# SUBVENCIONES PARA LA FINANCIACION DE LA I+D+i BIOMÉDICA Y EN CIENCIAS DE LA SALUD EN ANDALUCÍA

### MEMORIA CIENTÍFICO-TÉCNICA Y ECONÓMICA: PROYECTOS DE INNOVACIÓN

Orden de 19 de mayo de 2016 (BOJA nº 98 de 25 de mayo 2016, pp. 9-26. Disponible en <a href="http://juntadeandalucia.es/boja/2016/98/BOJA16-098-00051-9272-01\_00091809.pdf">http://juntadeandalucia.es/boja/2016/98/BOJA16-098-00051-9272-01\_00091809.pdf</a>)

LÍDER-JEFE DEL PROYECTO	
APELLIDOS	NOMBRE
Antón Gamero	Montserrat

### TÍTULO DEL PROYECTO

Creación de un Sistema de Apoyo a la Toma de Decisiones Clínicas para el manejo de la Enfermedad Renal Crónica no Terminal en niños basado en la aplicación de Minería de Datos sobre el Registro Español Pediátrico de Enfermedad renal crónica REPIR II.

### RESUMEN

(Máximo 250 palabras)

La Enfermedad Renal Crónica (ERC) supone un importante problema de salud, no sólo por su elevada morbimortalidad, sino por el elevado gasto sanitario que supone. Afecta al 10% de la población adulta, siendo los datos epidemiológicos en la infancia muy escasos. Se estima una incidencia y prevalencia europea de 11 y 70 casos respectivamente por millón de población pediátrica. La evolución natural de la enfermedad es progresiva hasta la pérdida total de la función renal, lo que hace necesario el tratamiento sustitutivo (diálisis o trasplante) en los estadios finales. La detección precoz y una adecuada intervención sobre los factores de progresión del daño renal pueden mejorar el pronóstico y ralentizar la evolución de la enfermedad. Actualmente, el control del coste económico del tratamiento de la ERC es un reto para el sistema sanitario.

Sin embargo, los estudios pediátricos que identifiquen estos factores y su relación con la progresión de la enfermedad son difíciles de desarrollar y no existen modelos predictivos validados para su aplicación clínica. Los existentes son extrapolaciones de modelos de la población adulta que no tienen en cuenta las características diferenciadoras como la diferente etiología del daño renal o su relación con el crecimiento.

El objetivo de este proyecto es desarrollar un conjunto de modelos pronósticos utilizando técnicas de Minería de Datos específicos para la edad pediátrica, que permitan establecer de forma temprana la mejor estrategia de intervención, minimizando la progresión y las comorbilidades asociadas, reduciendo con ello el gasto sanitario. Estos modelos son una base de conocimiento para un Sistema de Asistencia a la Toma de Decisiones Clínicas, con claro carácter traslacional.

Para este desarrollo se cuenta con la colaboración del REPIR II, un registro de niños con ERC de la Asociación Española de Nefrología Pediátrica de ámbito nacional, con casi 70 centros repartidos en 36 provincias.

Aunque hay mucho interés en la literatura por el desarrollo de estos modelos, el que se propone es original debido tanto al número de pacientes en el que se basa, mucho mayor que en otros estudios, como a su flexibilidad, más avanzados que los basados en estadística clásica.

Su objetivo no es meramente académico si no que busca el impacto en la práctica clínica. Se creará un software fácil de usar basado en los modelos pronósticos, que permita al clínico establecer la mejor línea de actuación para cada paciente, en función de los riesgos que tenga.

ADECUACIÓN E INTERÉS DE LA PROPUESTA CON ALGUNA DE LAS LÍNEAS ESTRATÉGICAS DE LA CONSEJERÍA DE SALUD
Marque y justifique la adecuación e interés de la propuesta con las siguientes líneas estratégicas:
☐ Estudios de intervención
El objetivo final del proyecto es la creación de un software de apoyo a la toma de decisiones clínicas en el manejo de niños con ERC. Esta herramienta estará basada en modelos validados en generalización mediante técnicas estadísticas. Su aplicación pretende disminuir la variabilidad de la práctica clínica y abrirá la puerta a un estudio de intervención posterior que valide su uso de forma general en poblaciones similares.
☐ Nuevos servicios integrados de atención basados en la aplicación de nuevas tecnologías, especialmente digitales (salud digital, e-salud)
□ Nuevas soluciones para la atención a pacientes crónicos y pluripatológicos
Los efectos de la ERC en el niño son devastadores. A medida que progresa el daño renal aumentan las comorbilidades asociadas como el retraso del crecimiento, la malnutrición, anemia, enfermedad mineral ósea e incremento del riesgo cardiovascular lo que supone un impacto significativo en la calidad de vida relacionada con la salud, alteraciones psicosociales y familiares con elevado consumo de los recursos sanitarios. El manejo de estos pacientes es complicado y requiere un abordaje multidisciplinario para conseguir no solo prolongar su vida en las mejores condiciones físicas sino el verdadero objetivo final de la integración del niño con ERC en la sociedad hasta alcanzar la edad adulta. Los modelos propuestos en este estudio permitirán identificar aquellos niños con mayor riesgo de progresión a ERC terminal que se pueden beneficiar de intervenciones específicas —de diagnóstico y tratamiento- en salud mediante la aplicación del Sistema de Apoyo a la Toma de Decisiones Clínicas. Esta herramienta es novedosa para la atención de niños con ERC y múltiples patologías asociadas permitiendo la aplicación de los resultados de la investigación en la práctica clínica habitual.
□ Terapias Avanzadas y medicina personalizada
Los modelos de predicción de la evolución del daño renal crónico en niños basados en datos reales permitirán ajustar las intervenciones en salud individualmente, con una mejora significativa de la atención clínica a estos pacientes sin la incorporación de nuevos servicios para los ciudadanos. Los esfuerzos dirigidos a mejorar los modelos actuales de la enfermedad permitirán tratamientos más dirigidos al control de los factores de riesgo asociados a la progresión del daño renal. A su vez, el ajuste de la frecuencia y características de las revisiones clínicas y analíticas a los realmente requeridos por cada paciente permitirá reducir el coste sanitario de forma significativa.
☑ Nuevas herramientas de gestión del conocimiento aplicadas a la decisión clínica y de salud poblacional
El objetivo final del desarrollo de los nuevos modelos predictivos es la construcción de un Sistema Experto de Asistencia a la Toma de Decisiones Clínicas para su uso en la práctica clínica diaria, por lo que el proyecto se alinea perfectamente con esta línea estratégica. Esta herramienta se integrará además con el registro REPIR II, facilitando la identificación de niños con mayor riesgo de progresión de daño renal, lo que puede contribuir no sólo a mejorar la utilidad del mismo sino la mejora de la salud global de la población registrada.
□ Soluciones de soporte al fomento de vida saludable y envejecimiento activo

## 1. LÍDER JEFE DEL PROYECTO Y EQUIPO DE INNOVACIÓN

1.1	CV DEL/ DEL LÍDER DEL PROYECTO y DEL EQUIPO DE INNOVACIÓN (incluidas las personas y empresas que colaboren en el desarrollo del proyecto mediante su participación activa e implicación en el diseño del mismo)
Se adjur	ntan los CVs en la aplicación informática Gestión de convocatorias

### 1.2 BREVE RESUMEN DEL EQUIPO DE INNOVACIÓN DE LOS ÚLTIMOS 5 AÑOS

Nombre y apellidos	Especialidad	Tipo colaboración en el equipo de innovación (IP o IC)	Nº años y ACTIVIDADES indexadas e I+D+i JCR del ISI		Nº art. revistas indexadas en otros índices	Nº Patentes	Nº proyectos de innovación y financiación obtenidos
Montserrat Antón Gamero	Pediatría	IP	17	4			
Marta Melgosa Hijosa	Pediatría	IC	10	16			
Víctor Manuel García Nieto	Pediatría	IC	32	24			

ANTECEDENTES DEL INVESTIGADOR/A PRINCIPAL Y DEL EQUIPO DE INNOVACIÓN EN LA APLICACIÓN DE RESULTADOS DE PROYECTOS ANTERIORES

### Lider-jefe de Proyecto

### Montserrat Antón Gamero

Facultativo especialista de área de Pediatría con dedicación desde hace 18 años a la Nefrología pediátrica en el Hospital Universitario Reina Sofía de Córdoba. Integrada en un equipo de investigación del Instituto Maimónides de Investigación Biomedica de Córdoba (IMIBIC). Profesora asociada de la Universidad de Córdoba. Secretaria de la Asociación Española de Nefrología Pediátrica (AENP) y coordinadora del REPIR II. Miembro del grupo de trabajo de ERC de la European Society of Pediatric Nephrology. Autora de capítulos de libros y publicaciones e investigador colaborador en proyectos de investigación sobre enfermedad renal en niños.

### Equipo innovación

### Marta Melgosa Hijosa

Facultativo especialista de área de Pediatría con dedicación desde hace 20 años a la Nefrología pediátrica en el Hospital Universitario La Paz de Madrid. Integrada en un equipo de investigación del IdiPaz. Coordinadora del REPIR II. Autora de capítulos de libros y publicaciones e investigador colaborador en proyectos de investigación sobre enfermedad renal en niños.

### Víctor Manuel García Nieto

Presidente de la Asociación Española de Nefrología Pediátrica. Pediatra y Nefrólogo con 42 años de dedicación en el Hospital Universitario Nuestra Señora de la Candelaria de Tenerife. Autor de cientos de publicaciones de impacto y editor de libros sobre enfermedades renales en niños.

### 2. ASPECTOS CIENTÍFICO-TÉCNOLÓGICOS DEL PROYECTO

### 2.1 BIBLIOGRAFÍA

Se valorará que la bibliografía sea pertinente al tema propuesto y que esté actualizada conteniendo las últimas publicaciones en relación al tema propuesto. También se valorará la citación de las referencias bibliográficas a lo largo del proyecto. (Máximo 2 páginas)

[Alcázar et al., 2008] Alcázar, J., Arenas, M., Álvarez-Ude, F., Virto, R., Rubio, E., Maduell, F., Fernández-Crespo, P., Angoso de Guzmán Delgado, R., Santamaría, C., et al. (2008). Resultados del proyecto de mejora de la calidad de la asistencia en hemodiálisis: estudio multic de indicadores de calidad de la sociedad española de nefrología (SEN). Nefrología, 28(6).

[Angoso et al., 2006] Angoso, M., Alcalde, G., Álvarez-Ude, F., and Arenas, M. (2006). Gestión de calidad en hemodiálisis. Nefrología, 26(8):73–87.

[Areses Trapote et al., 2010] Areses Trapote, R., Sanahuja Ibáñez, M., and Navarro, M. (2010). Epidemiología de la enfermedad renal cró terminal en la población pediátrica española: Proyecto repir ii. Nefrología (Madrid), 30(5):508–517.

[Bajo Rubio et al., 2010] Bajo Rubio, M. A., Selgas, R., Remón, C., Arrieta, J., Álvarez-Ude, F., Arenas, M., Borrás, M., Coronel, F., García R., Minguela, I., et al. (2010). Plan de calidad científico-técnica y de mejora continua de calidad en diálisis peritoneal. Nefrología (Madrid), 30(1):28–45

[Barbieri et al., 2015] Barbieri, C., Mari, F., Stopper, A., Gatti, E., Escandell-Montero, P., Martínez-Martínez, J. M., and Martín-Guerrero, J. (2015). A new machine learning approach for predicting the response to anemia treatment in a large cohort of end stage renal disease pat undergoing dialysis. Computers in biology and medicine, 61:56–61.

[Bolón-Canedo et al., 2016] Bolón-Canedo, V., Sánchez-Maroño, N., and Alonso-Betanzos, A. (2016). Feature selection for high-dimension Progress in Artificial Intelligence, 5(2):65–75.

[Cerqueira et al., 2014] Cerqueira, D. C., Soares, C. M., Silva, V. R., Magalhães, J. O., Barcelos, I. P., Duarte, M. G., Pinheiro, S. V., Colos A., Silva, A. C. S., and Oliveira, E. A. (2014). A predictive model of progression of ckd to esrd in a predialysis pediatric interdisciplinary pro Clinical Journal of the American Society of Nephrology, 9(4):728–735.

[SGC/CCAA, 2015] COHESIÓN, SUBDIRECCIÓN GENERAL DE CALIDAD Y DE LAS CCAA, C. D. S. (2015). Documento marco sobre Enfermedad Renal Crónica (ERC) dentro de la estrategia de abordaje a la cronicidad en el SNS.

[Collins et al., 2015] Collins, G. S., Reitsma, J. B., Altman, D. G., and Moons, K. G. (2015). Transparent reporting of a multivariable predict for individual prognosis or diagnosis (tripod): the tripod statement. BMC medicine, 13(1):1.

[Copelovitch et al., 2011] Copelovitch, L., Warady, B. A., and Furth, S. L. (2011). Insights from the chronic kidney disease in children (ckid) Clinical Journal of the American Society of Nephrology, 6(8):2047–2053.

