencias exploratorias en el subgrupo de pacientes con 6MWD 300-400 metros del estudio 007 que no tomaban corticoides, quienes no son representativos del estudio 020 en el que todos los pacientes tomaban corticoides. Estos hallazgos ilustran que los patrones de los resultados podrían ser muy probablemente resultado del azar.

Los análisis realizados por la FDA mostraron que el único efecto aparentemente consistente que se observó entre los dos estudios fue en el subgrupo con 6MWD de 300-400 metros. Sin embargo, como se indicó previamente, existen potencialmente importantes diferencias entre estos grupos respecto al uso de corticoides y la edad. En los otros subgrupos en base al 6MWD basal, los efectos del tratamiento fueron mayores en el estudio 007, a pesar de ser un estudio substancialmente más pequeño (para la comparación entre el placebo y el brazo con la menor dosis de ataluren). Asimismo, una observación preocupante fue que los subgrupos complementarios < 300 metros y  $\geq$  400 metros mostraron tendencias numéricas opuestas, es decir a favor de placebo, en ambos estudios. Estos hallazgos refuerzan la posibilidad de que los resultados obtenidos en el subgrupo de 6MWD de 300-400 metros podrían ser probablemente explicados por el azar y no por un efecto del tratamiento.

#### **Evaluación del North Star Ambulatory Assessment (NSAA)**

El NSAA consiste en 17 ítems que se enfocan en varios aspectos de la deambulación en DMD (e.g., habilidad para levantarse del suelo, habilidad para pasar de la posición supina a la sentada y de sentada a la erecta, saltar, etc.). Cada ítem se puntúa con 2 (normal – sin modificación obvia de la actividad), 1 (método modificado, pero consigue el objetivo sin asistencia de otros), o 0 (incapacidad para realizar la actividad independientemente). El puntaje total del NSAA varía entre 0-34, siendo los mayores puntajes los que representan mayores habilidades. El NSAA no fue evaluado en el estudio 007 y fue incluido como un desenlace exploratorio en el estudio 020, sin control estadístico del error tipo I. Se planeó explorar los resultados del NSAA usando tanto los puntajes ordinales (0-34) y con la transformación del puntaje ordinal a una lineal desde 0 (peor) a 100 (mejor).

Ninguno de estos dos análisis del NSAA en la población ITT resultó estadísticamente significativo (puntuación ordinal 0-34, valor  $p=0.13$  y la puntuación lineal  $p=0.27$ ), y fueron consistentes con los resultados no significativos del análisis del desenlace primario y con todos, excepto uno, de los resultados no significativos de los desenlaces secundarios.

El fabricante también presentó los resultados aparentemente significativos del NSAA en el subgrupo de pacientes con 6MWD basal de 300-400 metros ( $p$  valor con puntuación ordinal  $=0.04$  y  $p$  valor con puntuación lineal (0-100)  $=0.04$ ). Sin embargo, estos resultados están sujetos a las mismas observaciones y potenciales sesgos para este subgrupo que ya han sido discutidos, incluyendo el hecho de que estos resultados en el subgrupo > 400 metros numéricamente favorecieron al placebo.



El solicitante luego condujo un análisis post hoc que evaluó a los pacientes que habían declinado desde 2 a 1 o desde 1 a 0 durante el curso del estudio. El fabricante agregó un gráfico en el estudio publicado que tenía el fin de mostrar que más pacientes del grupo placebo declinaron de 2 a 1 o de 1 a 0 en cada uno de los 17 ítems del cuestionario NSAA.

Estos datos son el resultado de un análisis post hoc de un desenlace exploratorio que por sí mismo falló en demostrar, aunque sea una significancia estadística con dos métodos de análisis. Asimismo, los pacientes que declinaron de 2 a 1, constituyen un cambio clínico igualmente válido, pero no fueron presentados. En este análisis, considerando el paso de 2 a 1 punto, más pacientes del grupo de ataluren declinaron en 10 categorías, mientras que más pacientes del grupo placebo declinaron solamente en 2. Estos resultados resaltan las limitaciones inherentes en la presentación selectiva de los datos post hoc.

### Meta-Análisis

El único meta-análisis pre-especificado que fue incluido en el protocolo del estudio 020 combinó la población ADP post hoc del estudio 007 (solamente de la baja dosis) y la población ITT del estudio 020.

Al respecto los evaluadores de la FDA detallaron las razones por las que estos análisis eran inapropiados y no podían ser considerados como evidencia. La población ADP del estudio 007 se generó a partir de múltiples y simultáneos ajustes post hoc a la base de datos para encontrar un subgrupo con el menor valor p de significancia. Esta población fue luego prospectivamente evaluada en un ensayo clínico bien diseñado que fue claramente negativo para demostrar alguna diferencia con placebo. Es por tanto claramente inapropiado considerar los resultados de un ensayo negativo y combinarlos con los resultados post hoc de otro ensayo también negativo (estudio 007) con el objetivo de generar un análisis conjunto y con valores de p nominales, es decir sin control del error tipo I.

El fabricante también presentó los resultados de un meta-análisis post hoc que combinó la población ITT modificada post hoc del estudio 007 (dosis baja) con la población ITT del estudio 020. Este procedimiento de combinar datos sin cegamiento de dos estudios negativos de manera post hoc para generar hallazgos estadísticamente significativos con datos disponibles a la mano no deben ser considerados superiores a los resultados claramente negativos de los ensayos individuales cuando fueron analizados de acuerdo a un plan de análisis pre-especificado. Adicionalmente, este análisis combina dos poblaciones diferentes respecto al uso de corticoides (que varían desde ninguno a uso >12 meses), la edad ( $\geq 5$  vs  $\geq 7$  años), y el 6MWD basal ( $\geq 75$  metros vs  $\geq 150$  metros).



