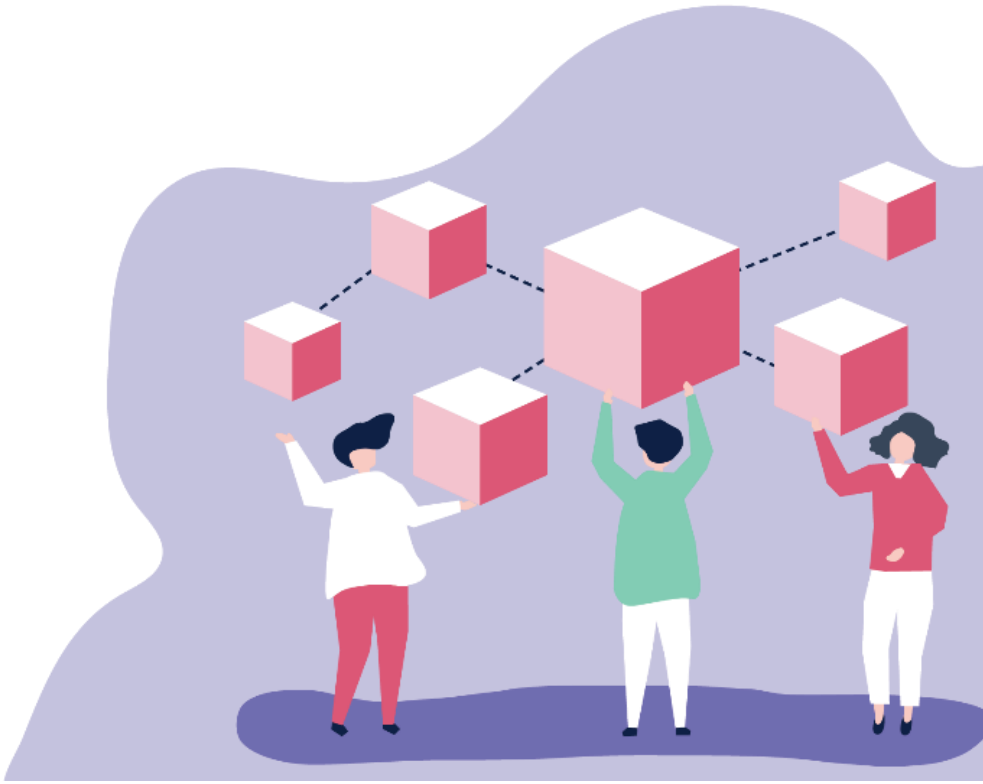
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 857223

GATE KEEPER

# User manual. GastricAITool

| | | | |
|---|---|---|---|
| **Deliverable No.** | | **Due Date** | |
| **Description** | User manual | | |
| **Type** | Report | **Dissemination Level** | CO |
| **Work Package No.** | WP3 | **Work Package Title** | |
| **Version** | 1.0 | **Status** | Final |



# Authors


| Name and surname | Partner name | e-mail |
|---|---|---|
| Rocío Aznar | ITAINNOVA | raznar@itainnova.es |
| Rubén Muñoz | ITAINNOVA | rmunoz@itainnova.es |
| Asunción García | IIS Aragón | agarcia.iacs@aragon.es |
| Patricia Carrera | IIS Aragón | pcarreralasfuentes@gmail.com |
| Vega Rodrigálvarez | ITAINNOVA | vrodrigalvarez@itainnova.es |
| Rafael del Hoyo | ITAINNOVA | rdelhoyo@itainnova.es |


# History

| Date | Version | Change |
|---|---|---|
| 13/04/2023 | 1.0 | Final version |

# Key data


| | |
|---|---|
| Keywords | **User, patient, tool, diagnosis, prognosis** |
| Lead Editor | |
| Internal Reviewer(s) | |


# Abstract

The objective of this document is to present the user manual of the GastricAITool (diagnostic and prognostic tool for gastric cancer).

# Statement of originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

# Table of contents

# List of figures

# 1 Introduction

## 1.1 Project objective

The main objective of the proposal is to develop and deploy an innovative diagnostic and prognostic tool for gastric cancer (GC) built on big data technologies and personalised in GATEKEEPER infrastructure that helps the clinician to take critical decisions and detect risks about patients in a specific moment, based on the information available at this specific time and on an advanced health monitoring based on historical data. This tool will allow the development of more personalised follow-up strategies, early interventions, and therapeutic actions, improving the healthcare system.