[Czock et al., 2015] Czock, D., Konias, M., Seidling, H. M., Kaltschmidt, J., Schwenger, V., Zeier, M., and Haefeli, W. E. (2015). Tailoring o substantially reduces the alert burden in computerized clinical decision support for drugs that should be avoided in patients with renal diseasournal of the American Medical Informatics Association, page ocv027.

[Drawz et al., 2015] Drawz, P. E., Archdeacon, P., McDonald, C. J., Powe, N. R., Smith, K. A., Norton, J., Williams, D. E., Patel, U. D., and A. (2015). Ckd as a model for improving chronic disease care through electronic health records. Clinical Journal of the American Society o Nephrology, 10(8):1488–1499.

[Flaquer et al., 2010] Flaquer, M., Romagnani, P., and Cruzado, J. (2010). Factores de crecimiento y regeneración renal. Nefrología (Madr 30(4):385–393.

[Floege et al., 2015] Floege, J., Mak, R. H., Molitoris, B. A., Remuzzi, G., and Ronco, P. (2015). Nephrology research [mdash] the past, prefuture. Nature Reviews Nephrology.

[García et al., 2015] García, S., Luengo, J., and Herrera, F. (2015). Data preprocessing in data mining. Springer.

[Ginsberg et al., 2014] Ginsberg, J. S., Zhan, M., Diamantidis, C. J., Woods, C., Chen, J., and Fink, J. C. (2014). Patient-reported and active safety events in ckd. Journal of the American Society of Nephrology, pages ASN–2013090921.

[Harambat et al., 2012] Harambat, J., Van Stralen, K. J., Kim, J. J., and Tizard, E. J. (2012). Epidemiology of chronic kidney disease in chil Pediatric nephrology, 27(3):363–373.

[Helldén et al., 2015] Helldén, A., Al-Aieshy, F., Bastholm-Rahmner, P., Bergman, U., Gustafsson, L. L., Höök, H., Sjöviker, S., Söderströn Odar-Cederlöf, I. (2015). Development of a computerised decisions support system for renal risk drugs targeting primary healthcare. BMJ 5(7):e006775.

[Hogg et al., 2003] Hogg, R. J., Furth, S., Lemley, K. V., Portman, R., Schwartz, G. J., Coresh, J., Balk, E., Lau, J., Levin, A., Kausz, A. T., (2003). National kidney foundations kidney disease outcomes quality initiative clinical practice guidelines for chronic kidney disease in child adolescents: evaluation, classification, and stratification. Pediatrics, 111(6):1416–1421.

[Ishikura et al., 2014] Ishikura, K., Uemura, O., Hamasaki, Y., Ito, S., Wada, N., Hattori, M., Ohashi, Y., Tanaka, R., Nakanishi, K., Kaneko (2014). Progression to end-stage kidney disease in japanese children with chronic kidney disease: results of a nationwide prospective coh Nephrology Dialysis Transplantation, page gfu012.

[Kaspar et al., 2016] Kaspar, C., Bholah, R., and Bunchman, T. (2016). A review of pediatric chronic kidney disease. Blood purification, 41: 3):211–217.

[Lakshmi et al., 2014] Lakshmi, K., Nagesh, Y., and Krishna, M. V. (2014). Performance comparison of three data mining techniques for probability. International Journal of Advances in Engineering & Technology, 7(1):242.

[Leung et al., 2013] Leung, R. K., Wang, Y., Ma, R. C., Luk, A. O., Lam, V., Ng, M., So, W. Y., Tsui, S. K., and Chan, J. C. (2013). Using a staged strategy based on machine learning and mathematical modeling to predict genotype-phenotype risk patterns in diabetic kidney dise prospective case—control cohort analysis. BMC nephrology, 14(1):1.

[Levey et al., 2005] Levey, A. S., Eckardt, K.-U., Tsukamoto, Y., Levin, A., Coresh, J., Rossert, J., Zeeuw, D. D., Hostetter, T. H., Lameire, N., and Eknoyan, G. (2005). Definition and classification of chronic kidney disease: a position statement from kidney disease: Improving global outcomes (kdigo). Kidney international, 67(6):2089-2100. [Levey et al., 2014] Levey, A. S., Inker, L. A., Matsushita, K., Greene, T., Willis, K., Lewis, E., De Zeeuw, D., Cheung, A. K., and Coresh, J. Gfr decline as an end point for clinical trials in ckd: a scientific workshop sponsored by the national kidney foundation and the US Food and Drug administration. American Journal of Kidney Diseases, 64(6):821-835. [López et al., 2012] López, V., Fernández, A., Moreno-Torres, J. G., and Herrera, F. (2012). Analysis of preprocessing vs. cost-sensitive learning for imbalanced classification. Open problems on intrinsic data characteristics. Expert Systems with Applications, 39(7):6585–6608. [Luo et al., 2013] Luo, L., Small, D., Stewart, W. F., and Roy, J. A. (2013). Methods for estimating kidney disease stage transition probabilities using electronic medical records. eGEMs (Generating Evidence & Methods to improve patient outcomes), 1(3):6. [Mak, 2007] Mak, R. H. (2007). Chronic kidney disease in children: state of the art. Pediatric Nephrology, 22(10):1687–1688. [Martín Hernández, 2011] Martín Hernández, R. (2011). Conocer y controlar los costes del tratamiento de la insuficiencia renal crónica: Una necesidad inaplazable. Nefrología (Madrid), 31(3):256-259. [Melgosa-Hijosa and Anton-Gamero, 2014] Melgosa-Hijosa, M. and Anton-Gamero, M. (2014). Epidemiological and clinical data from spanish pediatric pre-dialysis chronic kidney disease registry (repir ii). In PEDIATRIC NEPHROLOGY, volume 29, pages 1684–1684. SPRINGER 233 SPRING ST, NEW YORK, NY 10013 USA. [Mendonça et al., 2015] Mendonça, A. C. Q., Oliveira, E. A., Fróes, B. P., Faria, L. D. C., Pinto, J. S., Nogueira, M. M. I., Lima, G. O., Resende, P. L, Assis, N. S., e Silva, A. C. S., et al. (2015). A predictive model of progressive chronic kidney disease in idiopathic nephrotic syndrome. F Nephrology, 30(11):2011-2020. [Moons et al., 2015] Moons, K. G., Altman, D. G., Reitsma, J. B., Ioannidis, J. P., Macaskill, P., Steyerberg, E. W., Vickers, A. J., Ransohoff, D. F., and Collins, G. S. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod): explanation elaboration. Annals of internal medicine, 162(1):W1-W73. [Musso et al., 2013] Musso, C., Aguilera, J., Otero, C., Vilas, M., Luna, D., and de Quirós, F. G. B. (2013). Informatic nephrology. International urology and nephrology, 45(4):1033-1038 [Musso et al., 2015] Musso, C. G., Plazzotta, F., Otero, C., Aguilera, J., Campos, F., Diez, G. R., Luna, D., and de Quirós, F. G. B. (2015) Informatic nephrology: 17 years of one-center experience. International urology and nephrology, 47(9):1587–1588. [Pavlou et al., 2015] Pavlou, M., Ambler, G., Seaman, S. R., Guttmann, O., Elliott, P., King, M., and Omar, R. Z. (2015). How to develop a more [Plotnicki et al., 2012] Plotnicki, L., Höcker, B., Krupka, K., Kohl, C., Rahmel, A., Pape, L., Hoyer, P., Marks, S., Webb, N., Söylemezoglu, 🗘., et al. (2012). The certain registry: A novel, web-based registry and research platform for paediatric renal transplantation in europe: 2185. Transp 94(10S):1210. [Raghunathan, 2015] Raghunathan, T. (2015). Missing Data Analysis in Practice. CRC Press. [Samuel et al., 2010] Samuel, S. M., Tonelli, M. A., Foster, B. J., Nettel Aguirre, A., Na, Y., Williams, R., Soo, A., and Hemmelgarn, B. R. (2010). Overview of the canadian pediatric end-stage renal disease database. BMC nephrology, 11(1):1. [Schwartz et al., 2012] Schwartz, G. J., Schneider, M. F., Maier, P. S., Moxey-Mims, M., Dharnidharka, V. R., Warady, B. A., Furth, S. L., a Muñoz, A. (2012). Improved equations estimating gfr in children with chronic kidney disease using an immunonephelometric determination cystatin c. Kidney international, 82(4):445-453. [Singh et al., 2015] Singh, A., Nadkarni, G., Gottesman, O., Ellis, S. B., Bottinger, E. P., and Guttag, J. V. (2015). Incorporating temporal ellr data in predictive models for risk stratification of renal function deterioration. Journal of biomedical informatics, 53:220–228. [Siragy, 2010] Siragy, H. M. (2010). Escape: From hypertension to renal failure. Current hypertension reports, 12(4):207-209. [Staples and Wong, 2010] Staples, A. and Wong, C. (2010). Risk factors for progression of chronic kidney disease. Current opinion in pediatrics, 22(2):161. [Staples et al., 2010] Staples, A. O., Greenbaum, L. A., Smith, J. M., Gipson, D. S., Filler, G., Warady, B. A., Martz, K., and Wong, C. S. (2010). Association between clinical risk factors and progression of chronic kidney disease in children. Clinical Journal of the American Society of Nephrology, 5(12):2172-2179. [Staples et al., 2009] Staples, A. O., Wong, C. S., Smith, J. M., Gipson, D. S., Filler, G., Warady, B. A., Martz, K., and Greenbaum, L. A. (2009). Anemia and risk of hospitalization in pediatric chronic kidney disease. Clinical Journal of the American Society of Nephrology, 4(1):48-56 [Venkat-Raman et al., 2012] Venkat-Raman, G., Tomson, C. R., Gao, Y., Cornet, R., Stengel, B., Gronhagen-Riska, C., Reid, C., Jacquelir et, C., Schaeffner, E., Boeschoten, E., et al. (2012). New primary renal diagnosis codes for the era-edta. Nephrology Dialysis Transplantation, 27(12):4414-4419. [Ventura and Luna, 2016] Ventura, S. and Luna, J. M. (2016). Pattern mining with evolutionary algorithms. Springer. [Vijayarani et al., 2015] Vijayarani, S., Dhayanand, M. S., and Phil, M. (2015). Kidney disease prediction using svm and ann algorithms. International Journal of Computing and Business Research (IJCBR) ISSN (Online), pages 2229-6166. [Warady et al., 2015] Warady, B. A., Abraham, A. G., Schwartz, G. J., Wong, C. S., Muñoz, A., Betoko, A., Mitsnefes, M., Kaskel, F., Greenbaum, L. A., Mak, R. H., et al. (2015). Predictors of rapid progression of glomerular and nonglomerular kidney disease in children and adolescents: chronic kidney disease in children (ckid) cohort. American Journal of Kidney Diseases, 65(6):878-888. [Wong et al., 2012] Wong, C. J., Moxey-Mims, M., Jerry-Fluker, J., Warady, B. A., and Furth, S. L. (2012). Ckid (ckd in children) prospective cohort study: a review of current findings. American Journal of Kidney Diseases, 60(6):1002-1011. [Zhang et al., 2005] Zhang, S., Qin, Z., Ling, C. X., and Sheng, S. (2005). "Missing is useful": missing values in cost-sensitive decision trees. IEEE transactions on knowledge and data engineering, 17(12):1689-1693. Centers for Disease Control and Prevention. Chronic Kidney Disease Surveillance System—UnitedStates. Website. http://www.cdc.gov/ck

### 2.2 NECESIDAD NO CUBIERTA QUE SE PRETENDE SOLUCIONAR

Explicar la necesidad de la innovación en base a la conveniencia del problema que resuelve o mejora en lo que respecta a su prevalencia y las características y tamaño del grupo al que va destinado.

El objetivo de este proyecto es desarrollar un Sistema de Apoyo a las Decisiones Clínicas para el manejo de niños con Enfermedad Renal Crónica (ERC) no terminal basado en un conjunto de Modelos Pronósticos. A continuación se resume la situación actual del tema y se justifica la necesidad de este sistema, junto con la oportunidad única de realizarlo.

### Importancia de la Enfermedad Renal Crónica (ERC)

La enfermedad renal crónica (ERC) se define por la presencia de daño renal y la disminución del filtrado glomerular (FG) durante más de tres meses. Una vez establecida, su evolución es progresiva e irreversible hasta alcanzar el estadio terminal. Cuando el cese de la función renal es total es necesario iniciar un tratamiento sustitutivo como la diálisis o el trasplante.

Se estima que el 10 % de la población española tiene ERC y se prevé que esta cifra siga aumentando en los próximos años. El impacto de esta enfermedad es significativo, ocasiona muertes prematuras, aumenta la morbilidad cardiovascular y tiene gran repercusión social y en la calidad de vida así como un elevado consumo de recursos sanitarios. Aproximadamente 4 millones de personas tienen ERC en España con un coste social y económico muy elevado que alcanza el 3% del gasto sanitario total. Es considerado como un problema de salud pública mundial que ha generado el desarrollo de estrategias políticas para su control.