**NICE 2016- "Ataluren for treating Duchenne muscular dystrophy with a nonsense mutation in the dystrophin gene". Highly specialised technologies Published: 20 July 2016. nice.org.uk/guidance/hst3**

Esta ETS recomienda el uso de ataluren para el tratamiento de la distrofia muscular de Duchenne con mutación sin sentido en personas a partir de los 5 años que puedan caminar, y solamente si se cumplen las siguientes condiciones:



- El fabricante de ataluren provea del producto farmacéutico con el descuento acordado.
- Se cumpla el acuerdo de acceso gestionado, el fabricante recogería información de la calidad de vida y realizaría un análisis comparativo a los 2 años.



El comité de NICE concluyó que debido a las incertidumbres acerca de los beneficios clínicos en la población relevante, ataluren representaría una aceptable relación calidad-precio para el sistema nacional de salud del Reino Unido (NHS) únicamente si se ofrecía este medicamento en el contexto de un acuerdo de acceso gestionado a un precio con descuentos incorporados por el esquema de acceso al paciente y la aplicación de otros componentes financieros que reducen los costos totales para el NHS.



El comité concluyó que no había mejora significativa en la tasa de declive en el 6MWD con ataluren comparado con mejor tratamiento de soporte en la población ITT de los estudios 007 y 020. El fabricante propuso el esquema de acceso al paciente con el que ataluren podría ser brindado con un descuento; el descuento comercial que acordaron el NHS y el fabricante es confidencial y no fue reportando en este documento. El acuerdo de acceso gestionado incluye componentes financieros adicionales que reducen más los costos para el NHS y que están vigentes durante el período de acuerdo.

Aunque el comité de NICE permanecía preocupado de que el beneficio de ataluren no había sido mostrado en la población para la cual estaba dirigida en la práctica clínica, el comité consideró que, en base a la evidencia actualmente disponible, los potenciales beneficios asociados con ataluren justificarían su alto costo cuando se aplicaran importantes descuentos incluidos en el esquema de acceso al paciente y componentes financieros confidenciales del acuerdo de acceso gestionado.

## ii. ESTUDIOS PRIMARIOS

**McDonald et al., 2017- "Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial (estudio 020)**

Este estudio 020 fue un ECA doble ciego controlado con placebo. Se reclutaron a pacientes pertenecientes a la *fase de declive ambulatoria* (ADP, por sus siglas en inglés). Esta población en ADP tenía las siguientes características: edad 7 a 16 años, uso estable de corticoides, un 6MWD basal  $\geq 150$  metros, pero  $\leq 80$  % del predicho para la edad y

altura. Es decir, el estudio 020 tenía como objetivo demostrar los efectos sugeridos en un análisis post hoc, después del levantamiento del cegamiento de datos, en un subgrupo de pacientes del estudio previo 007.

El estudio se llevó a cabo desde marzo 2013 hasta agosto 2015 en varios países: Estados Unidos, España, Francia, Turquía, Chile, Alemania, Reino Unido, Italia, Australia, Polonia, Canadá, Suecia, Bélgica, Brasil, Corea del Sur, Suiza, República Checa e Israel.

Los criterios principales de inclusión fueron los siguientes:

- Sexo masculino de 7 a 16 años.
- Prueba fenotípica de distrofinopatía, en base a la presencia de signos y síntomas clínicos característicos (debilidad muscular proximal, marcha de pato, y signo de Gowers).
- Hasta los 6 años, aumento de la creatinina quinasa sérica y dificultad para caminar.
- Mutación sin sentido del gen de la distrofina determinada por secuenciación de genes.
- Uso de corticosteroides (prednisona, prednisolona o deflazacort) sistemáticamente desde  $\geq 6$  meses y dosis estable durante  $\geq 3$  meses.
- La distancia recorrida durante la prueba de 6 minutos de caminata  $\geq 150$  metros y  $\leq 80$  % de distancia esperada para la edad y altura del paciente.

Los criterios de exclusión principales fueron los siguientes:

- Tratamiento con un antibiótico amino-glucósido en los 3 meses anteriores.
- Cualquier cambio en la profilaxis o tratamiento de la insuficiencia cardíaca congestiva en los 3 meses anteriores.
- Tratamiento en curso con un anticoagulante a base cumarínico (por ejemplo, warfarina), fenitoína, tolbutamina o paclitaxel.
- Tratamiento previo con ataluren.
- Cirugía mayor en los seis meses previos o programada durante el estudio.
- Tratamiento continuo con un inmunosupresor (que no sea corticosteroides).
- Necesidad de asistencia ventilatoria diurna.
- Signos y síntomas no controlados de insuficiencia cardíaca congestiva (estadio ACC/AHA estadio C o D).

Para detectar una diferencia en la distancia recorrida durante el 6MWT de 30 metros a las 48 semanas, con una potencia del 85 %, y un riesgo  $\alpha$  bilateral = 0.05, se requería incluir a 105 pacientes en cada uno de los grupos. Un total de 230 pacientes fueron asignados aleatoriamente a dos grupos con una razón 1:1 para ser tratados por 48 semanas con

- Ataluren 10 mg/kg por la mañana, 10 mg/kg al mediodía y 20 mg/kg por la noche (para una dosis diaria total de 40 mg/kg) oralmente, durante 48 semanas (n=115).
- Placebo (n=115).