## 1.2 Purpose of the document

The purpose of this document is to present the user manual for the GastricAITool (the diagnostic and prognostic tool for gastric cancer). This tool provides diagnosis and/or prognosis of gastric cancer based on patient information. It is an intuitive and easy-to-use tool that has been designed in collaboration with project partners, always with clinical needs in mind. Collaborative work between partners (the Aragon Health Research Institute (IISA): the translational research in digestive pathology group, and ITAINNOVA from the Big Data and Cognitive systems side) has been crucial.

This document includes all the necessary information to use the tool, specifically related to:

- System login.
- User management.
- Creation and editing of patient data.
- Visualization of diagnostic and prognostic model results.

# 2 User manual

## 2.1 System login

Figure 1 displays the main screen of the tool. Here, the user must log in by entering the user name and password. The system will allow access to the tool only if the user name is on the list of users. Otherwise, the administrator must add it to the list.


Figure 1. Main screen. GastricAITool.

## 2.2 User management

Figure 2 shows the User Management System. Only the administrator has permissions to manage users (option located in the upper right corner: "Users management"). Therefore, the administrator is responsible for adding new users to the list, editing already created users, or removing them from the list.

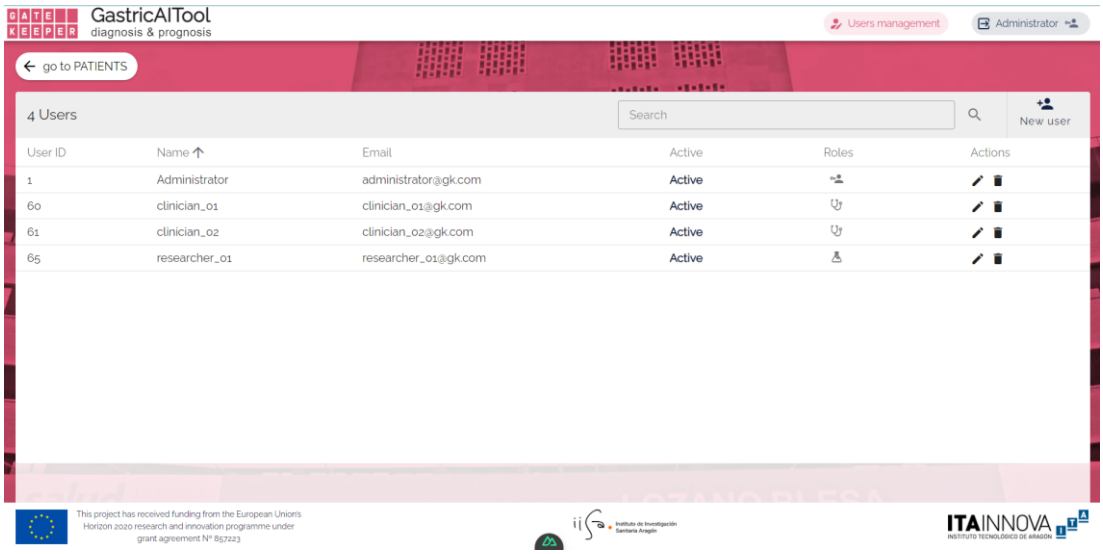



Figure 2. Users' management. GastricAITool.

To create a new user (option located in the upper right corner), the following screen appears as shown in Figure 3. Each created user is automatically assigned an identifier.

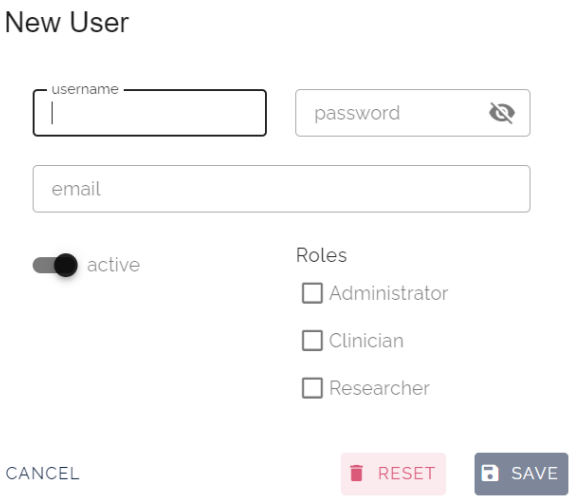



Figure 3. Create new user. GastricAITool.

