

# SHARE: Spanish cAnceR data space

—  
Dr María Torrente, MD, PhD

Hospital Universitario Puerta de Hierro-  
Majadahonda

---

# CÁNCER: PRINCIPAL CAUSA DE MORTALIDAD EN ESPAÑA EN 2023

## Principales causas de muerte en España en 2023

Defunciones	Fallecidos	%
Tumores	115 013	26,55
Enfermedades del sistema circulatorio	114 865	26,52
Enfermedades del sistema respiratorio	46 807	10,81
Enfermedades del sistema nervioso y de los órganos de los sentidos	25 774	5,95
Enfermedades del sistema digestivo	22 923	5,29
Trastornos mentales y del comportamiento	21 970	5,07
Causas externas de mortalidad	17 737	4,09
Enfermedades del sistema genitourinario	17 228	3,98
Enfermedades endocrinas, nutricionales y metabólicas	15 355	3,54
Enfermedades infecciosas y parasitarias (1)	14 865	3,43
Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte	9 698	2,24

# EL PROBLEMA DE LOS DATOS EN SALUD



**Falta de datos de salud estructurados** relacionados con el cáncer y su estandarización



**Falta de diseño prospectivo en los estudios clínicos:** intervenciones reactivas, no anticipatorias



**Elevada complejidad y heterogeneidad de los datos recopilados en oncología:** notas clínicas, resultados de laboratorio, datos genómicos, datos histopatológicos y de imágenes.



**Falta de sistemas de datos** para ayudar al clínico a reducir la incertidumbre



**Calidad de vida:** No sabemos lo que le importa al paciente: no solo es tratar.

**CASO DE USO: Paciente con cáncer de pulmón, varón 55 años, fumador, HTA, sobrepeso. Nuevo diagnóstico de nódulo pulmonar maligno.**

Multiple PET-CT scans

Multiple blood tests

Huge inter-patient variability

Cost-effectiveness of multiple treatments

Acute toxicities and drug interaction

Resistance to treatments

Comorbidities

Family burden

Social/working disruption

Non-evaluated family risk of cancer incidence

Undetected risk of developing secondary malignancies and toxicities

Multiple and/or unnecessary visits to the ER

Real life patients (out of Clinical Trials)

Complex cancer care (with a non evidenced follow-up)

Treatment duration

Long physical and psychological recovery

Management approach towards end of life

Financial hardship

# CLARIFY Platform



Plataforma digital que integra datos clínicos y genómicos de **35,000 pacientes con cáncer de pulmón de 187 hospitales** de España



Analizar subpoblaciones de interés, analizar morbilidad mortalidad cáncer-específica, identificación de variables pronósticas.



Normalización de los datos clínicos para su acceso y análisis.

# CLARIFY PUBLICATIONS WITH LUNG CANCER DATA



## Negation and uncertainty detection in clinical texts written in Spanish: a deep learning-based approach

Oswaldo Solarte Pabón<sup>1,2</sup>, Orlando Montenegro<sup>2</sup>, María Torrente<sup>3</sup>, Alejandro Rodríguez González<sup>2</sup>, Mariano Provencio<sup>1</sup> and Ernestina Menasalvas<sup>1</sup>

<sup>1</sup> Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, Spain

<sup>2</sup> Escuela de Ingeniería de Sistemas y Computación, Universidad del Valle, Cali, Colombia

<sup>3</sup> Hospital Universitario Puerta de Hierro, Madrid, Spain

### ABSTRACT

Detecting negation and uncertainty is crucial for medical text mining applications. Otherwise, extracted information can be incorrectly identified as real or factually incorrect. Although several approaches have been proposed to detect negation and uncertainty in clinical texts, most efforts have focused on the English language. Most previous approaches have used rule-based methods or machine learning techniques.

## Journal of Clinical Oncology

An American Society of Clinical Oncology Journal

Enter words / phrases / DOI / ISBN / authors / keywords / etc.