La aleatorización se estratificó en función de la edad, duración de la terapia con corticosteroides y la distancia recorrida durante el 6MWT al inicio del estudio. Las evaluaciones se realizaron durante las visitas cada ocho semanas. Al final de la fase de doble ciego del tratamiento, todos los pacientes fueron elegibles para recibir ataluren 10,10, 20 mg/kg de manera abierta en un estudio de seguimiento (PTC124-GD-020e-DMD). Cuatro pacientes en el grupo de placebo interrumpieron el tratamiento y cinco en el grupo de ataluren.

El desenlace principal fue el cambio respecto al basal de la distancia caminada en el test de caminata de 6 minutos a las 48 semanas (6MWD). Los secundarios fueron:

1. Variación de la función motora proximal evaluada por las pruebas funcionales cronometradas (EFM): el tiempo necesario para recorrer 10 m, el tiempo necesario subir o bajar 4 escalones de escaleras.
2. Retraso en el empeoramiento persistente del 10% de la distancia recorrida durante el 6MWT.

Se incluyó también los siguientes desenlaces exploratorios:

1. Variación de la función física medida por la puntuación NSAA.
2. Variación de la calidad de vida usando la escala PDCOI y un cuestionario elaborado por el fabricante del medicamento.

El análisis de eficacia se basó en la población ITT, consistente en todos los pacientes que fueron aleatorizados y que tenían un valor válido de 6MWD basal y al menos un valor válido posterior.

Después de imputar 100 datos faltantes, se utilizó la prueba estadística ANCOVA (evalúa si la media de la variable primaria *distancia recorrida en 6 minutos* son iguales entre el grupo de ataluren y el de placebo; ajustado por otras variables continuas que no son de interés primario). El modelo de ANCOVA incluyó los siguientes factores de estratificación: la edad, la duración del uso de corticoides, la categoría 6MWD basal y el valor basal de 6MWD.

Se agregaron nueve análisis por subgrupos utilizando el modelo ANCOVA con un factor de interacción de tratamiento por subgrupo, con el fin de explorar el efecto en el subgrupo del 6MWD. Estos fueron los nueve subgrupos, cinco de los cuales se basaron en el valor basal del 6MWD:

- Factor de estratificación según 6MWD basal (> 350 metros vs < 350 metros).
- 6MWD basal (< 300 metros, ≥ 300 a < 400 metros, ≥ 400 metros).
- Duración del uso de corticoides previo (≥6 a <12 meses vs ≥12 meses).
- Edad al inicio (< 9 años vs ≥ 9 años).



## Resultados

Las características demográficas de los pacientes en ambos grupos fueron comparables. La edad media fue 9 años ( $\pm 1,7$ ) con edades comprendidas entre 7 y 14 años. La edad media al diagnóstico de la enfermedad fue de 4,4 años ( $\pm 2,1$ ). La media de la distancia caminada en 6 minutos fue aproximadamente 363 metros. El 37 % de los pacientes tenía una 6WMT < 350 metros. Cerca del 84 % de los pacientes recibían corticosteroides por más de un año.

### Desenlace primario: 6MWD

Después de 48 semanas, no se encontraron diferencias entre los grupos de ataluren y placebo en el desenlace principal (cambio en el 6WMT entre ataluren y placebo: 13 metros; IC 95 % -7.44 a 33.39,  $p=0.21$ ). Es decir, los resultados del análisis primario fueron negativos.

### Desenlaces secundarios

La proporción de pacientes con al menos un 10 % de empeoramiento en el 6MWD en la semana 48 fue 46 % y 43% en el grupo de placebo y ataluren, respectivamente. El Hazard Ratio (HR) para ataluren versus placebo fue 0.75 ( $p=0.160$ ).

A las 48 semanas, solo una prueba cronometrada (descenso de 4 escalones) alcanzó significancia estadística ( $p=0.012$ ). Aunque el resto de las pruebas de función motriz mostraron que los tiempos se incrementaron ligeramente en el grupo de ataluren respecto al grupo placebo, ninguna alcanzó diferencia estadísticamente significativa. Los análisis realizados consideraron que, si el tiempo para realizar la prueba excedía los 30 segundos o que, si el paciente no podía realizar la prueba debido a progresión de la enfermedad, se utilizaría un valor de 30 segundos.

### Desenlaces exploratorios

No hubo diferencias entre ataluren y placebo en términos de la calidad de vida evaluada con el cuestionario PODCI ni en las actividades de la vida diaria evaluada con ADL (activities of daily living).

## Ensayos clínicos en curso

En la página de ClinicalTrials.gov está registrado el siguiente estudio aún en progreso:

1. A Phase 3 Randomized Double Blind Placebo Controlled Efficacy and Safety Study of Ataluren in Patients with Nonsense Mutation Duchenne Muscular Dystrophy (nmDMD) and Open Label Extension, con número de registro NCT03179631 y auspiciado por PTC Therapeutics.

Este estudio fue registrado el 7 de junio del 2017 y fue actualizado al mes siguiente el 14 de julio del 2017. Se trata de un estudio fase III, aleatorizado, doble ciego y controlado con placebo. El tiempo de duración serán 72 semanas y tendrá un tiempo de extensión de 72 semanas más. Actualmente se encuentra en fase de reclutamiento.