The fields that identify each user are:

- Username and password: Required to access the tool.
- Email
- Status (active or inactive): Only active users and results of their patients will be displayed in the tool.
- Roles: They refer to the user's permissions to create new patients, delete patients, view patients' diagnostic and prognostic results, and manage users. The options are as follows:
    - Administrator: They have all permissions. Responsible for managing users. Can create and delete patients and view results of all patients.
    - Clinician: They can add new patients, view and delete the patients they have created. They cannot view patients of other clinicians.
    - Researcher: They have only read permissions. Can view the results of all patients.

## 2.3 List of patients

Once the user logs in, the first screen that appears is the list of patients as shown in Figure 4.

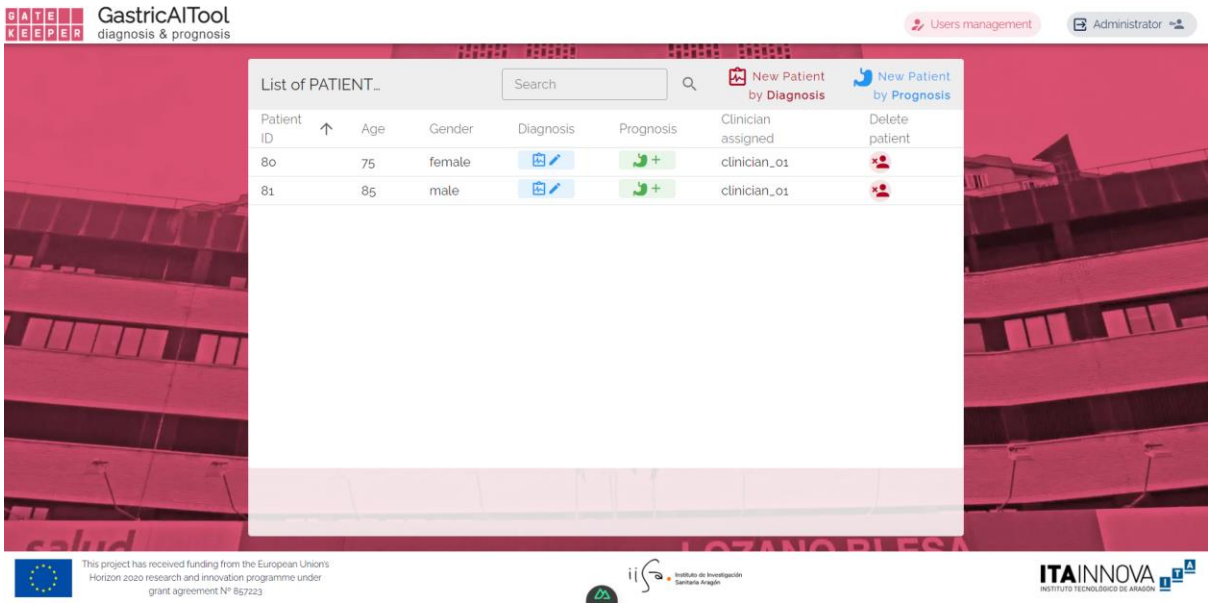



Figure 4. List of patients.

It displays the patients already included with the following information:

- Patient's tool identifier.
- Patient's age.
- Patient's gender.
- Their assigned clinician.
- Option to edit the patient's data related to diagnosis or prognosis or view the results (blue button).
- Option to add patient's data created related to diagnosis or prognosis (green button).
- Delete the patient.
- Include new patients to predict their diagnosis or prognosis ("New Patients by Diagnosis/Prognosis").

Also, a free text search for the patient by identification, age, or sex is offered.

## 2.4 New Patient by Diagnosis/Prognosis

Figure 5 and Figure 6 show the screens that appear when adding a new patient to predict their diagnosis ("New Patient by Diagnosis") or prognosis ("New Patient by Prognosis"), respectively.