En los niños españoles se estima una incidencia y prevalencia de 10 y 103 pacientes por millón de población pediátrica respectivamente (datos obtenidos del Registro Español Pediátrico de Enfermedad renal crónica REPIR II). Sus consecuencias son aún más devastadoras que en los adultos, afecta a la salud y bienestar de los pacientes hasta la vida adulta con gran impacto familiar, psicosocial y un elevado consumo de cuidados en salud.

A diferencia de los adultos en que la nefropatía diabética y la hipertensión arterial son las causas más frecuentes de ERC, en los niños más de la mitad son secundarias a malformaciones estructurales de la vía urinaria.

### Progresión de la ERC y factores de riesgo asociados

El inicio del daño renal crónico pone en marcha mecanismos de progresión que conducen inequívocamente al daño renal terminal independientemente de la causa que lo originó. El término progresión hace referencia al deterioro gradual de la función renal medido por la disminución del FG. La variabilidad del ritmo de descenso del FG sugiere la existencia de factores biológicos relevantes que pueden influir en la evolución del daño renal.

Sin embargo, existe un vacío de conocimiento importante en el análisis de estos factores de riesgo, su influencia en la enfermedad y las relaciones existentes entre ellos [Warady et al., 2007; Staples et al., 2010; Cerqueira et al., 2015; Warady et al., 2015; Kaspar et al., 2016] que está ampliamente reconocida en la literatura [Floege et al., 2015]. En el siglo XXI, la estandarización de la definición y estratificación de la ERC en las guías KDOQI en el año 2002 [Hogg et al., 2003] y posteriormente las guías KDIGO en 2012 surgidas a partir de [Levey et al., 2005] – ambas iniciativas de calidad-, así como la estandarización de los métodos de estimación del FG [Schwartz et al., 2012] han incrementado el interés por el estudio de la progresión del daño renal.

En la población adulta la hipertensión arterial, la proteinuria y la enfermedad renal primaria son factores de riesgo asociados con mayor progresión de la ERC. Sin embargo la menor incidencia y prevalencia de la ERC en la población pediátrica hace que los estudios sobre la progresión del daño renal crónico sean escasos y que precisen la realización de trabajos colaborativos multicéntricos con cohortes europeas y americanas (ESCAPE y CKiD [Siragy, 2010; Wong et al., 2012]). Los factores mejor conocidos son el grado de severidad del estadio de la ERC, la proteinuria, hipertensión arterial, anemia y otros menos estudiados como el bajo peso al nacimiento o la prematuridad, la hiperuricemia, hiperlipemia, acidosis metabólica o la enfermedad mineral ósea. Se ha demostrado que el control estricto de la tensión arterial y la disminución de la proteinuria en niños enlentecen la progresión del daño renal. Actualmente se han convertido en objetivos terapéuticos en la práctica clínica habitual.

Sin embargo, mucha de la información utilizada en la práctica clínica actual deriva de extrapolaciones a partir de estudios sobre adultos. Estos modelos son inadecuados porque están extraídos en poblaciones con unas características diferentes, y no tienen en cuenta factores muy importantes y diferenciadores como la fuerte interacción entre la ERC y el proceso de crecimiento del niño. Esto limita la información disponible por los clínicos a la hora de decidir el mejor tratamiento para cada paciente, ya que no cuentan con información objetiva [Mak, 2007].

Los registros pediátricos multicéntricos han aportado una valiosa fuente de datos para la investigación de los factores de riesgo asociados a la ERC en el niño, Cabe citar, como ejemplos, [Staples et al., 2010], [Cerqueira et al., 2015], [Mendoça et al., 2015], [Warady et al, 2015] que con poblaciones de distinto

tamaño y estadíos de ERC realizan el análisis de entre 8 y 18 variables con métodos estadísticos clásicos estableciendo modelos de regresión logística.

Estos estudios tienen limitaciones muy importantes. En primer lugar, la mayoría están realizados sobre un conjunto muy pequeño de sujetos, que no es representativo de la población general. En segundo lugar, estos estudios utilizan análisis basados en métodos estadísticos clásicos, como modelos de riesgos de Cox o regresión logística. Estas técnicas son adecuadas para análisis concretos de pequeños conjuntos de datos de tipos homogéneos, pero en general son difíciles de aplicar a grandes bases de datos reales, donde existe un gran número de variables heterogéneas de tipos muy diferentes (nominales, continuas, textuales, etcétera), omisiones en los datos en patrones complejos debidos a multitud de causas (por ejemplo, no todas las medidas sobre un paciente son aconsejables o practicables), etcétera. Estos problemas se acentúan al analizar bases de datos médicas, que presentan estas características de forma intrínseca, debido a la variabilidad de las intervenciones médicas y de la disposición de los pacientes, a la privacidad de la información y a la distinta susceptibilidad a proporcionar ciertos datos para estudio. Por último, los que desarrollan modelos presentan ciertos problemas de validación, ya que pocos de ellos están caracterizados mediante evaluaciones sobre datos diferentes a aquellos de los que se han extraído, lo que en general da lugar a un sesgo de sobreestimación del rendimiento, a menos que se utilicen técnicas estadísticas avanzadas, como validación cruzada o bootstrapping. Sin embargo, muchos de ellos se limitan a realizar un análisis de identificación de factores de riesgo independientes, que no permite desarrollar modelos predictores completos ni investigar la relación entre estos factores.

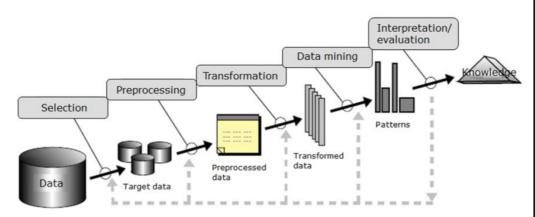
Finalmente, el conocimiento de estos factores, y sobre todo, el de modelos predictivos que permitiesen determinar de forma anticipada y con precisión la evolución que tendrá cada paciente individualizado (ritmo de pérdida de filtración, probabilidad de aparición de enfermedades asociadas, respuesta a tratamientos, potencial variabilidad en la evolución, etcétera) permitiría considerar las diversas opciones de tratamiento de forma personal, así como establecer un ritmo de control y un conjunto de intervenciones adecuado a cada paciente, optimizando con ello el uso de recursos, pudiéndolos dedicar allí donde más impacto tengan.

### Minería de datos

Frente a las técnicas estadísticas clásicas han surgido en las últimas décadas nuevas técnicas de análisis denominadas descubrimiento o extracción de conocimiento en bases de datos o, de forma general, minería de datos. Estas técnicas han sido desarrolladas precisamente para extraer conocimiento nuevo y no evidente que sea útil en procesos de toma de decisiones, como puede ser establecer la mejor actuación sobre un paciente, a partir de colecciones masivas de datos heterogéneos. Actualmente estas técnicas están en auge gracias a los avances en las Tecnologías de la Información y la Comunicación, que han permitido que se disponga en todos los ámbitos de grandes bases de datos de todo tipo, así como de recursos suficientes de computación con un coste relativamente pequeño. Por ello, se han convertido en un elemento estratégico para los sistemas sanitarios y, en general, para la investigación e innovación en biomedicina.

Existen muchas aproximaciones y variaciones a estas técnicas, incluyendo algunas que se han estandarizado en muchos ámbitos, como CRISP-DM (Cross Industry Standard Process for Data Mining), pero todas comparten un conjunto central de actividades comunes:

Selección de datos, donde se establecen qué fuentes de datos se utilizarán en el proyecto.



**Preparación de datos**, en la que se hace un análisis exploratorio para conocer las características de los datos, que guiarán los procesos posteriores, y se corrigen las incidencias que puedan aparecer en los datos, como valores anómalos, "outliers", inconsistencias, valores erróneos, ruido u omisiones. Para cada uno de estos problemas se selecciona y aplica la mejor estrategia, con el

fin de optimizar el rendimiento de los modelos a generar. Finalmente, se estructuran lo datos según los requerimientos de los procesos de minería que se vayan a aplicar. Es un proceso difícil de automatizar, del que dependen los resultados de los proyectos, y frecuentemente consume la mayor parte del tiempo.

**Modelado:** Es la etapa central de la minería de datos. Consiste en aplicar un conjunto de modelos con parámetros libres y ajustarlos a los datos existentes para se adapten de forma óptima a la resolución del problema de información que se tenga. Este ajuste de los modelos puede ser costoso y el rendimiento alcanzado depende de la flexibilidad de los modelos, de las características de los datos y del tipo de búsqueda que se realice, por lo que no se puede estimar con precisión de antemano. Por esto, habitualmente se hace una selección inicial de modelos en función del tipo de problema y de las características de los datos observadas en el análisis y durante esta fase se evalúan y seleccionan entre ellos aquel modelo con el que se obtengan mejores resultados. Existen diversos grupos de problemas, cada uno de los cuáles requiere técnicas diferentes.

**Evaluación e interpretación:** Durante esta etapa se evalúa el conocimiento generado (reglas, patrones, clasificadores, etcétera) en la fase anterior para estimar su utilidad, su novedad, así como el rendimiento que se espera al aplicarlos sobre nuevos datos, es decir, si es aplicable a datos que no se hayan utilizado para desarrollar los modelos, y en qué grado, lo cual requiere de técnicas estadísticas complejas. El conocimiento que pase estos filtros se interpreta para expresarlo en una forma comprensible y útil para la finalidad del estudio.

El conocimiento generado se expresa en forma de modelos y puede adoptar formas muy diversas, dependiendo de la técnica concreta que se aplique. Las más utilizadas en medicina son:

**Modelos de clasificación**: dado un conjunto de entidades (como pacientes) con unas etiquetas (como enfermo o no enfermo), se buscan reglas que permitan conocer las etiquetas en función del resto de datos. Dentro de estos modelos pueden considerarse dos grandes grupos: los modelos de caja negra, que hacen cálculos complejos para calcular la etiqueta (como las redes neuronales); y los modelos interpretables, que pueden representarse en formas que los investigadores pueden comprender fácilmente, como árboles de decisión. Estos modelos, aunque normalmente menos precisos, suelen ser más útiles tanto para guiar la investigación futura (por ejemplo, para identificar por qué funciona el modelo) como para que los médicos clínicos puedan evaluarlo e integrarlo con el resto de conocimiento en su campo.

**Modelos de regresión:** de forma similar a la clasificación, proporcionan un medio de calcular variables cuantitativas en función del resto de datos de cada entidad. Igualmente pueden ser de caja negra, normalmente más precisos, o interpretables.

Reglas de asociación frecuente: en este caso, los algoritmos producen patrones que aparecen con gran frecuencia en los datos, lo que sugiere una relación no conocida. Por ejemplo, se podría encontrar que el valor de una medición sobre un paciente es alto siempre que el de otras dos variables lo es.

**Agrupamiento o clústering:** este tipo de procesos de minería busca identificar agrupamientos que surgen de forma natural entre las entidades estudiadas (como pacientes). Esto permite, por ejemplo, identificar perfiles de pacientes característicos o representativos, y establecer un tratamiento para cada uno por separado.

El proceso de minería no siempre es lineal, de modo que los resultados en una fase pueden determinar que se vuelva a una fase anterior. Por ejemplo, si los modelos evaluados no son satisfactorios, pueden considerarse nuevas transformaciones de los datos y repetir el proceso hasta obtener modelos adecuados. Cada etapa es compleja e implica numerosas técnicas matemáticas y estadísticas.

Hay que señalar que los mejores modelos dependen del problema concreto y de los datos, y no hay un modelo que funcione con todos los problemas, existiendo varios teoremas en este sentido. Además, estimar el rendimiento futuro de forma fiable requiere técnicas complejas. Por ello es fundamental la etapa de selección y evaluación.

El desarrollo de estas técnicas ha estado ligado desde su comienzo a la investigación biomédica. Su uso en el ámbito de la investigación en ERC es prometedor, aunque reciente y aun por desarrollar. Así, pueden encontrarse estudios como [Leung et al., 2013], que utiliza diversos algoritmos de aprendizaje automático para predecir la aparición de enfermedad renal diabética; [Lakshmi et al., 2014] investiga predictores para la supervivencia de los pacientes sometidos a diálisis, mientras que [Vijayarani et al., 2015] lo hace para predecir enfermedades renales. [Barbieri et al., 2015] utiliza una red neuronal para predecir el nivel de hemoglobina futura en enfermos con enfermedad renal terminal. También podemos citar [Singh et al., 2015], que utiliza información longitudinal contenida en los historiales electrónicos de 6435 pacientes para desarrollar un modelo predictivo de pérdida de tasa de filtración glomerular estimada. Estos estudios optan por el uso de técnicas como SVM o ANN que no son fácilmente interpretables, lo que dificulta su aplicación clínica y la definición de nuevas líneas de investigación.