El objetivo principal es examinar los efectos a largo plazo del tratamiento con ataluren 40 mg/k/día, dividido en tres dosis: 10, 10, 20 mg/kg. Se planea reclutar un total de 250 pacientes con los siguientes criterios de inclusión: niños varones  $\geq$  de 5 años con DMD con mutación sin sentido, con una prueba basal de la DC6M  $\geq$  150 metros, uso estable de corticoides por un mínimo de 12 meses y habilidad de realizar las pruebas funcionales dentro de 30 segundos. En el total de pacientes se incluirá a 160 (64 %) pacientes que cumplan con los siguientes criterios de inclusión: niños varones de 7 a 16 años con una prueba basal de la DC6M  $\geq$  300 metros, tiempo para pasar de la posición supina a la erecta  $\geq$  5 segundos).

El desenlace principal será el cambio en la *distancia caminada en 6 minutos* (DC6M) desde el inicio del tratamiento hasta las 72 semanas después. Los desenlaces secundarios evaluados son los siguientes: la función física, la evaluación con el North Star Ambulatory Assessment y la calidad de vida medida con EQ-5D, tras 72 semanas de tratamiento.

Según el registro del estudio la fecha estimada para el cumplimiento con la información principal será en marzo del 2020, y el cumplimiento del estudio en diciembre del 2021.



## V. DISCUSIÓN

Ataluren fue evaluado previamente por el IETSI en el Dictamen Preliminar de Evaluación de Tecnología Sanitaria N.° 020-SDEPFyOTS-DETS-IETSI-2016. En tal evaluación se consideró la evidencia derivada del estudio 007 (Bushby 2014). La eficacia de ataluren no mostró ser diferente a placebo en términos de la DC6M definido a priori como desenlace principal en este ensayo, tanto en el análisis por intención a tratar predefinido como en el modificado. Por lo que, se consideró que no existía evidencia sobre algún beneficio clínico de ataluren respecto al tratamiento de soporte para aprobar ataluren en el tratamiento de pacientes con DMD con mutación sin sentido.

El presente dictamen tiene como objetivo reevaluar la eficacia y seguridad de ataluren 40 mg en el tratamiento de pacientes DMD con mutación sin sentido y en fase de declive ambulatorio, tras de la publicación de los resultados del estudio NCT01826487 (McDonald 2017, estudio 020) en julio del 2017. En este estudio, se reclutaron a pacientes en fase de declive ambulatorio, población derivada de un análisis post hoc del estudio 007 en el que se sugirió un potencial efecto con el tratamiento con ataluren. Asimismo, se hizo una nueva búsqueda de la literatura científica, que resultó en la identificación de dos ETS, realizadas por NICE (NICE 2017) y por el comité de transparencia de la Alta Autoridad de Salud de Francia (Haute Autorité de Santé 2017), y una evaluación metodológica de la FDA. Por otro lado, la EMA no ha publicado ninguna nueva evaluación de ataluren con los nuevos datos disponibles del estudio 020.

El fabricante afirma que existe un efecto de tratamiento de ataluren a la dosis de 10/10/20 mg/kg, en base a un análisis por subgrupo con un 6MWD basal  $\geq 300$  y  $< 400$  metros, en el que se sugiere un efecto de  $\sim 50$  metros en el 6MWD por haber alcanzado significancia estadística, sin ningún control del error tipo I. También sugirió un efecto de tratamiento en este subgrupo en la mayoría de los desenlaces secundarios del estudio 020. En el estudio 020, el tiempo en descender cuatro escalones (un desenlace secundario) alcanzó significancia estadística nominal ( $p=0.012$ ). Los análisis post hoc que examinaron el cambio en la puntuación de "1" o "2" (capaz de realizar la función) a "0" (incapaz de realizar la función) sugirieron que menos pacientes en el grupo de ataluren perdieron la habilidad para realizar las funciones de NSAA. Sin embargo, todos estos resultados fueron obtenidos de manera secundaria, derivado de múltiples comparaciones y sin ajuste del control del error tipo I, o fueron análisis post hoc que tienen la desventaja de realizarse con los datos a la mano y sin ningún tipo de cegamiento, y con datos derivados de un estudio cuyo resultado del análisis primario en la población ITT no fue significativo. A continuación, se discuten cada uno de estos resultados en términos de sus limitaciones.



### Desenlace primario de eficacia

Ambos estudios pivotaes fallaron en mostrar un efecto del tratamiento con ataluren en el desenlace primario: cambio del 6MWD en la semana 48. El estudio 020 fue enriquecido para incluir a pacientes en la fase de declive ambulatoria, en base a los resultados de un subgrupo post hoc del estudio 007, y tuvo una muestra dos veces más grande que el estudio 007. Sin embargo, el desenlace primario también falló en mostrar significancia estadística; se observó un efecto no significativo de 13 metros en el 6MWD ( $p=0.21$ ). Estos resultados ilustran claramente la naturaleza engañosa de los análisis post hoc con datos a la mano y completamente abiertos, con el inconveniente de ser análisis exploratorios con datos de un estudio negativo.

El fabricante justificó este resultado negativo en esta población alegando que, a pesar de los esfuerzos de enriquecer el estudio con pacientes en la fase de declive ambulatoria, la población del estudio 020 permanecía siendo heterogénea, específicamente en la media del 6MWD basal. Ante esto, los evaluadores de la FDA realizaron un análisis en el subgrupo de pacientes con el promedio del 6MWD basal del subgrupo ADP del estudio 007 ( $< 433$  metros); los resultados de este análisis mostraron un ligero incremento del efecto, pero sin significancia estadística (15 metros,  $p=0.23$ ). Por lo tanto, los evaluadores del comité consultor de la FDA concluyeron que este resultado sugiere que la discrepancia entre los efectos del tratamiento entre ambos estudios no puede ser explicado por las diferencias en la media del 6MWD basal.