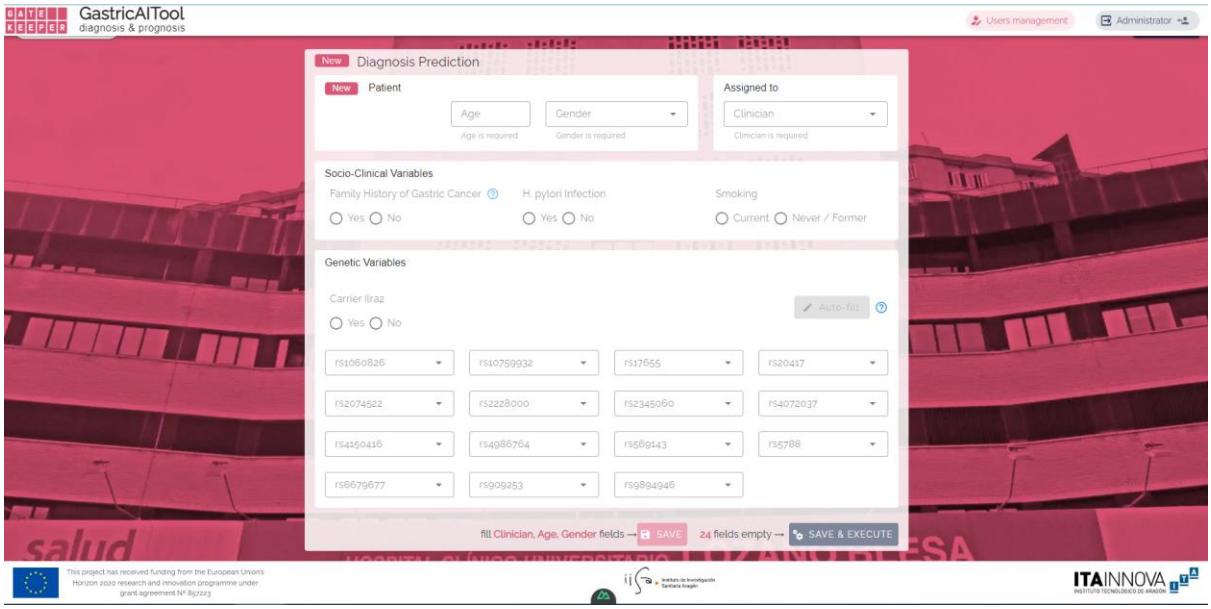



Figure 5. New Patient by Diagnosis. GastricAITool.

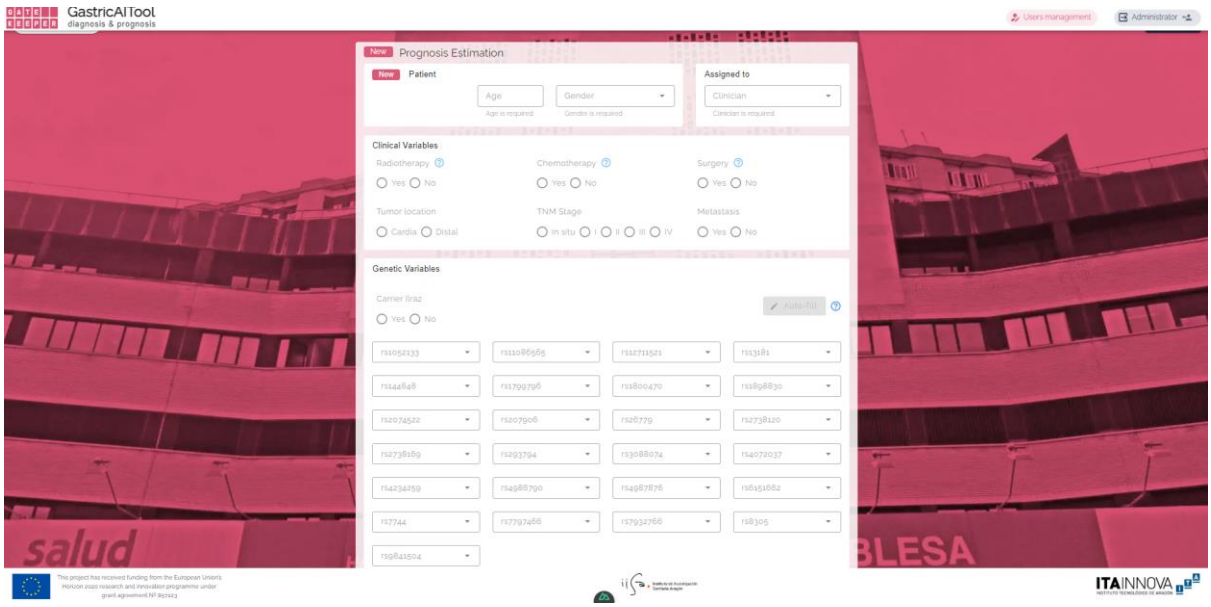



Figure 6. New Patient by Prognosis. GastricAITool.