Estos estudios demuestran el interés actual por la aplicación de minería de datos a la enfermedad renal en adultos. Sin embargo, no es posible encontrar estudios específicos sobre esta condición en población pediátrica. Los modelos existentes están adaptados a las características específicas de estos pacientes, que se encuentran en una etapa de desarrollo que determinará su calidad de vida en el futuro. La investigación de predictores específicos supondrá un paso importante hacia la medicina personalizada,

donde los tratamientos se determinan en función de las características individuales y la situación de cada paciente, y no mediante estadísticas generales. Esto permitiría mejorar los pronósticos, reducir costes, y aplicar los recursos donde son más efectivos.

#### REPIR II

La Asociación Española de Nefrología Pediátrica ha creado y soporta dos registros nacionales pediátricos de ERC. El REPIR I recoge aquellos pacientes menores de 18 años con ERC en estadío terminal que recibe tratamiento sustitutivo (diálisis o trasplante), aporta sus datos al registro europeo ERA/EDTA. El REPIR II creado en 2007, está basado en un sistema de información online que recoge los datos demográficos, clínicos, analíticos y de tratamiento de los niños y jóvenes menores de 18 años residentes en España con ERC estadios 2-5 sin tratamiento sustitutivo.

Los datos son aportados voluntariamente por los pediatras nefrólogos mediante el registro de los pacientes incidentes y el seguimiento anual de los prevalentes. En cada centro se ha nombrado a un coordinador que es el responsable principal del registro en dicho centro.

El REPIR cumple con la LOPD y el código de inscripción en el Registro Español de la Agencia de Protección de Datos es 2052500136. Los pacientes y/o sus tutores legales son informados de que sus datos serán almacenados y analizados por ordenador, que la normativa española sobre la gestión de datos computerizada será cumplida y descrita en la información escrita que se les proporciona previamente a su inclusión y que la identificación de datos de los pacientes sólo será posible realizarla por el investigador. Existe un modelo de consentimiento informado de cumplimentación obligatoria a disposición de todos los participantes. A cada responsable de centro se le proporciona una clave y contraseña con la que puede acceder a la página web del registro.REPIR II: https://www.clindemiurg.com/IRC-REPIR/IRCREPIR.asp para incluir los datos de los pacientes.

En la actualidad el REPIR II cuenta con más de 1400 pacientes prevalentes, de los que el 17% son andaluces y 67 centros participantes. El registro contiene un número significativo de variables por paciente.

### Sistemas informáticos de soporte a la decisión clínica

En la práctica clínica la toma de decisiones se realiza basándose en el conocimiento de la fisiopatología, estimaciones de probabilidad, guías de práctica clínica y recomendaciones de sociedades científicas y en la experiencia del clínico. Esto condiciona una gran variabilidad en la práctica clínica que repercute en la seguridad del paciente y dificulta la medición y establecimiento de medidas de mejora.

En los últimos años el desarrollo de Sistemas de software de Apoyo a la Toma Clínica de Decisiones ha comenzado a introducirse en poco a poco en la práctica clínica. De forma general, estos programas consisten en un Sistema Inteligente compuesto por una base de conocimiento con un modelo de la enfermedad o enfermedades a las que se aplican, junto con un motor de inferencia. El motor de inferencia toma los parámetros de los pacientes proporcionados por el clínico y determina uno o varios pronósticos en función de los modelos de la base de conocimiento, junto con sus probabilidades en su caso. En los casos más avanzados, en los que existen variables de intervención que modelan las distintas actuaciones que los clínicos pueden realizar sobre los pacientes, el motor de inferencia puede incluir algoritmos para determinar los valores de estas variables que optimizan los resultados del proceso. La base de conocimiento puede provenir bien de la experiencia de los expertos, recogida de diversas formas, o bien mediante un conjunto de reglas extraídas mediante minería de datos o procesos de machine learning. Como resultado, este tipo de software proporcionan al clínico un pronóstico y un conjunto de recomendaciones de actuación, y posiblemente una explicación extraída del modelo. También se pueden incorporar alarmas que detecten anomalías, ya sea en los datos del paciente o en el seguimiento del proceso asistencial.

Aunque existen intentos para basar estos sistemas en la historia clínica electrónica [Luo et al., 2013; Drawz et al., 2015], la situación en la sanidad española es compleja. Es conocido que los sistemas de historia clínica electrónicos cuentan con mucha información en forma de texto libre, y muy poca información estructurada. La extracción de información a partir de texto libre es un problema abierto, que necesita mejorar antes de poder ser usado en la práctica.

También existen intentos de llevar estos sistemas a los propios pacientes y sus responsables. En este sentido cabe citar a [Swallow et al., 2014], que proporciona a los padres guías de actuación y alarmas para niños con ERC en estadío III.

### Necesidad no cubierta y oportunidad

Tras este repaso por la situación actual de la ERC, se ha justificado que existe una necesidad de desarrollar modelos complejos de alto valor predictivo que permitan pronosticar la evolución de los pacientes pediátricos con esta condición. La obtención de estos modelos se ha visto clásicamente limitada por la difícil aplicación de los modelos estadísticos clásicos. Sin embargo actualmente están madurando las técnicas basadas en *data mining*, que permiten extraer conocimiento en diversas formas a partir de datos reales, y evaluar los modelos de forma rigurosa.

En este sentido, este proyecto dispone, además de la larga experiencia en *Nefrología Pediátrica* de los investigadores que la llevarán a cabo, de la UCAIB de Bioinformática del Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC), que cuenta con Ingenieros y Matemáticos especializados en este tipo de análisis, además de la infraestructura tecnológica necesaria. El IMIBIC cuenta además con

un grupo especializado en *minería de datos* en colaboración con la Universidad de Córdoba, que pueden prestar su apoyo en caso de requerirse el desarrollo de nuevos algoritmos. La aplicación de estas técnicas al desarrollo de modelos de ERC infantil es novedosa, y no hay literatura relevante al respecto.

La gran limitación para este tipo de análisis viene dada por la dificultad para encontrar fuentes de datos fiables, de calidad, y con un número elevado de variables y casos. La disponibilidad de la colaboración con el proyecto REPIR II, un registro pediátrico multicéntrico con 10 años de recorrido y ámbito nacional, supone por tanto una oportunidad única y una ventaja estratégica para la realización de este proyecto.

Por último, para que los modelos desarrollados lleguen realmente a la práctica clínica y no queden en el ámbito académico, se plantea este proyecto como un **proyecto de innovación**, en el que se desarrollará un Sistema de Software de Apoyo a la Toma de Decisiones Clínicas que traslade este conocimiento a las consultas pediátricas.

### 2.3 OBJETIVOS E IMPACTO QUE SE PRETENDE CONSEGUIR CON LA SOLUCIÓN

Enumerar brevemente, con claridad, precisión, por orden de prioridad y de forma acorde con la duración prevista del proyecto, los objetivos concretos que se persiguen. Se valorará la claridad, relevancia y novedad de los objetivos. Recuerde que en este apartado solo deben enumerarse, pudiendo desarrollarse en los apartados siguientes.

### Objetivos generales:

- Mejorar el conocimiento actual de la Enfermedad Renal Crónica no Terminal en pacientes pediátricos en base a datos reales, identificando los factores de riesgo y su influencia e interacción, en forma de modelos útiles que se puedan aplicar a la práctica clínica para mejorar los procesos asistenciales considerando las mejores opciones de tratamiento de forma adaptada a las características y estado de cada paciente, mejorando el pronóstico de estos sujetos, y disminuyendo el coste económico global.
- 2. Dotar a los clínicos de un Sistema de Apoyo a la Toma de Decisiones Clínicas en forma de software que permita aplicar sencillamente este nuevo conocimiento específico, de forma que los pacientes reciban los beneficios del mejor conocimiento de la enfermedad, y resolviendo los problemas de falta de homogeneidad en la práctica clínica en este campo.

### Objetivos específicos:

- 1. Migrar la base de datos del registro REPIR II desde un sistema cerrado no normalizado que no permite su explotación mediante herramientas de análisis de datos avanzadas a un sistema abierto e integrable con dichas herramientas, de forma que se puedan realizar las tareas de minería de datos y adicionalmente se pueda continuar el proyecto sin coste adicional para poder ampliar la muestra y mejorar progresivamente los modelos, y realizar una validación clínica posterior de las herramientas que se construirán durante la ejecución de este proyecto.
- 2. Desarrollar un conjunto de modelos predictivos multivariantes basados en la aplicación de Técnicas de Extracción del Conocimiento en bases de datos (minería de datos) sobre el registro pediátrico REPIR II. Estos modelos buscarán predecir el pronóstico de la enfermedad (si habrá progresión o no para un paciente); el tiempo hasta fallo renal, en su caso; la fracción de pérdida de filtración glomerular que determina la progresión; y la probabilidad de aparición de enfermedades asociadas y eventos adversos, como anemia, problemas de crecimiento, malnutrición, etcétera. Estos modelos se basarán en tres conjuntos de variables: características del paciente, estado de su enfermedad, y variables de tratamiento (medicación, frecuencia de las revisiones, intervenciones realizadas, etcétera), y permitirán a los especialistas determinar qué pacientes tienen mayor riesgo de progresión o de aparición de comorbilidades y por tanto requieren una mayor atención y control, y qué intervenciones minimizan estos riesgos.
- 3. Validar estos modelos en una población externa ajena a aquella sobre la que se han desarrollado, para tener una estimación fiable y no sesgada de su rendimiento predictivo.
- 4. Implementar estos modelos como base de conocimiento en un CDSS (sistema de apoyo a las decisiones clínicas) que informe al médico de los riesgos que presenta cada paciente y sugiera un curso de intervención óptimo basado en ellos, y que pueda validarse clínicamente en el futuro.

No es un objetivo de este proyecto la validación clínica del CDSS para determinar la mejora que supone frente a la práctica clínica actual debido a que la enfermedad tiene, por sus propias características, un curso normalmente lento. Puesto que el proyecto propuesto requerirá dos años de ejecución para conseguir datos adicionales para validar los modelos sobre una población externa, se considera necesario dejar la validación clínica formal del software para una investigación posterior.

### 2.4 METODOLOGÍA Y PLAN DE TRABAJO

Detallar y justificar las actividades o tareas que se van a desarrollar, con indicación de la/s persona/as y/o organizaciones que van a desarrollar cada tarea y con un cronograma de hitos previstos (no inferior al trimestre ni superior al año). Se valorará la viabilidad del proyecto de innovación: adecuación de la metodología, el diseño y el plan de trabajo a los objetivos (Máximo 5 páginas).

Dada la importancia para el proyecto que supone el desarrollo de los modelos predictores, se utilizarán las mejores prácticas recientes en este campo. Se tomarán como base las recomendaciones de TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis; véanse [Collins et al., 2015; Moons et al., 2015]), para eliminar los posibles sesgos y garantizar la reproducibilidad del proceso, evitando los problemas de sobreestimación del rendimiento predictivo de muchos estudios de minería de datos (sesgo de selección, verificación sobre los datos de entrenamiento, etcétera).

Para abordar el proyecto propuesto se ha diseñado un plan de trabajo dividido en 6 etapas. A continuación se detalla la metodología que se seguirá en cada una de ellas.

### 1) REPIR II

Para la realización de este proyecto se partirá del registro pediátrico REPIR II, un sistema de información online que permite el registro de los datos demográficos, clínicos, analíticos y de tratamiento de los niños y jóvenes menores de 18 años residentes en España con ERC estadíos 2-5. Fue creado dentro de la Asociación Española de Nefrología Pediátrica (AENP) en el año 2007 con una comisión constituida por un presidente, dos coordinadores y 10 representantes de los centros participantes. En él registran pacientes 67 centros hospitalarios de diferentes niveles asistenciales y se recogen datos de todas las comunidades autónomas. En cada centro existe un pediatra con dedicación a la nefrología pediátrica responsable de incluir en el registro todos los pacientes incidentes y aportar anualmente los datos evolutivos de los pacientes prevalentes. Actualmente tiene incluidos mas de 1400 pacientes prevalentes. Aporta datos epidemiológicos, clínicos, analíticos y del tratamiento conservador de la ERC que se recoge en un informe anual tras un análisis estadístico.