A partir de un resultado no significativo en el análisis de eficacia primario, debido a un orden jerárquico de análisis en el que incluye el ajuste del valor  $p$ , todos los resultados de análisis secundarios solo pueden ser considerados como exploratorios, como es el caso del estudio 020, que no especificó ningún orden de importancia de las múltiples comparaciones que realizó, y no hubo tampoco un plan del control estadístico del error tipo I.

Los desenlaces secundarios incluyeron pruebas motrices cronometradas, de las cuales solo una, el tiempo en descender cuatro escalones, fue significativo ( $p=0.01$ ) en la población ITT. Este resultado también debe ser considerado como generador de hipótesis en el contexto de un desenlace primario negativo, y no puede sustentar la efectividad de ataluren.

Otra limitante adicional de este desenlace es su falta de robustez. Los autores del estudio decidieron usar el tiempo máximo de 30 segundos si los pacientes que no podían realizar la prueba, o empleaban un mayor tiempo. Los resultados de los análisis de las pruebas funcionales cronometradas eran sensibles al manejo de máximo valor, y se perdió la significancia estadística de la prueba cuando se usó el mayor valor observado.



### Análisis por subgrupos

No se puede apoyar la eficacia de un agente en base a un análisis por subgrupo debido a múltiples limitaciones y sesgos en sus resultados. El plan de análisis estadístico del estudio 020 definió cinco subgrupos en base al 6MWD y en cuatro subgrupos basados en otras características basales. El fabricante no pre-especificó si uno de estos subgrupos sería incluido en el análisis primario con ajustes para controlar el error tipo I. El fabricante eligió resaltar el subgrupo con 6MWD basal de 300 a 400 metros, en el que la media del cambio en el 6MWD a las 48 semanas comparado con el basal fue 47.2 metros, favoreciendo a ataluren (n=47) comparado con placebo (n=52; p no ajustado=0.01). Aunque estos análisis por subgrupos fueron identificados en el manuscrito del estudio, la no disponibilidad del protocolo impide determinar si estos análisis fueron pre-especificados y si estas comparaciones forman parte de un número aún más grande de comparaciones. Esto es relevante tener en cuenta debido a que las múltiples comparaciones sin ajuste del valor p, y por consiguiente del error tipo I, aumenta la probabilidad de encontrar diferencias en algunas de ellas solo por azar. Tampoco se brindaron las razones para la elección de estos subgrupos ni se especificaron las direcciones de las comparaciones. Otro punto importante a considerar es que cuando se decide formar subgrupos en base a una característica, después de la aleatorización, estos nuevos subgrupos no estarán balanceados respecto a variables que se relacionan con el desenlace, porque no fueron aleatorizados. Estos nuevos grupos pueden mostrar diferencias estadísticamente significativas, que pueden aparecer debido a que los mismos pacientes son diferentes, y no por un efecto del tratamiento. Aun así, se obtuvieran valores de p significativos, estos análisis solo deben ser considerados únicamente como generadores de hipótesis para ser testeada en futuros ensayos clínicos.

Por otro lado, los evaluadores de la FDA notaron que los resultados en los subgrupos complementarios < 300 metros y  $\geq$  400 metros favorecieron numéricamente a placebo. Estas observaciones sugirieron que los resultados obtenidos en el subgrupo de 300 a 400 metros hayan sido solo por azar, o el resultado de excluir a pacientes con resultados negativos o no favorables para ataluren. Por último, los análisis por subgrupos, aunque tienen el potencial de generar hipótesis de futuros estudios, no proveen evidencia de la efectividad de ataluren en el contexto de un desenlace primario negativo.

### Evaluación exploratoria NSAA

La evaluación de la deambulacion con el NSAA fue incluida como una comparación exploratoria. Ninguno de los dos análisis que consideró el fabricante (usando la puntuación nominal y una transformación lineal del puntaje) en la población ITT resultó estadísticamente significativo, guardando consistencia con los resultados no significativos del análisis del desenlace primario y con la mayoría de los resultados no significativos de los desenlaces secundarios.



### Análisis post hoc

Muchos de los otros análisis post hoc fueron realizados usando una variedad de desenlaces y subgrupos, y algunos fueron positivos. Por ejemplo, el análisis del tiempo en descender cuatro escalones en el subgrupo de pacientes con 6MWD basal de 300-400 metros favoreció a ataluren con un valor de  $p < 0.001$ . Es bien reconocido que los análisis estadísticos post hoc no pre-especificados de múltiples subgrupos y con múltiples desenlaces resultaran en falsos hallazgos positivos o significativos y no pueden ser tomados como diferencias reales. Por lo tanto, los análisis post hoc no apoyan la efectividad de ataluren cuando los análisis pre-especificados del desenlace primario es negativo o no significativo.