The following information appears:

- Patient: First, the patient's information (age and gender) and their assigned clinician appear. In this example (Figure 5 and Figure 6), since the role is that of an administrator, it allows assigning from the list of clinicians. If the user were a clinician, they would be automatically assigned to themselves.
- Clinical variables: Second, the clinical variables information appears. Help is provided to detail each information.
- Genetic variables: Finally, the genetic variables to be completed are shown, providing possible options.

- Autofill: In this section, the option to autofill genetic variables is offered if this information is not available, provided that at least 80% of the genetic variables have been completed.

Options:

- **Save**: This button must be clicked first. It allows saving the completed information of the new patient in the database. To do so, it is required to have completed at least the age, gender, and assigned clinician.

  After saving the information initially, an identifier is automatically created in the tool for the new patient, without the possibility of identifying them with the real patient. It will be the responsibility of the assigned clinician to establish this identification in a way that complies with GDPR.

- **Save & Execute**: It allows applying the diagnosis or prognosis model given the patient's information. All information must be completed. As a result, the predicted probabilities by the model are provided, as well as global and local explainability graphs that allow the clinician to interpret the results and understand how the model arrived at that outcome.

Likewise, the tool provides help functions that offer detailed explanations in the points where they are needed.

The same screen shown in Figure 5 and Figure 6 also appears with the option to edit or add patient information from the patient list (blue and green buttons, Figure 4). If the edit option is selected (blue button), the information will appear already completed, along with the results provided by the model, below.

## 2.5 Results

Once all the patient information has been completed, the "Save & Execute" button applies the diagnostic or prognostic model and provides the following information:

- **Diagnosis**
  - Probability of gastric cancer risk.
  - Global Explanation. Features contribution: A graph that provides an explanation of the general behaviour of the model.
  - Local Explanation. Features contribution: A graph that provides an explanation of how the model made the prediction for the particular individual.
  - Local vs. Global Explanation. Features contribution: A graph that incorporates absolute local and global contribution.
  - Genetic Risk Score (GRS): The GRS, normalized between 0 and 10, resulting from the patient's genetic information.
- **Prognosis**
  - Survival probability at 1.5, 3, 5 and 10 years.
  - Predicted survival curve.
  - Global Explanation. Features contribution.
  - Local Explanation. Features contribution.

- Genetic Risk Score (GRS): The GRS, normalized between 0 and 10, resulting from the patient's genetic information.

All graphs are accompanied by explanatory text.

These results can be saved as a PDF report. The following pages show examples of diagnosis and prognosis, respectively.