REPIR II registra actualmente las siguientes variables principales:

Datos	Variables	
Epidemiológicos	Edad Edad de diagnóstico de la ERC Sexo Peso al nacimiento Provincia de nacimiento	Provincia de seguimiento clínico Etiología de la ERC Síndrome polimalformativo Otras enfermedades asociadas
Clínicos	Somatometría	Peso Talla Perímetro craneal IMC Tensión arterial
Analíticos Función renal		Urea/ creatinina FG estimado por fórmula de Schwartz FG estimado por otras técnicas Sodio/potasio Equilibrio ácido/base Proteinuria Albuminuria
	Anemia	Hemoglobina Hematocrito Hierro Ferritina Índice de saturación de transferrina
	Metabolismo Ca/P	Calcio Fósforo PTH Calcidiol

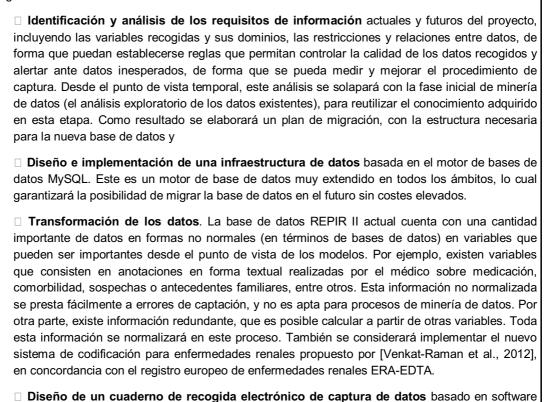
		Calciuria	
Tratamiento (si/no)	Estatinas	Ácido fólico	
	Quelantes del fósforo	Alopurinol	
	Vitamina D3	Sonda nasogástrica	
	Hierro	Gastrostomía	
	Eritropoyetina	Suplementos nutricionales	
	Bicarbonato	Fluorhidrocortisona	
	Resinas de intercambio	Hipotensores	
	Cloruro sódico	Antiproteinúricos	
	Hormona de crecimiento	Aporte de calcio	

El registro REPIR se encuentra implementado en un sistema propietario cerrado que no permite su integración en una plataforma de análisis avanzado de datos, tanto por la incompatibilidad de los sistemas como por estructura no normalizada. Por ejemplo, presenta numerosos campos de texto libre que no se prestan a un análisis automatizado de la información, o campos que se han medido por métodos diferentes. Por ello, el siguiente paso será la migración del sistema a una base de datos abierta y la normalización de la información.

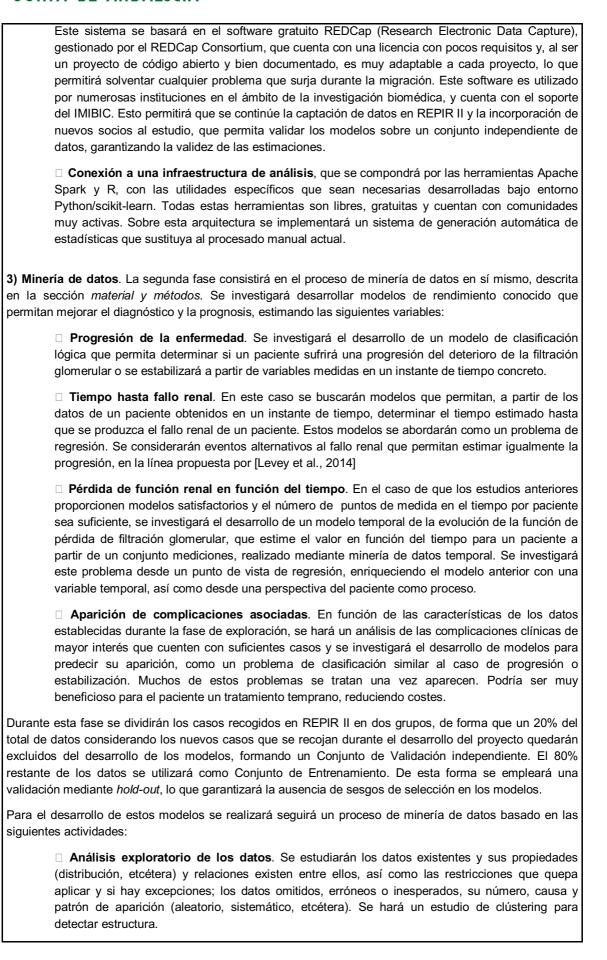
2) Migración de la base de datos. Actualmente, REPIR II se encuentra en un sistema propietario y cerrado que no se presta fácilmente a alimentar un sistema de análisis y minería de datos. El primer paso será migrar estos datos a una base de datos documentada basada en estándares abiertos. Se empleará el motor de gestión de bases de datos MySQL, compatible con el software RedCAP que se utilizará para la gestión de los Cuadernos de Recogida de Datos electrónicos.

Consideramos especialmente importante migrar los datos a una plataforma de software libre para garantizar la supervivencia a largo plazo del proyecto. Los proyectos de minería de datos dependen fundamentalmente de la cantidad de datos disponibles. Utilizar una infraestructura de software estándar y abierta garantizará la disponibilidad de los datos, la calidad del proyecto (por la propia naturaleza del software de fuentes abiertas), y ahorrará tanto costes directos (licencias de software) como indirectos, al eliminar las dependencias de proveedores de servicios específicos, garantizando la interoperabilidad.

Esta migración incluirá:



de fuentes abiertas conectado con la base de datos que permita continuar la recogida de datos.



□ Preparación y transformación de los datos. Se realizará una limpieza de los datos para eliminar el ruido, los *outliers* y los datos anormales que puedan afectar al rendimiento de los algoritmos, decidiendo una estrategia de imputación. Aunque es muy habitual en minería de datos imputar estos valores mediante diversas técnicas (sustitución por la media o moda del atributo, por un valor neutro para los algoritmos, por un valor aleatorio, utilizando un estimador, etcétera), en medicina pueden ser importantes las causas de omisión de datos, como que la medición tenga un coste elevado, riesgo o efectos secundarios para el paciente, de forma que sólo sea justificable la medición ciertos casos. En estas ocasiones es mejor considerar que la falta de datos proporciona información adicional para el modelo (Zhang et al., 2005), pudiendo ser apropiado penalizar aquellos modelos que usen el atributo, utilizando modelos resistentes a casos omitidos (como C5.0), o reservando el atributo en los modelos en forma de árbol de decisión sólo para casos ambiguos que requieran de información adicional para discriminar en ramas de decisión concretas.

Finalmente, los datos se transformarán para que su estructura se preste al proceso de minería, en forma de ejemplos de entrenamiento. Los atributos de estas entidades se transformarán con el objetivo de maximizar el rendimiento de los modelos en función de los datos y el rendimiento obtenido, considerándose discretización (supervisada, no supervisada, basada en entropía, etcétera), normalización y estandarización para los valores numéricos, proyección en componentes y factores, transformación de atributos nominales en lógicos, entre otros.

□ Minería de Reglas de Asociación y Patrones Frecuentes. Se hará un análisis de patrones frecuentes de asociación que permitan descubrir relaciones desconocidas en los datos, es decir, valores de variables se presenten simultáneamente con mucha frecuencia. Para ello se partirá de una discretización de los atributos (forma transaccional), según se describió en la etapa de preparación de los datos, y se aplicará el algoritmo *Apriori* (o de una de sus variantes escalables, como FPGrowth o [Ventura & Luna, 2016]), considerando su extensión a series temporales. Los patrones frecuentes serán evaluados por las métricas habituales (soporte, confianza, *lift, informativeness*), no trivialidad, redundancia y novedad. Aquellos que pasen estos filtros serán evaluados por el equipo de investigadores para establecer su significado y utilidad. Estas relaciones se considerarán para refinar la fase de transformación.

□ Generación de modelos. Teniendo en cuenta que muchos pacientes tienen una observación al año, se considerará la evaluación de modelos predicción que tengan en cuenta la serie temporal de observaciones de cada paciente. En concreto, se considerará el desarrollo de *Modelos Ocultos de Markov* en una variante con observables multivariantes para reconocer los patrones temporales asociados a progresión y a estabilización de la enfermedad renal. Estos modelos permitirían asignar una probabilidad a cada paciente de progresar o estabilizarse en función de un número arbitrario de medidas en el tiempo.

Sin embargo, el número de observaciones por paciente es muy bajo, con una media para los pacientes de 2015 de 3.3 observaciones por paciente, y menos de un 30% de ellos tiene 5 o más observaciones. Los algoritmos para generar los modelos anteriores (como Baum-Welch) requieren un número muy elevado de observaciones para su desarrollo, ya que el número de parámetros libres de los modelos es grande (número de estados de un paciente, probabilidad de transición entre estados, distribución de probabilidad de los observables en cada estado), y es previsible que no sean aplicables. Además, clínicamente es deseable realizar un diagnóstico con el menor número de observaciones posibles, sobre todo atendiendo a la frecuencia anual de las visitas en el estudio.

Por esto, optamos por modelos estáticos basados en la observación de variables de pacientes en un único instante de tiempo o, en función del rendimiento de los anteriores, de un número fijo y pequeño de observaciones consecutivas (dos o tres), que serán tratadas como variables independientes para cada sujeto.

Dado el número relativamente pequeño de variables, no se hará un filtrado inicial, obteniéndose los conjuntos óptimos para cada modelo mediante búsqueda evolutiva en el espacio de atributos.

Modelos de clasificación. Los problemas de progresión o no de la enfermedad y de aparición de enfermedades asociadas se tratarán como un problema de clasificación lógica, con probabilidades asociadas. En este caso, se probarán las familias más representativas de algoritmos de clasificación en aprendizaje automático: vecinos más cercanos, máquinas de soporte vectorial, modelos bayesianos, regresión logística y modelos lineales generalizados mixtos, árboles de decisión (J48, CART), Random Forests, redes neuronales. Para cada modelo se realizará selección de atributos mediante wrapping (por búsqueda evolutiva) y embedding en aquellos métodos que establezcan un mérito para las variables. Para la evaluación durante la selección se utilizará validación cruzada de 10 conjuntos con un número de repeticiones adaptado a la varianza observada, utilizando las métricas de área bajo la curva ROC y κ de Cohen para garantizar que el desbalanceo posible de clases no tenga una influencia negativa. Se realizará un muestreo previo en este caso para equilibrar las clases (pacientes en los que progresa la enfermedad frente a aquellos estables, por ejemplo) mediante submuestreo del conjunto más representado o mediante la generación sintética de instancias mediante el algoritmo SMOTE, en función de las características de los datos observada durante la fase de análisis exploratorio y de la aplicabilidad de este último.

Finalmente, el equipo investigador realizará una evaluará adicional estos modelos de clasificación, favoreciendo aquellos modelos interpretables frente a los no interpretables, en función de los rendimientos obtenidos por ambos. Esta selección se comparará siempre con aquellos modelos de más rendimiento obtenidos con el menor número de restricciones adicionales, de forma que pueda ponderarse adecuadamente la conveniencia y la pérdida de rendimiento. Se considerará igualmente ponderar durante la búsqueda de atributos cada variable por el coste de obtención de la misma, para favorecer modelos fácilmente trasladables a la práctica clínica.

El rendimiento final de estos modelos se evaluará a través de la exactitud, la sensibilidad, la especificidad, el área bajo la curva ROC, la κ de Cohen, los valores predictivos positivo y negativo, J de Youden y el número de variables usadas, así como el coste de obtenerlas.

**Modelos de regresión.** La estimación de tiempo hasta fallo renal y pérdida de función renal en función del tiempo se tratará como un problema de regresión.

Al igual que en el caso anterior, se desglosará cada entrada en la base de datos de aquellos pacientes con información de fallo renal en una instancia independiente, a la que se le añadirá como variable dependiente (a predecir por el modelo) el tiempo desde la visita hasta este evento.

Se procederá a una búsqueda de modelos de regresión más comunes (redes neuronales, modelos lineales generalizados, etcétera, incluyendo modelos de clasificación que permitan regresión, como *random forests*). Los modelos se evaluarán y seleccionarán de nuevo mediante validación cruzada.

El equipo investigador realizará un cribado final entre los modelos encontrados para seleccionar aquellos más útiles desde el punto de vista clínico.

En función de los resultados, se considerará utilizar una discretización de las variables de salida (los tiempos hasta fallo renal o la pérdida de filtración) en lugar de una predicción continua, utilizando términos lingüísticos de utilidad médica y una orientación de clasificación difusa en lugar de regresión.

La evaluación final de estos modelos de regresión se caracterizará mediante la complejidad, el número de variables de entrada, el error cuadrático medio, el error absoluto medio y el error absoluto medio relativo, la varianza relativa media, así como el coeficiente de regresión.

□ Evaluación e interpretación. El paso final de la minería de datos será la evaluación de los

modelos y la extracción de conclusiones a partir de ellos, de forma que se tenga una buena estimación del comportamiento que cabe esperar al usarlos en la práctica clínica y que se pueda utilizar este conocimiento para sugerir nuevas líneas de investigación.

Para ello, y en función de los modelos anteriores y a las variables de cada paciente que utilice, así como a la importancia o mérito de la variable en la predicción en el caso de algoritmos que permitan calcularlo (como C5.0 o Random Forest), y estudiando la influencia de la variable en el aumento de riesgo predicho por los modelos, se establecerán un conjunto de factores de riesgo para cada caso. En esta selección se atenderá tanto a la relevancia de los factores como a su no redundancia dentro del conjunto.

El proceso de minería de datos se realizará utilizando el software RStudio Server, Apache Spark, Microsoft R Server y python/scikit-learn para las implementaciones específicas necesarias.