Otro análisis post hoc que el fabricante condujo fue en el NSAA, considerando a los pacientes que habían declinado desde 2 a 1 o desde 1 a 0 durante el curso del estudio, con el que mostró que más pacientes del grupo placebo declinaron de 2 a 1 o de 1 a 0 en cada uno de los 17 ítems del cuestionario NSAA. Sin embargo, estos datos son el resultado de un análisis post hoc de un desenlace exploratorio que por sí mismo falló en demostrar, aunque sea una significancia estadística con dos métodos de análisis. Asimismo, los pacientes que declinaron de 2 a 1, constituyen un cambio clínico igualmente válido, pero no fueron presentados. En este análisis, considerando el paso de 2 a 1 punto, más pacientes del grupo de ataluren declinaron en 10 categorías, mientras que más pacientes del grupo placebo declinaron solamente en 2. Estos resultados resaltan las limitaciones inherentes en la presentación selectiva de los datos post hoc.

### Meta-análisis

Aunque algunos de los meta-análisis mostraron alguna significancia a favor de ataluren, sin embargo, ellos no apoyan la efectividad de ataluren debido a las siguientes razones:

- Estos meta-análisis no fueron pre-especificados ni planeados, y fueron hechos después de que se supieran los resultados del estudio, y solo pueden ser considerados como generadores de hipótesis.
- Los métodos de análisis fueron diferentes de los análisis pre-especificados en el estudio 007.
- La omisión del brazo de alta dosis del estudio 007 es un ajuste post hoc adicional.

Por tanto, es inapropiado considerar los resultados de un ensayo negativo y combinarlos con los resultados post hoc de otro ensayo también negativo (estudio 007) con el objetivo de generar un análisis conjunto y con valores de  $p$  nominales, es decir sin control del error tipo I, por lo que no pueden ser considerados como evidencia para afirmar un efecto significativo del tratamiento con ataluren.

La Comisión de transparencia del sistema de salud de Francia evaluó la información proveniente de estudio 020. Los evaluadores encontraron muchas limitaciones metodológicas y de difícil interpretación en este estudio, por lo que concluyeron que no

aportaba ninguna evidencia respecto a la eficacia de ataluren respecto a placebo. En este estudio el resultado del análisis primario de eficacia, no pone en evidencia diferencias entre placebo y ataluren sobre la base del declive de la distancia caminada según el 6MWT. Todos los resultados de los desenlaces secundarios y por subgrupos fueron considerados exploratorios debido a que el análisis estadístico era jerárquico, es decir su validez dependía de la significancia en el desenlace primario de eficacia. El fabricante proporcionó datos de análisis en subgrupos definidos a posteriori según el 6MWT, estos análisis fueron incorporados al plan de análisis estadístico en 2015. Por lo que, estos análisis solo pueden considerarse como información exploratoria que podría sugerir eficacia heterogénea según el 6MWT basal. Consideraron preocupante el uso de un plan de análisis estadístico que fue finalizado después de levantar el cegamiento, y sin explicación convincente de ello por qué.

A pesar de las limitaciones metodológicas de los análisis de subgrupos y post hoc, se decidió renovar la autorización de comercialización, pero condicionada a la realización de un ensayo más grande y de mayor duración (18 meses) y con seguimiento posterior de 18 meses adicionales. Además, en acorde al bajo nivel de confianza de los resultados, la comisión opinó favorablemente a la inscripción en la lista de medicamentos reembolsables a los asegurados, pero con una tasa de reembolso del 15%.

Por otro lado, aunque al comité evaluador de NICE le preocupó que el beneficio de ataluren no había sido mostrado en la población para la cual estaba dirigida en la práctica clínica, el comité consideró que, en base a la evidencia actualmente disponible, los potenciales beneficios asociados con ataluren justificarían su alto costo cuando se aplicaran importantes descuentos incluidos en el esquema de acceso al paciente y componentes financieros confidenciales del acuerdo de acceso gestionado. Por ello, decidió aprobar el uso de ataluren solamente si se aplicaban los descuentos acordados confidencialmente entre el fabricante del medicamento y el NHS al precio de ataluren.

Ambas evaluaciones resaltaron las limitaciones metodológicas del estudio 020, y aprobaron su uso de manera condicional, reflejo de la falta de robustez de los resultados que sugieran un beneficio clínico con ataluren. Así, la evaluación de la comisión de transparencia solo ha propuesto un reembolso del 15 % y ha comprometido al fabricante a la realización de otro estudio aleatorizado, doble ciego, controlado con placebo, de mayor duración (72 semanas) y un periodo de seguimiento posterior de 72 semanas adicionales. Asimismo, la aprobación de NICE claramente se basó en acuerdos que incluyen descuentos importantes al costo del medicamento para equilibrar la relación calidad-precio de esta tecnología.

En la evaluación metodológica que hizo la FDA, ya se había resaltado la necesidad de realizar otro estudio bien diseñado. Se señaló que la acumulación de varios análisis adicionales y exploratorios solo pueden ser considerados como generadores de hipótesis tal como están presentadas, y que estos hallazgos pueden posiblemente proveer información de apoyo acompañante de los resultados de un único estudio adecuado y



bien controlado. De presentarse tal evidencia a la FDA podría tener el potencial de proveer evidencia substantiva de efectividad. En la carta de revisión la FDA expresó su disposición de trabajar cercanamente con PTC, el fabricante de ataluren en desarrollar un protocolo de un ensayo en base a las hipótesis generadas de los estudios 007 y 020.

El fabricante afirmó que la efectividad de ataluren había sido establecida en base a análisis post hoc en la población ADP del estudio 007. Sin embargo, cuando esta conclusión fue evaluada prospectivamente en el estudio 020, los resultados fueron claramente negativos. Esta situación resalta la naturaleza frecuentemente engañosa de los análisis exploratorios de los ensayos negativos. El fabricante sostiene la efectividad de ataluren solamente con los resultados de numerosos análisis sin control del error tipo I, post hoc y exploratorios que tienen el fin de mitigar dos ensayos negativos.