# Diagnosis Prediction

## Patient

Patient ID: 80

Age: 75

Gender: female

## Assigned to

Clinician: clinician_01

## Socio-Clinical Variables

Family History of Gastric Car: ○ Yes ◉ No

H. pylori Infection: ◉ Yes ○ No

Smoking: ◉ Current ○ Never / Former

## Genetic Variables

Carrier Ilra2: ○ Yes ◉ No

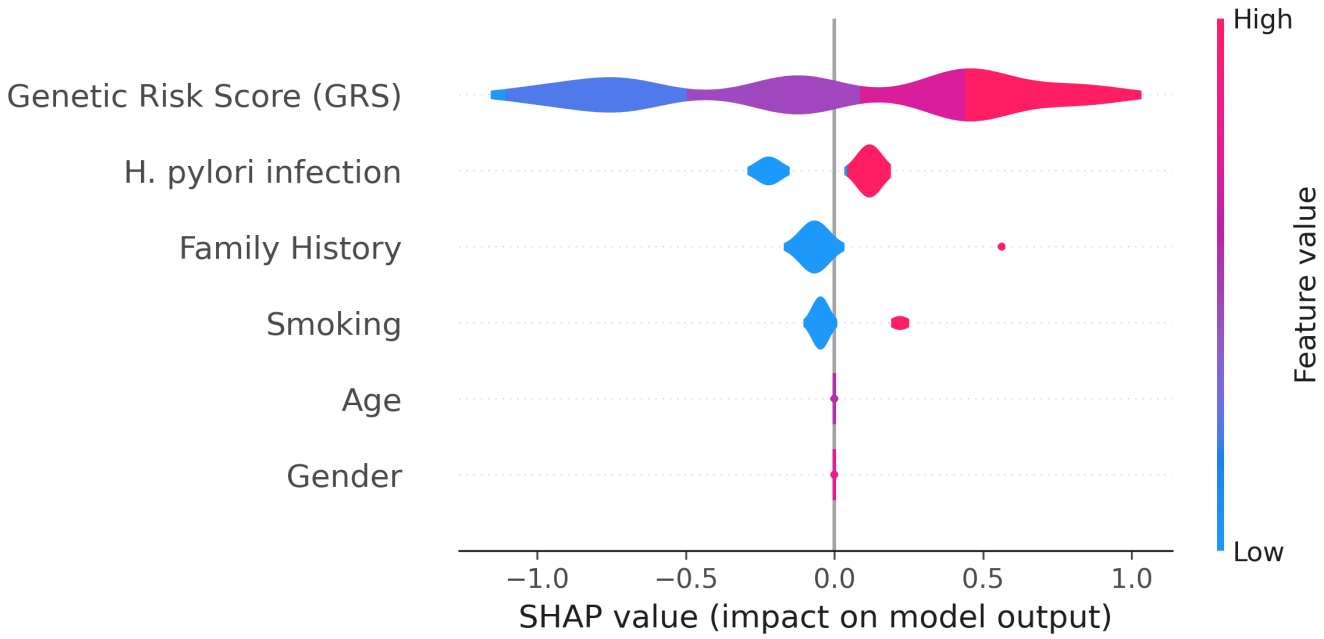
| | | | |
|---|---|---|---|
| rs1060826: AG | rs10759932: CT | rs17655: CG | rs20417: GG |
| rs2074522: AA | rs2228000: CT | rs2345060: GG | rs4072037: TT |
| rs4150416: TT | rs4986764: CT | rs569143: CG | rs5788: CC |
| rs6679677: CC | rs909253: AA | rs9894946: GG | |

RESULTS


Probability of gastric cancer risk

## Global Explanation. Features contribution

This graph represents the global explainability of the model through the contribution that each variable has had in the constructed model. The position of the x-axis is determined by the Shapley value, and the position of the y-axis is determined by the variable, ordered from most to least important. Positive shap values indicate risk of gastric cancer, and negative values indicate the opposite. The color indicates the value of the variable, represented by blue for the lowest values and red for the highest values.