- 4) Recopilación de nuevos datos de evaluación (Equipo investigador, participantes REPIR II). Durante esta fase, los investigadores continuarán con la recogida de datos de pacientes más allá de los existentes en la base de datos actual y con la que se desarrollan los modelos.
- 5) Evaluación final de los modelos sobre población externa. Los modelos obtenidos y validados mediante validación cruzada desarrollados en las fases anteriores se evaluarán sobre los nuevos datos independientes recogidos en la fase anterior. A partir de estas evaluaciones se extraerán estimaciones independientes del rendimiento de generalización de los algoritmos. Para cada una de las métricas pertinentes se establecerán los intervalos de confianza al 95 %. En caso de imputación de datos, se realizará un análisis de sensibilidad a la imputación mediante imputación múltiple para aquellos algoritmos que no manejen omisiones en los datos de forma intrínseca, en los casos en los que las omisiones sean no ignorables y no pueda modelarse su causa. Se investigarán así mismo las curvas de aprendizaje de los modelos obtenidos, para determinar si los modelos mejorarían con una muestra mayor, y determinar las futuras líneas de investigación.
- **6)** Refinamiento de los modelos y conclusiones (Equipo investigador, IMIBIC). A partir de los resultados que se obtengan en la fase de evaluación se realizará una última etapa de refinamiento de los modelos, seleccionando aquellos que se hallan comportado bien sobre los datos independientes, se extraerán conclusiones, y se determinarán las mejores líneas de investigación para el futuro, con recomendaciones sobre qué datos tomar y con qué frecuencias de muestreo en el futuro.
- 7) Desarrollo de un Sistema de Apoyo a las Decisiones Clínicas. Los modelos predictivos desarrollados y validados en la fase anterior se implementarán en forma de base de conocimiento. Esta será integrada dentro de un software de apoyo a la toma de decisiones clínicas. Este software se implementará mediante servicios web online, de forma que el médico pueda especificar las características de un paciente y el sistema le informe de la probabilidad de progresión de la enfermedad, de aparición de enfermedades asociadas, y le sugiera un curso de actuación (medicación, frecuencia de controles, análisis, etcétera) que optimice la intervención.

Este software se desarrollará bajo la plataforma Java EE utilizando servicios web restful bajo JAX-RS. Se realizará una conexión directa con la base de datos MySQL mediante JPA/Hibernate y se integrará con el Cuaderno de Recogida de Datos RedCAP, de forma que las alarmas y recomendaciones se integren dentro del proceso de registro normal.

Se prestará especial atención al diseño de la usabilidad y la UX (experiencia de usuario), ya que, junto con el rendimiento de los modelos, será esencial para la adopción del sistema.

### Cronograma

A continuación se presenta el cronograma con la planificación temporal del proyecto:

	ENE2017	FEB2017	MAR2017	ABR2017	MAY2017	JUN2017	JUL2017	AGO2017	SEP2017	OCT2017	NOV2017	DIC2017	ENE2018	FEB2018	MAR2018	ABR2018	MAY2018	JUN2018	JUL2018	AGO2018	SEP2018	OCT2018	NOV2018	DIC2018
Nueva DB: Migración de la base de datos REPIR II																								
Análisis y estudio de requisitos de información																								
Diseño e implementación de la infraestructura de datos																								
Normalización y transformación de los datos																								
Implementación de los cuadernos de recogida de datos																								
Modelos: Minería de datos																								
Preparación y transformación																								
Minería de reglas de asociación																								
Generación de modelos																								
Evaluación e interpretación			П.																					
Recopilación de nuevos datos de validación																								
Evaluación final de los modelos																								
Desarrollo de software																								

### 2.5 ASPECTOS ÉTICOS, LEGALES Y ECONÓMICOS DE LA –SOLUCIÓN

Se valorará que se especifiquen los aspectos éticos que se deben tener en cuenta para realizar el proyecto.

El registro REPIR II en que se basa este proyecto cumple con la LOPD y el código de inscripción en el Registro Español de la Agencia de Protección de Datos es 2052500136. Los pacientes son informados de que sus datos serán almacenados y analizados por ordenador, que la normativa española sobre la gestión automática de datos será cumplida y descrita en la información escrita que se les proporciona previamente a su inclusión en el registro y que la identificación de datos de los pacientes sólo será posible realizarla por el investigador. Tras esta información se les solicita explícitamente la firma de un consentimiento informado.

Este consentimiento informado de REPIR II se integrará dentro del CRDe del proyecto, de modo que se recoja explícitamente la firma del paciente o su tutor.

Aunque el proyecto no implica cambios en la intervención sobre los pacientes, requerirá el procesamiento de datos para investigación. Se informará y pedirá autorización al Comité de Ética de la Investigación correspondiente para la realización del proyecto. Siendo un proyecto que implica a menores de edad, se informará por escrito a la Fiscalía de Menores, de acuerdo con la legislación vigente.

### 3. ALCANCE DEL PROYECTO

### 3.1 NOVEDAD, RELEVANCIA E IMPACTO

- Explicar los antecedentes y estado actual del tema del tema (tanto los trabajos previos sobre el tema del proyecto realizados por el equipo de innovación como los realizados por otros grupos nacionales e internacionales (Máximo 3 páginas). Se valorará la relevancia y novedad de en relación con el estado de conocimiento del área científico-tecnológica.
  - Se ha analizado el estado del arte en la sección 2.2), como parte de la justificación de la necesidad no cubierta que se quiere solucionar. Por ello, remitimos al evaluador a esta sección, que contiene una descripción extensa y con citas relevantes.

Como resumen, podemos afirmar que existe un consenso en la literatura actual sobre la necesidad de modelos predictivos específicos para la Enfermedad Renal Crónica no terminal en población pediátrica, de forma que permitan determinar cuáles son los factores de riesgo que explican la gran variabilidad de las respuestas de los pacientes a los tratamientos, y cuál es la relación. Estos modelos permitirían basar las decisiones asistenciales sobre datos concretos. Actualmente estas decisiones se basan en extrapolaciones a partir de estudios sobre adultos, en la experiencia de los médicos, y en Guías de práctica clínica de reciente creación, que a su vez se basan en las anteriores. Esta necesidad ha dado origen a la creación de registros pediátricos sobre la enfermedad, que recogen datos epidemiológicos, clínicos, analíticos y de tratamiento. Sin embargo, la mayoría de los análisis sobre estos datos se basan en técnicas estadísticas clásicas, que no son adecuadas, dado el gran volumen de datos, la heterogeneidad de las variables, y las características propias de las bases de datos médicas (omisiones en los datos, número de parámetros variables para cada paciente, etcétera). Por ello, los análisis se basan en subconjuntos pequeños de variables numéricas o categóricas, y se centran en la identificación de factores de riesgo o en la elaboración de modelos sencillos, como los basados en regresión

logística. Para el análisis de este tipo de bases de datos complejas se ha comenzado a aplicar con éxito un conjunto de técnicas diferentes conocidas como Minería de Datos. Aunque se han empezado a realizar estudios sencillos en el área de la Nefrología, aun no existe ningún estudio importante específico para la población infantil, donde son más necesarios si atendemos a la falta de conocimiento del campo. La principal dificultad para la realización de este tipo de estudios es la disponibilidad de fuentes de datos de calidad.

Por ello, este estudio está alineado con el interés de la literatura actual, cubre un problema en investigación actual, y cuenta con los elementos necesarios para su conclusión con éxito: un registro pediátrico de ámbito nacional (REPIR II), investigadores pediatras expertos en nefrología, y un equipo con experiencia en análisis de datos y en el desarrollo de software (IMIBIC).

Por otra parte, no existe ningún software asistencial para la práctica clínica relevante para esta patología, y sin embargo existe una gran variabilidad en el tratamiento que se aplica a los pacientes. Por ello, un software que permita trasladar directamente el conocimiento extraído mediante estos modelos a la consulta resulta especialmente novedoso.

De nuevo, remitimos al lector a la sección 2.2 para un análisis más detallado.

2) Exponer la existencia de alternativas que resuelven el mismo problema

Aunque existen modelos pronósticos sobre progresión de ERC en pacientes menores en la literatura (como los citados [Staples et al., 2010, Cerqueira et al., 2015, Mendoça et al., 2015, Warady et al, 2015], por ejemplo), estos tienen un alcance limitado, debido fundamentalmente al empleo de modelos clásicos sobre bases de datos relativamente pequeñas. Debido a esto, estos modelos han tenido un impacto reducido en la práctica clínica, sirviendo fundamentalmente para identificar los factores de riesgo para la progresión que se usan en las Guías Asistenciales. Sin embargo, la clasificación de pacientes que permiten es grosera, y no permiten ajustes significativos en los pacientes.

Respecto a modelos basados en minería de datos, sólo existen algunos ejemplos para ERC en adultos, y adolecen de problemas metodológicos. Cabe citar, por ejemplo, [Leung et al., 2013; Lakshmi et al., 2014; Vijayarani et al., 2015; Barbieri et al., 2015Singh et al., 2015], aunque la mayoría están publicados en revistas de bajo impacto. No existen publicaciones significativas para población infantil, por lo que la investigación propuesta constituiría una aplicación marcadamente novedosa de la minería de datos.

Respecto a los sistemas de software asistencial, no existen alternativas relevantes para estos pacientes.

 Indicar la previsión de los efectos a largo plazo, positivos y negativos, primarios y secundarios, producidos directa o indirectamente por la propuesta.

A corto plazo esperamos que los nuevos modelos permitan determinar perfiles de pacientes que respondan de forma diferente a los tratamientos. La identificación de aquellos pacientes con menor riesgo de progresión permitirá reducir la frecuencia de los controles clínicos y analíticos, reduciendo el número de pruebas y visitas innecesarias, con la consiguiente reducción en el gasto para el Sistema de Salud y de las inconveniencias para los pacientes. Por otra parte, la identificación de los pacientes con mayor riesgo de progresión o de aparición de eventos adversos en el curso de su enfermedad permitirá prestar una mayor y más precoz atención a este grupo con lo que esperamos que mejoren los resultados de la intervención. A esto también contribuirá el efecto divulgador de las mejores prácticas que producirá la disposición de un software de asistencia clínica.

En cuanto a la investigación, la identificación de los factores de riesgo y la disponibilidad de modelos que permitan estudiar su influencia conjunta permitirá la postulación de nuevas hipótesis de investigación.

A largo plazo, la disponibilidad de un software de asistencia clínica integrado con el registro REPIR II permitirá aumentar significativamente el número de casos registrados, incorporando nuevos centros al estudio. Esto, a su vez, permitirá la incorporación de nuevas variables para los pacientes, lo que permitirá el desarrollo de modelos más complejos con un coste reducido y amortizado por la utilidad de la herramienta.

4) Explicar si considera que el proyecto tiene una alta probabilidad de obtener resultados protegibles a través del registro de la propiedad intelectual o industrial, en forma de patentes, modelos de utilidad, etc.

El software desarrollado será susceptible de protección a través de un Registro de Propiedad

Intelectual. El equipo del proyecto contará con el asesoramiento de la Unidad de Innovación del Instituto Maimónides de Investigación Biomédica de Córdoba para determinar si hay resultados adicionales que puedan ser protegidos.

### 3.2 APLICABILIDAD DE LOS RESULTADOS

Se valorarán las expectativas de APLICACIÓN A la práctica ASISTENCIAL, DE LA SOLUCIÓN DENTRO DEL SSPA Y DEL SNS

1) LA SOLUCIÓN PROPUESTA ES aplicable e incorpora mejoras en la práctica del Sistema Sanitario.

SI/NC

Justifique la respuesta e indique ámbito de aplicación:

Sí, la propuesta es directamente aplicable a la práctica clínica. Como se ha mencionado, el Proceso Asistencial actual en el tratamiento de la ERC infantil se ha uniformado en los últimos años gracias a la Guías Clínicas como KDOQI o KDIGO, pero aún se basa significativamente en extrapolaciones a partir de estudios sobre poblaciones diferentes y en el criterio particular de cada especialista, existiendo variabilidad en la práctica clínica. La disponibilidad de modelos desarrollados y validados a partir de datos reales permitirá basar las decisiones clínicas en información de calidad, y adaptar los tratamientos a los pacientes de forma individual. Por otra parte, el software de apoyo clínico facilitará la traslación de estas mejoras a las consultas pediátricas. Actualmente no existe un software de este tipo para el manejo de la ERC.

2) LA SOLUCIÓN PROPUESTA ES TRASLADABLE AL SSPA EN SU CONJUNTO Y/O EN Parte a corto, medio largo plazo.

SI/NO

Justifique la respuesta

- Sí. El software de asistencia se implementará mediante servicios web y se podrá integrar en el Sistema Corporativo del SSPA para su uso en todos los servicios de Pediatría que traten a estos pacientes.
- 3) La solución propuesta es trasladable a otros sistemas de salud y tiene valor para el SNS c.

SI/NO

Justifique la respuesta.

- Sí. Los modelos predictores utilizados por el software asistencial estarán validados en una cohorte de pacientes de ámbito nacional, y por tanto serán aplicables a nivel nacional, gracias a la colaboración del registro REPIR II. La implementación como servicios web on-line facilitará la distribución sin modificaciones importantes a otros servicios de salud, incluido el nacional.
- 4) Los resultados tienen potencial para ser adquiridos para su explotación por terceros.

SI/NC

Justifique la respuesta.

Sí. Como hemos justificado, la necesidad de mejores modelos y de software asistencial para el manejo de esta patología es global, por lo que el mercado potencial de la aplicación no se limita al nacional. Países como Portugal o Italia cuentan con registros pediátricos similares a REPIR II, que podrían permitir validar y en su caso reentrenar los modelos para ajustarlos a las características de su población.