## VI. CONCLUSIONES

- El presente dictamen tiene como objetivo reevaluar la eficacia y seguridad de ataluren 40 mg en el tratamiento de pacientes DMD con mutación sin sentido y en fase de declive ambulatorio, tras de la publicación de los resultados del estudio 020. Asimismo, se incluyeron dos ETS, realizadas por NICE (NICE 2017) y por el comité de transparencia de la Alta Autoridad de Salud de Francia (Haute Autorité de Santé 2017), y una evaluación metodológica de la FDA.
- Los estudios 007 y 020 claramente fallaron en el análisis primario de la *distancia caminada en 6 minutos (6MWD)* en todos los pacientes. A pesar del esfuerzo de enriquecer al estudio 020 basado en los resultados del estudio 007, con una muestra dos veces más grande, el análisis primario de eficacia no mostró diferencias entre ataluren y placebo.
- Los análisis de un subgrupo de pacientes según el valor basal del 6MWD 300-400 metros del estudio 020 mostraron una diferencia estadísticamente significativa en el desenlace primario, y en los desenlaces secundarios en este subgrupo que podrían sugerir un efecto favorable de ataluren. Sin embargo, el mirar dentro de un estudio en subgrupos no pre-especificados incrementa el error tipo I substancialmente; el problema es aún mayor cuando existen muchos subgrupos, como en este caso. Los revisores de la FDA consideraron que el subgrupo con 6MWD 300-400 metros no había sido pre-especificado, ya que no fue identificado claramente en el análisis de efectividad planeado con un control planeado del error tipo I, en el plan de análisis estadístico.
- Aunque resulten tentadores, los análisis por subgrupos conllevan múltiples limitaciones. El más importante a ser considerado es que cualquier subgrupo que se forme no produciría grupos comparables por la falta de aleatorización, de lo cual se desprende que cualquier diferencia que se encuentre puede deberse a varios factores (no balanceados) aparte de la intervención. Los análisis por subgrupos solo pueden ser considerados como generadores de hipótesis, y para ello deben ser medianamente creíbles. Esto significa que, estos resultados deben provenir de un número limitado y pre-especificado de comparaciones, para los cuales los autores han brindado las explicaciones de su elección y han especificado a priori las direcciones de los resultados. Adicionalmente, deben contar con apoyo de otros hallazgos similares y existir una plausibilidad biológica.
- En el caso del estudio 020, no se encuentra disponible el protocolo que permita determinar cuántas comparaciones se pre-especificaron y cuáles fueron las razones de su elección. De nueve comparaciones que reportaron los autores, solo uno resultó siendo significativa, y con el agravante de que no fue ajustado por



múltiples comparaciones, incrementando el error tipo I. No sabemos si estas nueve comparaciones forman parte de un número mayor de comparaciones, lo cual incrementa aún más la posibilidad de encontrar alguna diferencia significativa solo por azar. Una observación que hicieron los revisores de la FDA fue la dirección opuesta, a favor de placebo en los subgrupos complementarios (6MWD < 300 o  $\geq$  400 metros) en el estudio 020, sugiriendo que el efecto de tratamiento observado en el subgrupo con 6MWD basal de 300-400 metros haya podido ser el resultado de la exclusión de los subgrupos de pacientes con resultados negativos.

- Los resultados en los desenlaces secundarios (funciones motoras cronometradas) que sugirieron ser favorables con ataluren en el subgrupo con 6MWD basal de 300-400 metros provienen de un análisis post hoc. Es bien reconocido que los análisis estadísticos post hoc no pre especificados de múltiples subgrupos y con múltiples desenlaces resultarán en falsos hallazgos positivos o significativos y no pueden ser tomados como diferencias reales. Por lo tanto, los análisis post-hoc no apoyan la efectividad de ataluren cuando los análisis pre-especificados del desenlace primario es negativo o no significativo.
- Las recomendaciones a favor del uso de ataluren en las ETS consultadas, exponen claramente las limitaciones de la evidencia disponible, condicionando su aprobación en base a una disminución importante en su inversión. Así, la evaluación de NICE ha condicionado su aprobación a la aplicación de importantes descuentos confidenciales al precio del agente, mientras que la comisión de transparencia de Francia ha propuesto el reembolso solo del 15 % y ha comprometido al fabricante a la realización de un ensayo más grande y de mayor duración.
- La evaluación previa que hace la FDA, antes de aprobar el registro de un medicamento para ser evaluado formalmente, ya había expresado las limitaciones de la evidencia presentada e instó al fabricante a realizar otro ensayo, para el cual manifestó que prestaría todo el apoyo metodológico necesario. Este ensayo ha sido registrado recientemente en junio del 2017. Se planea reclutar a 250 pacientes para evaluar los efectos a mayor plazo (72 semanas) seguido de una fase abierta de 72 semanas adicionales y controlado con placebo.
- Cuando los planes del análisis son desarrollados con los datos en las manos, sin un plan prospectivo para controlar el error tipo I, el estudio carece entonces de rigor estadístico necesario para que el estudio sea considerado como adecuadamente controlado. En el presente caso, el plan estadístico claramente definió el análisis primario del estudio 020, en una población específicamente seleccionada como una población enriquecida según el estudio 007. Los estudios