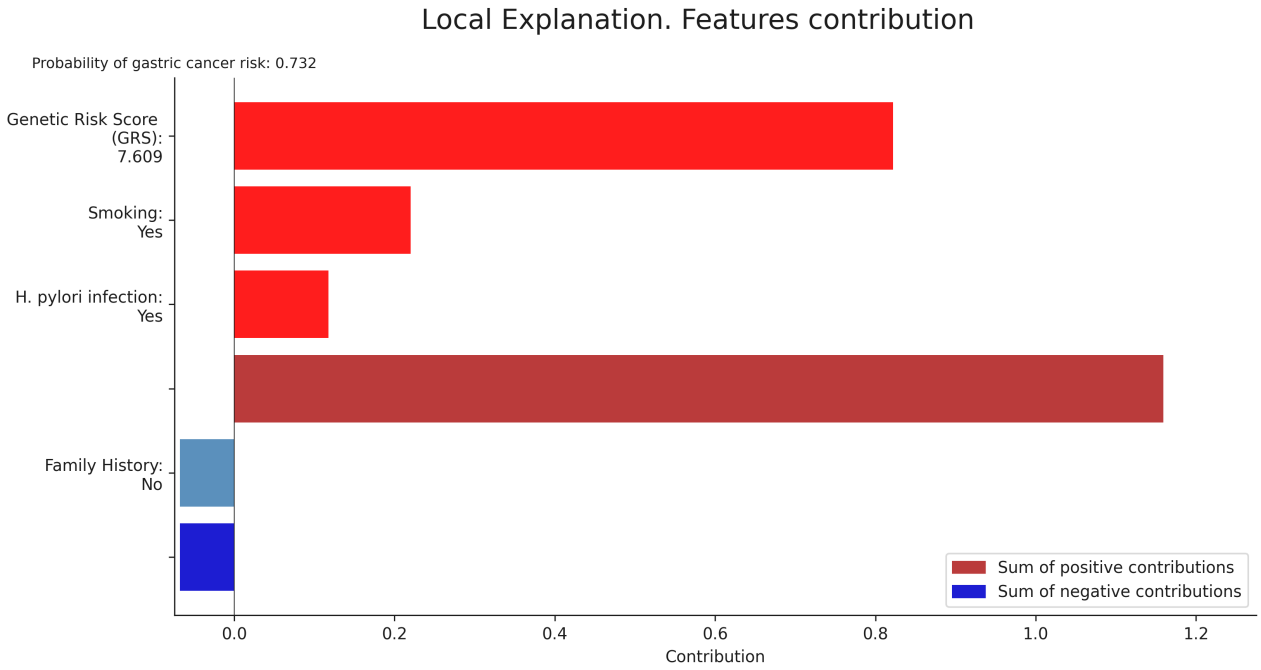



The variables of H. pylori infection, family history or smoking are binary variables where 0 indicates absence and 1, the opposite. The graph indicates that genetics, H. pylori infection and family history are the factors that have most globally influenced the model's decision. High values of the genetic risk score and having a family history of gastric cancer have an increased risk of gastric cancer. Conversely, low values of the genetic risk score or not being infected by H. pylori bacteria are protective factors.

## Local Explanation. Features contribution

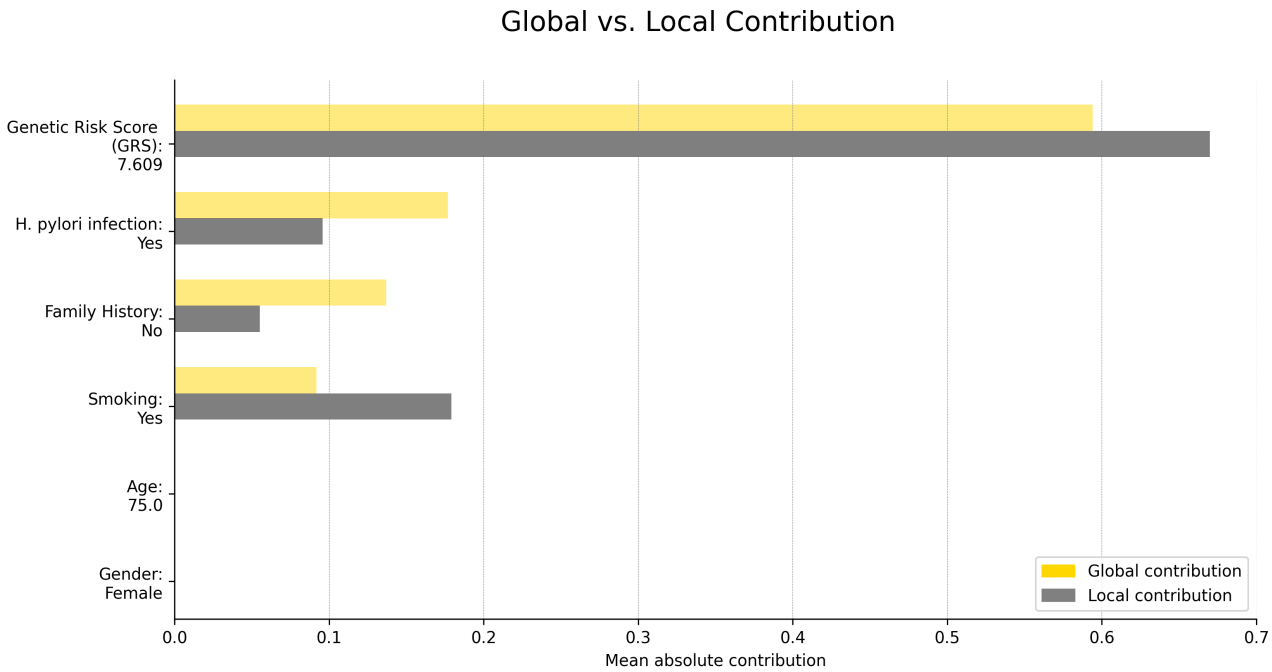
This graph represents the local explainability, i.e. the contribution of each variable (>0) that influenced the model's decision for the particular patient. Positive contributions indicate gastric

cancer risk and negative contributions the opposite.




## Local vs. Global Explanation. Features contribution

This graph represents both the local and the global explainability of the model. The mean absolute contribution of each variable to the model prediction is presented. The sum of the contributions of all variables is 1. This graph allows to compare the importance of the variables globally and in the prediction of the individual.