Por otra parte, hay que tener en cuenta que el software resultante de este proyecto será un piloto. El objetivo del plan de trabajo ha sido maximizar la efectividad del desarrollo minimizando el coste, dado el presupuesto relativamente bajo respecto al número de tareas a realizar (migración, análisis de datos, validación y desarrollo de software). Pero dada la necesidad del software y el potencial mercado, así como el alto coste que supone el tratamiento crónico de estos pacientes, es factible esperar que alguna entidad se interese en colaborar con el SSPA y la AENP para

mejorar, consolidar y explotar la innovación.	

### 4. DIFUSIÓN

### 4.1 PLAN DE DIFUSIÓN Y EXPLOTACIÓN

1) La solución desarrollada es susceptible de ser incorporada a la práctica del SSPA mediante regulación en el SSPA y/o transferencia a proveedores externos y/o compra pública y/o cambio organizativo y/o cambio en el perfil de competencias de los profesionales.

SI/NO Justifique la respuesta.

Como se ha mencionado en el apartado anterior (3.2.4), la aplicación propuesta tiene muchas posibilidades de extensión, y puede dar origen tanto a la internacionalización de la aplicación propuesta como a la expansión de las patologías cubiertas, completando su valor asistencial.

2) Indicar el <u>plan de difusión y divulgación</u> de los resultados del proyecto de innovación en forma de publicaciones en revistas científicas indexadas en JCR, eventos de divulgación de la ciencia, de innovación, con empresas, etc.

La investigación propuesta es acorde con el estado actual de la literatura, y consiste en el desarrollo de modelos novedosos mediante el empleo de técnicas actuales a un registro consolidado. Como se ha mostrado, existe un gran interés en la investigación actual en este problema, con numerosas publicaciones en los últimos años en el primer cuartil, aunque la mayoría de las publicaciones están limitadas por la falta de datos o las técnicas empleadas. Por ello, el equipo investigador espera que el resultado de este proyecto de origen a al menos un artículo que se publique en una revista de alto impacto, relativo al desarrollo y validación de los modelos predictivos.

3) Indicar el <u>plan de explotación</u> de los resultados y el potencial de los resultados de I+i para ser incorporado a la práctica del SSPA. Para ello realice una breve descripción del mercado en la medida que lo conozca (nº de pacientes o usuarios actuales y previstos a 5 años a nivel autonómico, nacional o mundial, nº de empresas dedicadas a la comercialización de productos/soluciones centradas en estos pacientes/usuarios, inversiones en I+i realizadas para dar soluciones y otros datos relevantes que conozca) y de las oportunidades organizativas y barreras para su implantación.

Dada la aplicabilidad e impacto esperado por este proyecto, tras la consecución del mismo, se espera que se generen resultados innovadores susceptibles de ser protegidos para su posterior explotación comercial por parte de empresas interesadas en el sector.

Este proceso de protección y explotación de resultados se llevará a cabo en colaboración con la Unidad de Innovación del IMIBIC, que cuenta con especialistas en transferencia y gestión de la innovación.

El equipo de Innovación de IMIBIC analizará las posibilidades comerciales y de protección, y, en su caso, iniciará el proceso de protección de los resultados, bien a través de la Oficina de Transferencia de Tecnología del SSPA o la Oficina de Transferencia de Resultados de Investigación de la Universidad de Córdoba.

### 5. MEDIOS DISPONIBLES Y PRESUPUESTO SOLICITADO

### 5.1 MEDIOS Y RECURSOS DISPONIBLES PARA REALIZAR EL PROYECTO

### A) MATERIAL INVENTARIABLE

Para la realización del proyecto se contará con la actual base de datos REPIR II. Para el soporte informático, la UCAIB de Bioinformática del Instituto Maimónides de Investigación Biomédica de Córdoba podrá a disposición del proyecto sus infraestructuras tecnológicas, y entre ellas sus servidores de internet, centro de cálculo, e instalaciones de software necesario (RedCAP, Apache Spark, RStudio Server, MySQL, etcétera). Se contará con el apoyo del resto de infraestructuras y servicios disponibles en el Instituto en caso de ser necesarios.

### B) MATERIAL BIBLIOGRÁFICO

La red de hospitales andaluces en los que se realizará la investigación cuentan con un conjunto de bibliotecas que permiten consultar toda la bibliografía necesaria para realizar el estudio, así como un servicio de intranet que permite, con mayor comodidad, consultar en red todas las publicaciones científicas que se encuentran en la biblioteca. Además, también están disponibles, para su consulta on-line, todas las publicaciones científicas de la web de la Universidad de Córdoba así como de la Biblioteca Virtual del Sistema Sanitario Público de Andalucía.

### C) PERSONAL

Para la realización de este proyecto, además de los investigadores del proyecto, y siendo representados por estos, se contará con la colaboración de los investigadores que participan en el registro REPIR II de la Asociación Española de Nefrología Pediátrica. Actualmente este registro cuenta con la participación de especialistas en pediatría y nefrología repartidos en 67 centros sanitarios a lo largo de 36 provincias. La disponibilidad de este registro y la colaboración de sus responsables representa una oportunidad estratégica para la realización de este proyecto, y permitirá desarrollar modelos predictivos validados sobre una cohorte de ámbito nacional, y con que permitirá desarrollar los modelos y validarlos sobre un número de sujetos mucho mayor de lo que sería posible con una fuente de datos a nivel local, con una mayor validez poblacional. A este respecto hay que tener en cuenta que, en minería de datos, una mayor cantidad de datos con una mayor variedad dentro de la población en estudio es preferible al empleo de algoritmos más complejos.

Se cuenta igualmente con

Se lista a continuación los participantes en el proyecto REPIR II:

Profesional	Centro	Provincia
Ana Campos Aguilera	H. Torrecárdenas	Almería
Francisco Vela Enríquez	Hospital de Poniente. El Ejido	Almería
Emilia Hidalgo Barquero del Rosal	H. Materno-Infantil Infanta Cristina de Badajoz	Badajoz
Marta Carrasco Hidalgo-Barquero	H. Materno-Infantil Infanta Cristina de Badajoz	Badajoz
Angel Vila Lopez	Centro de Nefrología	Barcelona
Margarita Catalá Puigbo	H. General de Granollers	Barcelona
Jordi Vila Cots	H. Sant Joan de Deu	Barcelona
Juan A. Camacho Díaz	H. Sant Joan de Deu	Barcelona
Gloria María Fraga Rodríguez	H. Santa Creu i S. Pau	Barcelona
Gulliem Pintos Morell	H. U. Germans Trias i Pujol - Badalona	Barcelona
Alvaro Madrid Aris	H. Vall d'Hebrón	Barcelona
Gema Ariceta Iraola	H. Vall d'Hebrón	Barcelona
Luis Enrique Lara Moctezuma	H. Vall d'Hebrón	Barcelona
Escribano Subías Joaquín	Hospital Universitari San Joan Deu. Reus	Barcelona

María Herrero Goñi	H. de Cruces	Bilbao
		Bilbao
Mireia Aguirre Meñica	H. de Cruces	2243
Nélida García Perez	Hospital Universitario de Basurto	Bilbao
Gregorio de la Mata Franco	Hospital Universitario Burgos	Burgos
Virgina Roldán Cano	H. Puerta del Mar	Cádiz
Francisco J Romero Salas	H. San Pedro de Alcántara	Cáceres
Sonia Quecuty Vela	Hospital de especialidades de Jerez	Cádiz
Mª Pilar Benito Juvue	H. de la Plana de Vilarreal	Castellón
Jose Luis Badia Mayor	H. General de Castellón	Castellón
Montserrat Anton Gamero	H. Universitario Reina Sofía	Córdoba
Elena García Martínez	H. Universitario Reina Sofía	Córdoba
Cristina Diz Lois	H. Juan Canalejo	Coruña
Inés Vergara Pérez	Hospital Materno-Infantil Teresa Herrera	Coruña
David Barajas de Frutos	Hospital Virgen de las Nieves	Granada
Pablo de Diego Fernández	Hospital Virgen de las Nieves	Granada
Juan Cruz Len Aguilera	Hospital Universitario de Álava. Txagorritxu	Vitoria-Gasteiz
María Amalia Pérez Sáez	Hospital Zumarraga	Guipúzcoa
África Ávila Casas	Hospital Materno Infantil Jaen	Jaén
Tomás del Campo Muñoz	Hospital Materno Infantil Jaen	Jaén
Blanca Valenciano Fuente	H. U. Materno Infantil de Las Palmas	Las Palmas
Marta Fernández Fernández	Complejo Asistencial Universitario de León	León
Araceli García Pose	Hospital Universitario Lucus Augusti	Lugo
Zaira Ibars Valverde	Hospital Universitario Arnau de Vilanova de Lleida	Lleida
Marta Rupérez Lucas	Fundación Hospital Alcorcón	Madrid
Carmen de Lucas Collantes	H. del Niño Jesús	Madrid
Cristina Aparicio Lopez	H. del Niño Jesús	Madrid
Julia Vara	H. Doce de Octubre	Madrid
Mª del Mar Espino Hernández	H. Doce de Octubre	Madrid
Angustias Fernández Escribano	H. Gregorio Marañón	Madrid
Elvira izquierdo García	H. Gregorio Marañón	Madrid
Ma Dolores Morales	H. Gregorio Marañón	Madrid
Marta Melgosa Hijosa	H. La Paz	Madrid
Beatriz Huertes Díaz	H. U. de Getafe	Madrid
Adoración Clara Granados Molina	Hospital Infanta Elena Valdemoro	Madrid
Jorge Luján Martínez	Hospital Sanitas La Moraleja	Madrid
José María De Cea Crespo	Hospital Severo Ochoa de Leganes	Madrid
Carlos María García-Vao Bel	Hospital Universitario del Tajo	Madrid
Teresa Alarcón Alacio	Hospital Universitario Fuenlabrada	Madrid
Pablo Bello Gutiérrez	Hospital Universitario Rey Juan Carlos	Madrid
Daniel Ruiz Díaz	Hospital Quirón Málaga. Centro Integral Pediátrico de Córdoba (Clínica CIP)	Málaga
Alberto Bueno Fernández	H. Carlos Haya	Málaga
Francisco Nieto Vega	H. Carlos Haya H. Carlos Haya	Málaga
Juan David González Rodríguez	<u> </u>	
Carmen de Vicente Calderón	H. Santa Mª del Rosel de Cartagena	Murcia
	H. V. de la Arrivaca	Murcia
Juan Piñero Fernández	H. V. de la Arrixaca	Murcia
Alfredo Rodriguez Do Forno	Complejo Hospitalario de Ourense	Ourense
Isabel Castro López	Complejo Hospitalario de Ourense	Ourense
Fernado Santos Rodríguez	H. Central de Asturias	Oviedo
Flor Angel Ordoñez Alvarez	H. Central de Asturias	Oviedo
		I Dalma da Mallaraa
Mª Dolores Rodrigo Jiménez	H. Son Espases	Palma de Mallorca
Mª Dolores Rodrigo Jiménez Javier Lumbreras Fernández	H. Son Espases	Palma de Mallorca
Mª Dolores Rodrigo Jiménez		

Alfonso Bao Corral	Hospital Xeral-Cíes	Pontevedra
Carla Criado Muriel	H Clínico Universitario Salamanca	Salamanca
Jorge Martínez Jubera	H. Donostia	San Sebastián
Mercedes Ubetagoyena Arrieta	H. Donostia	San Sebastián
Marta Gil Calvo	Complejo Hospitalario de Santiago	Santiago de Compostela
Ana María Sánchez Moreno	H. Virgen del Rocío	Sevilla
Francisco de la Cerda Ojeda	H. Virgen del Rocío	Sevilla
Julia Fijo López-Viota	H. Virgen del Rocío	Sevilla
Elena Pérez González	Hospital Virgen de la Macarena	Sevilla
Fátima Fraga Bilbao	Hospital Universitario de Canarias	Tenerife
María Isabel Luis Yanes	Hospital Virgen de la Candelaria	Tenerife
Ester Parada Ricart	Hospital universitari Joan XXIII	Tarragona
Mª Angeles Fernández	Hopital Virgen de la Salud	Toledo
Empar Lurbe Ferrer	Consorcio H. General Universitario	Valencia
Francisco Jose Aguilar Bacallado	Consorcio H. General Universitario	Valencia
Juan Marín Serra	H. Clinico Universitario Valencia	Valencia
Susana Ferrando	H. Clinico Universitario Valencia	Valencia
Roberto Hernández Marco	H. Dr. Peset	Valencia
Pedro J Ortega Lopez	H. La Fe	Valencia
Santiago Mendizabal Oteiza	H. La Fe	Valencia
Elena Lucas Saez	Hospital de Manises	Valencia
Sara Puente Montes	Hospital Río Hortega de Valladolid	Valladolid
Mª Luisa Justa Roldan	H. Miguel Servet	Zaragoza

### PRESUPUESTO SOLICITADO Y JUSTIFICACIÓN.

5.2

Se deberá desglosar y justificar cada partida del presupuesto solicitado indicando los conceptos, unidades, precio unitario, etc., y, si se dispone de la información, es recomendable señalar el proveedor. En caso de no coincidir con el presupuesto introducido en la aplicación informática prevalecerá el que allí se indicó (Ver anexo).