Newest Articles Issues Special Content Authors Subscribers About ASCO

Journal of Clinical Oncology > List of Issues > Volume 36, Issue 15, suppl >

Meeting Abstract | 2018 ASCO Annual Meeting I

LUNG CANCER—NON-SMALL CELL LOCAL-REGIONAL/SMALL CELL/OTHER THORACIC CANCERS

### Big data for supporting precision medicine in lung cancer patients.

Check for updates

Journal of Thoracic Oncology

P08 HEALTH SERVICES RESEARCH/HEALTH ECONOMICS - FACTORS AFFECTING CLINICAL OUTCOMES  
10, SUPPLEMENT 1, S991-S992, OCTOBER 01, 2021

P08.01 Building Personalized Follow-Up Care Through AI by the Lung Cancer Patient, Data Scientist and Oncologist Together

M. Torrente • F. Franco • V. Calvo • ... L. Costabello • M. Pocs • M. Provencio • Show all authors

DOI: <https://doi.org/10.1016/j.jtho.2021.08.294> • Check for updates



Article

## An Artificial Intelligence-Based Tool for Data Analysis and Prognosis in Cancer Patients: Results from the Clarify Study

María Torrente<sup>1,2,\*</sup>, Pedro A. Sousa<sup>3,†</sup>, Roberto Hernández<sup>1</sup>, Mariola Blanco<sup>1</sup>, Virginia Calvo<sup>1</sup>, Ana Collazo<sup>1</sup>, Gracinda R. Guerreiro<sup>4</sup>, Beatriz Núñez<sup>1</sup>, Joao Pimentao<sup>3</sup>, Juan Cristóbal Sánchez<sup>1</sup>, Manuel Campos<sup>5,6</sup>, Luca Costabello<sup>7</sup>, Vit Novacek<sup>8</sup>, Ernestina Menasalvas<sup>9</sup>, María Esther Vidal<sup>10</sup> and Mariano Provencio<sup>1</sup>

frontiers | Frontiers in

Check for updates

### OPEN ACCESS

EDITED BY  
Pasquale Pisapia,  
University of Naples Federico II, Italy

REVIEWED BY  
Jing-Sheng Cai,  
Peking University People's Hospital, China  
Jorge J. Nieva,  
University of Southern California,  
United States

\*CORRESPONDENCE  
María Torrente  
[mtorrente80@gmail.com](mailto:mtorrente80@gmail.com)

## JCO® Clinical Cancer Informatics

An American Society of Clinical Oncology Journal

Enter words / phrases / DOI / ISBN / authors / keywords / etc.

Newest Content Archive Special Content Authors Subscribers About ASCO Publications

JCO Clinical Cancer Informatics > List of Issues > Volume 7 >

ORIGINAL REPORTS | Artificial Intelligence

### Machine Learning-Assisted Recurrence Prediction for Patients With Early-Stage Non-Small-Cell Lung Cancer

Check for updates

Adrianna Janik<sup>1</sup>, MS<sup>1</sup>; María Torrente<sup>2</sup>, MD, PhD<sup>2</sup>; Luca Costabello<sup>3</sup>, PhD<sup>1</sup>; Virginia Calvo, MD, PhD<sup>2</sup>; Brian Walsh, PhD<sup>3,4</sup>; Carlos Camps<sup>5</sup>, MD, PhD<sup>5</sup>; ...

Show More



Artificial Intelligence in Medicine  
Volume 105, May 2020, 101860



### Reconstructing the patient's natural history from electronic health records ☆

Marjan Najafabadi<sup>1,2,3,4</sup>, Massimiliano Zanin<sup>5,6</sup>, Alejandro Rodríguez-González<sup>7,8</sup>, María Torrente<sup>9,10</sup>, Beatriz Nuñez García<sup>11,12</sup>, Juan Luis Cruz Bermudez<sup>13</sup>, Mariano Provencio<sup>14</sup>, Ernestina Menasalvas<sup>15</sup>

Show more

### Clinical factors influencing long-term survival in a real-life cohort of early stage non-small-cell lung cancer patients in Spain

María Torrente<sup>1,2,\*</sup>, Pedro A. Sousa<sup>3</sup>, Gracinda R. Guerreiro<sup>4</sup>, Fabio Franco<sup>1</sup>, Roberto Hernández<sup>1</sup>, Consuelo Parejo<sup>1</sup>, Alexandre Sousa<sup>3</sup>, José Luis Campo-Cañaverl<sup>5</sup>, João Pimentão<sup>6</sup> and Mariano Provencio<sup>1</sup>

# SOLUCIÓN: SHARE: Spanish cAnceR data spacE

- más datos
- más compartición de datos
- más casos de uso (otros tumores)



# SOLUCIÓN: SHARE: Spanish cAnceR data space

- Desarrollo de estrategias para **detectar el cáncer en etapas tempranas**
- Desarrollar **modelos predictivos de respuesta y riesgo de recaída con datos de población real española.**
- Conocer situación, manejo y **necesidades de largos supervivientes con cáncer.**
- Desarrollar **índices pronósticos de riesgo**
- Entrenamiento de **modelos avanzados de machine learning** en entornos federados de datos.



**GRACIAS!!**