DICTAMEN PRELIMINAR DE EVALUACIÓN DE TECNOLOGÍA SANITARIA N.º 056-SDEPFYOTS-DETS-IETS-2017  
EFICACIA DE ATALUREN EN PACIENTES CON DIAGNÓSTICO DE DISTROFIA MUSCULAR DE DUCHENNE PORTADORES DE UNA  
MUTACIÓN SIN SENTIDO EN EL GEN DE DISTROFINA - ACTUALIZACIÓN

007 y 020, analizados de la manera prospectiva planeada, claramente no proveen ninguna base de evidencia importante de la efectividad de ataluren

- Por lo expuesto, el Instituto de Evaluación de Tecnologías en Salud e Investigación- IETSI ratifica lo estipulado en el Dictamen Preliminar de Evaluación de Tecnología Sanitaria N.º 020-SDEPFyOTS-DETS-IETSI-2016, donde no se aprueba el uso de ataluren en el tratamiento de pacientes con DMD con mutación sin sentido en el gen de distrofina.



## VII. RECOMENDACIONES

A la fecha no existe evidencia sobre la eficacia de ataluren comparado con el tratamiento de soporte. Se insta a los médicos especialistas sugerir nuevas tecnologías para esta enfermedad.



## VIII. REFERENCIAS BIBLIOGRÁFICAS

FDA Briefing Document. Peripheral and Central Nervous System Drugs. Advisory Committee Meeting. September 28, 2017. Disponible en: <https://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/PeripheralandCentralNervousSystemDrugsAdvisoryCommittee/ucm577348.htm>.

FDA. Manual of policies and procedures. Center for drug evaluation and research. Policy and procedures. Good Review Practice: Refuse to File. Disponible en: <https://www.fda.gov/downloads/aboutfda/centersoffices/officeofmedicalproductsandtobacco/cder/manualofpoliciesprocedures/ucm370948.pdf>.

McDonald CM, Campbell C, Torricelli RE, Finkel RS, Flanigan KM, Goemans N et al; Clinical Evaluator Training Group; ACT DMD Study Group. Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. Lancet 2017 Sep 23;390(10101):1489-1498. (Estudio 020).

Haure Autorité de Santé. Commission de la Transparence - Opinión. Publicado el 5 de julio del 2017. Disponible en [https://www.has-sante.fr/portail/jcms/c\\_2009663/en/translama-ataluren-ma-interference-agent](https://www.has-sante.fr/portail/jcms/c_2009663/en/translama-ataluren-ma-interference-agent).

Bushby K, Finkel R, Wong B, Barohn R, Campbell C, Comi GP, et al; PTC124-GD-007-DMD STUDY GROUP. Ataluren treatment of patients with nonsense mutation dystrophinopathy. Muscle Nerve. 2014 Oct;50(4):477-87. (Estudio 007).

European Medicines Agency. EMA. Summary of the risk management plan (RMP) for Translarna (ataluren). Disponible en [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/EPAR\\_-\\_Risk\\_management-plan\\_summary/human/002720/WC500167551.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Risk_management-plan_summary/human/002720/WC500167551.pdf).

NICE 2016- "Ataluren for treating Duchenne muscular dystrophy with a nonsense mutation in the dystrophin gene". Highly specialised technologies Published: 20 July 2016. nice.org.uk/guidance/hst3.

Translarna. London: European Medicines Agency;2014.



## IX. ANEXOS

Tabla N.º 1. Resultados de eficacia en los principales puntos finales del estudio 020 . Población ITT

Desenlaces	Ataluren (n=115)	Placebo (n=115)
<b>Primario</b>		
Cambio en la distancia recorrida en 6 minutos. Medias (EE)		
<ul style="list-style-type: none"> <li>• Basal</li> <li>• Cambio intra-grupo a las 48 semanas</li> <li>• Promedio de diferencias vs placebo</li> <li>• Valor p</li> </ul>	364.0 (73.3) -42.2 (9.2) 15.4 0.21	362.7 (81.4) -57.6 (9.3)
<b>Secundarios (exploratorios)</b>		
Tiempo de ocurrencia del deterioro persistente del 10% de la distancia caminada (días) Mediana (IC 95%)	NA [261;NA]	NA [224;NA]
<b>Pruebas funcionales cronometradas:</b>		
Tiempo requerido para recorrer 10 metros caminando o corriendo (segundos) promedio (EE)		
<ul style="list-style-type: none"> <li>• Basal</li> <li>• Alargamiento promedio intra-grupo a las 48 semanas</li> </ul>	6.6 (3.1) 2.3 (5.2)	6.8 (2.9) 3.5 (6.4)
Tiempo requerido para subir 4 escalones (segundos)	5.9 (5.1)	6.4 (5.7)
<ul style="list-style-type: none"> <li>• Basal</li> <li>• Alargamiento promedio intra-grupo a las 48 semanas</li> </ul>	2.7 (5.3) 5.0 (5.0)	4.5 (7.3) 4.8 (4.5)
Tiempo requerido para bajar 4 escalones (segundos)	2.2 (5.3)	4.0 (7.9)
<ul style="list-style-type: none"> <li>• Basal</li> <li>• Alargamiento promedio intra-grupo a las 48 semanas</li> </ul>		
Puntuación total NSAA, media (EE)		
<ul style="list-style-type: none"> <li>• Basal</li> <li>• Cambio intra-grupo a las 48 semanas</li> </ul>	22.2 (7.8) -2,7 (4,0)	21.9 (8.0) -3,7 (4,3)