CONCEPTOS	PRESUPUEST	O SOLICITAD	0	
CONCEPTOS	AÑO 1	AÑO 2	AÑO 3	TOTAL

Bienes y Servicios:			
Material para el desarrollo del proyecto:			
Material Fungible:			
Material Bibliográfico no disponible en la BVSSPA:			
Contratación de servicios externos y arrendamiento de equipamiento: (traducciones, realización de encuestas, mensajería, realización de técnicas específicas, etc.) Empresas colaboradoras y proveedoras de servicios. Se deberá indicar el tipo de servicio a contratar y la necesidad del mismo.			
Subcontratación de la UCAIB de Bioinformática del Instituto Maimónides de Investigación Biomédica de Córdoba para las siguientes tareas:  • migración de la base de datos REPIR II a un sistema abierto adecuado para los procesos de data mining  • creación y alojamiento de los Cuadernos electrónicos de Recogida de Datos  • desarrollo de modelos predictores mediante técnicas de minería de datos  • validación de los predictores sobre una población externa  • creación de un Sistema de Apoyo a las Decisiones Clínicas basado en el software anterior  Total Bienes y Servicios:	9000€	4500 €	13500 €
Personal:  Indique el tipo de personal a contratar (postdoctoral, predoctoral, técnico superior,	13500 €		13500 €
Contratación de un investigador predoctoral como coordinador de la migración / data manager durante seis meses a tiempo completo.  Viajes y Dietas: (Justificación y detalle de nº días y destino) Viajes y dietas para Congresos nacionales			
Viajes y dietas para Congresos internacionales Reuniones de grupo Total Viajes y dietas:			
Otros Gastos:			
Seguros ensayos clínicos			
Difusión de resultados: (inscripciones a congresos para difundir los resultados).  Presentación en congreso europeo.		1500 €	1500 €
Estudios de mercado y del estado del arte de las soluciones			
Total Otros gastos:			

|--|

**Descripción y justificación de la necesidad**: (indicar conceptos, unidades, congresos, reuniones, etc., si no se encuentra justificada se eliminará del presupuesto)

La justificación de la contratación de un investigador predoctoral como data manager se justifica en el siguiente apartado, 5.3. Datos del Personal Contratado.

El proyecto requiere de un análisis de la base de datos REPIR II mediante técnicas de *minería de datos*. Estas técnicas son complejas y necesitan técnicos especialistas y recursos de cálculo que no están disponibles en el equipo de investigación. Por ello, se considera justificada la subcontratación de la UCAIB de Bioinformática del Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC), que cuenta con Ingenieros Informáticos y Matemáticos especializados en el tratamiento de datos. La unidad también proporcionará Ingenieros de Software para el desarrollo de la aplicación final, así como las infraestructuras hardware y software (servidores, instalaciones de RedCAP, Apache Spark, MySQL, RStudio Server, etcétera) necesarias para migrar y alojar la base de datos REPIR II y los Cuadernos de Recogida de Datos electrónicos necesarios para el desarrollo del proyecto.

### 5.3 DATOS DEL PERSONAL SOLICITADO

Cumplimentar solo en caso de que se solicite personal.

Tipo de personal: Investigador predoctoral

Duración del contrato/beca: 6 meses

Horas de dedicación al proyecto: 40 horas semanales durante 6 meses, 960 horas

### Actividades a realizar en el proyecto:

- Definición de los requisitos de información para la migración de REPIR II a un sistema data mining
- Definición del plan de calidad para los datos de REPIR II y determinación del estado actual
- Homogeneización de los datos actuales, que provienen de mediciones con diferentes criterios
- Especificación del Cuaderno electrónico de Recogida de Datos
- Ejecución de un plan de formación para los investigadores de REPIR II
- Coordinación de la transición de REPIR II de su estructura actual a la nueva
- Resolución de incidencias relativas a la migración de los datos

### Justificación de la necesidad:

La contratación de un investigador predoctoral que actúe como *data manager* o gestor de datos se justifica por la complejidad de coordinar la migración de la base de datos REPIR II, que actualmente está en uso por casi 90 investigadores repartidos a lo largo de la geografía nacional, desde un sistema cerrado y cuya estructura no se presta a la realización del tratamiento automatizado que requiere el proceso de *minería de datos*, contando por ejemplo con numerosos campos de texto libre, hacia una plataforma abierta y susceptible de ser integrada con las herramientas necesarias, como Apache Spark. El cambio implicará no sólo el diseño, documentación y prueba de la nueva estructura de datos, junto con sus restricciones de calidad, etcétera, sino la creación de nuevos Cuadernos electrónicos de Recogida de Datos. La definición de los requisitos de información, la conversión de los datos y la coordinación de todos los investigadores durante este cambio no puede ser asumida por los investigadores, dada la alta carga clínica. Por ello, se considera necesario contar con un coordinador expresamente contratado durante esta migración. Se estima en 6 meses a jornada completa la necesidad de este apoyo.

El presupuesto se debe indicar tanto en la aplicación informática de gestión de convocatorias <a href="http://www.juntadeandalucia.es/fundacionprogresoysalud/gestionconvocatorias/inicio.convocatorias.jsp">http://www.juntadeandalucia.es/fundacionprogresoysalud/gestionconvocatorias/inicio.convocatorias.jsp</a> como en la memoria científico-técnica de la propuesta. En ambos casos se deben especificar los importes solicitados por anualidades y por partidas.

Los importes solicitados a través de la herramienta informática y los indicados en la memoria deben ser coincidentes y en caso de no coincidir, prevalecerá el introducido en la aplicación informática.

Las partidas presupuestarias son cuatro: Bienes y Servicios, Personal, Viajes y Dietas, y Otros Gastos

En la memoria, cada partida presupuestaria se deberá desglosar en los conceptos que se van a solicitar. Los conceptos deben estar razonados y adecuados a los objetivos propuestos, indicando las unidades que se van a adquirir, precio unitario, etc., y, si se dispone de la información, es recomendable señalar el proveedor. En este sentido, es recomendable evitar, tanto las globalizaciones, que no aportan información, como las descomposiciones exhaustivas, que resultan innecesarias.

A continuación se detallan algunos de los conceptos subvencionables dentro de cada partida y los importes máximos a subvencionar. Si se solicita un concepto dentro de una partida que no sea la adecuada se pasará su importe a la partida correspondiente.

NOTA: Cualquier importe que sea superior al aquí establecido deberá estar perfectamente justificado y detallado, para que los evaluadores puedan determinar la necesidad del gasto.

### > PARTIDA DE BIENES Y SERVICIOS. En líneas generales se incluyen:

Material para el desarrollo del proyecto: exclusivamente para su uso exclusivo en el proyecto, cuando no esté disponible en el entorno.

Material Fungible: se consideran así los fungibles de laboratorio, papelería y material de oficina, pequeño material informático, fungibles de informática (cartuchos de tinta, etc.), gastos de impresión de posters, etc.

Material Bibliográfico:si no se encuentra disponible en la Biblioteca Virtual del SSPA.

Contratación de servicios externos y arrendamiento de equipamiento: cualquier contratación de servicio (traducciones, realización de encuestas, mensajería, realización de técnicas específicas, etc.). Se deberá indicar el tipo de servicio a contratar y la necesidad del mismo.

CIENTÍFICO	IMPORTE MÁXIMO
Pipetas automáticas	
- Unidad	200-300€
- Multicanal	700€
- Juego de pipetas	600-700€

FUNGIBLE	IMPORTE MÁXIMO
Material de oficina (papelería, informática, etc)	300€ anuales
Gastos de impresión de posters	30-50€/unidad

FORMACIÓN Y DIFUSIÓN DE RESULTADOS*	IMPORTE MÁXIMO
Congreso nacional	600€
Congreso internacional	1.000€
Cursos	600€

CONTRATACIÓN DE SERVICIOS EXTERNOS	IMPORTE MÁXIMO
Traducciones	500€/artículo

### > PARTIDA DE PERSONAL

Los importes máximos necesarios para realizar las contrataciones serán los que tengan fijados cada entidad contratante en función de las categorías profesionales. A modo de orientación, la Consejería de Salud, financiará los importes máximos que se indican a continuación para las siguientes categorías, aunque estas puedan variar. Por ejemplo, para investigadores postdoctorales, hasta 32.000€ incluida cuota patronal de Seg. Social; investigadores predoctorales y tecnólogos, hasta 27.000€ incluida cuota patronal de Seg. Social; técnicos de apoyo, hasta 22.000€ incluida cuota patronal de Seg. Social. En este grupo último grupo se incluye personal que va a dar apoyo en la realización del proyecto (personal administrativo, técnicos de laboratorio, encuestadores, etc.)

En la memoria científico-técnica del proyecto se debe indicar si la financiación solicitada es total o parcial y las actividades previstas para el personal a contratar.

### > PARTIDA DE VIAJES Y DIETAS

Las cuantías para viajes y dietas son las que se encuentran estipuladas por la Orden de 11 de julio de 2006, por la que se actualizan las cuantías de determinadas indemnizaciones por razón del servicio.

### A modo de resumen:

INDEMNIZACIÓN POR UTILIZACIÓN DE VEHÍCULO	IMPORTE MÁXIMO
PARTICULAR	
Automóvil	0,19 €/ Km.
Motocicleta	0,078€/ Km.
DIETAS EN TERRITORIO NACIONAL	
Alojamiento	64,27€
Manutención pernoctando	40,82€
Manutención sin pernoctar	26,67€
½ Manutención	20,41€
DIETAS EN TERRITORIO NACIONAL (MADRID)	
Alojamiento	96,41€
Manutención pernoctando	41,78€
Manutención sin pernoctar	26,67€
½ Manutención	20,89€

PAÍS	POR ALOJAMIENTO (€)	POR MANUTENCIÓN PERNOCTANDO (€)
Alemania	142,04	63,63
Australia	86,77	54,64
Austria	102,19	62,99
Bélgica	158,75	88,70
Canadá	100,91	55,28
Dinamarca	131,12	69,41
Estados Unidos	152,97	74,56
Francia	131,12	70,06
Italia	140,11	67,49
Noruega	142,04	86,13
Reino Unido	167,75	88,70
Suecia	157,47	80,34
Suiza	158,75	65,55

Nota: para información sobre el resto de países se puede consultar la Orden 11 de Julio de 2006.

PARTIDA	IMPORTE MÁXIMO APROXIMADO
DESPLAZAMIENTO EN TERRITORIO NACIONAL	Hasta un máximo de 600€ dependiendo del medio de transporte utilizado.
DESPLAZAMIENTO EN EUROPA	Hasta un máximo de 1.000€ dependiendo del lugar de destino.
DESPLAZAMIENTO RESTO DEL MUNDO	Hasta un máximo de 2.000€ dependiendo del lugar de destino

La cuantía a subvencionar dependerá de las necesidades del proyecto, por lo que <u>se deberán detallar de la</u> forma más aproximada posible el número de viajes, número de días y los destinos previstos.

### Ejemplo:

- 1. Reunión de grupo (conceptos a tener en cuenta):
  - Trayecto: Sevilla-Córdoba-Sevilla
  - Gastos de desplazamiento: Si se realiza en vehículo particular la Orden de 11 de julio de 2006 recoge la cuantía por kilómetro. Si es otro medio de transporte se deberá reflejar un importe aproximado del billete.
  - Gastos de manutención: Según lo establecido en la Orden de 11 de julio, se recoge la cuantía en concepto de manutención pernoctando (si se pasa la noche en el lugar de destino) o sin pernoctar (si no se pasa la noche en el lugar de destino).
- 2. Congreso (conceptos a tener en cuenta):

Trayecto: Sevilla-Chicago-Sevilla

Duración: 5 días

### **CONSEJERÍA DE SALUD**

- Gastos de desplazamiento: Avión (importe máx. aproximado 2.000])
- Gastos de manutención: Según lo establecido en la Orden de 11 de julio, se recoge la cuantía en concepto de manutención pernoctando (si se pasa la noche en el lugar de destino) o sin pernoctar (si no se pasa la noche en el lugar de destino).

Importe aproximados para aquellos casos en los que a fecha de elaboración de la propuesta se desconozca el lugar de celebración del congreso.

- Congreso nacional (3 días aprox.): 650€
- Congreso europeo (5 días aprox.): 1.500€
- Congreso en el resto del mundo (7 días aprox.): 2.500€

### > PARTIDA DE OTROS GASTOS

En esta partida sólo se incluyen costes de seguros para los ensayos clínicos y formación y Difusión de resultados: inscripciones a congresos, cursos puntuales de aprendizaje de técnicas concretas y costes de publicación en revistas especializadas.

No se subvencionarán conceptos como regalos o tarjetas de regalo, ni