Genetic Risk Score


Value from 0 to 10

# Prognosis Estimation

## Patient

Patient ID: 82

Age: 80

Gender: male

## Assigned to

Clinician: clinician_01

## Clinical Variables

Radiotherapy: ○ Yes ◉ No

Chemotherapy: ○ Yes ◉ No

Surgery: ○ Yes ◉ No

Tumor location: ◉ Cardia ○ Distal

TNM Stage: ○ In situ ◉ I ○ II ○ III ○ IV

Metastasis: ◉ Yes ○ No

## Genetic Variables

Carrier Ilra2: ○ Yes ◉ No

| | | | |
|---|---|---|---|
| rs1052133: CG | rs11086565: AG | rs12711521: AA | rs13181: GT |
| rs144848: GT | rs1799796: GG | rs1800470: TT | rs1898830: AG |
| rs2074522: GG | rs207906: AG | rs26779: AG | rs2738120: GG |
| rs2738169: AA | rs293794: CT | rs3088074: GG | rs4072037: TT |

rs4234259
AG

rs4986790
AG

rs4987876
GT

rs6151662
GG

rs7744
AA

rs7797466
GG

rs7932766
CC

rs8305
AA

rs9841504
CC

RESULTS


Survival probability at 1.5 years

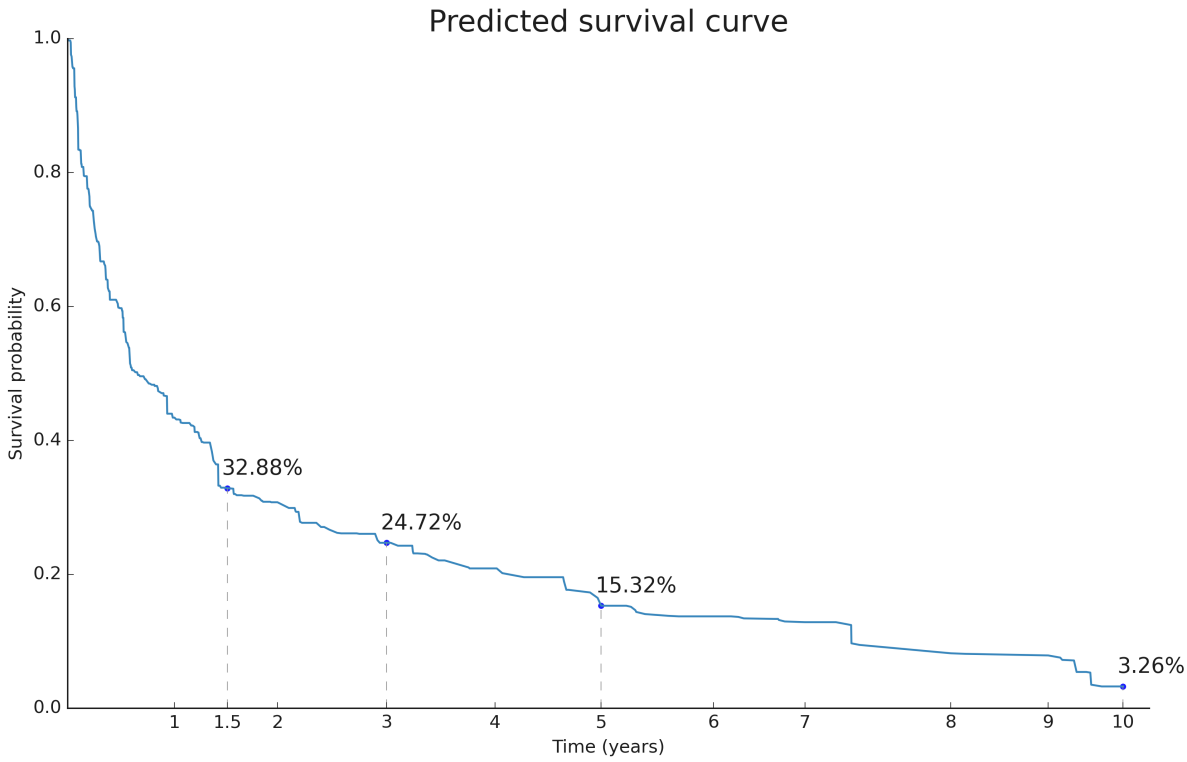

Survival probability at 3 years


Survival probability at 5 years


Survival probability at 10 years

## Survival curve

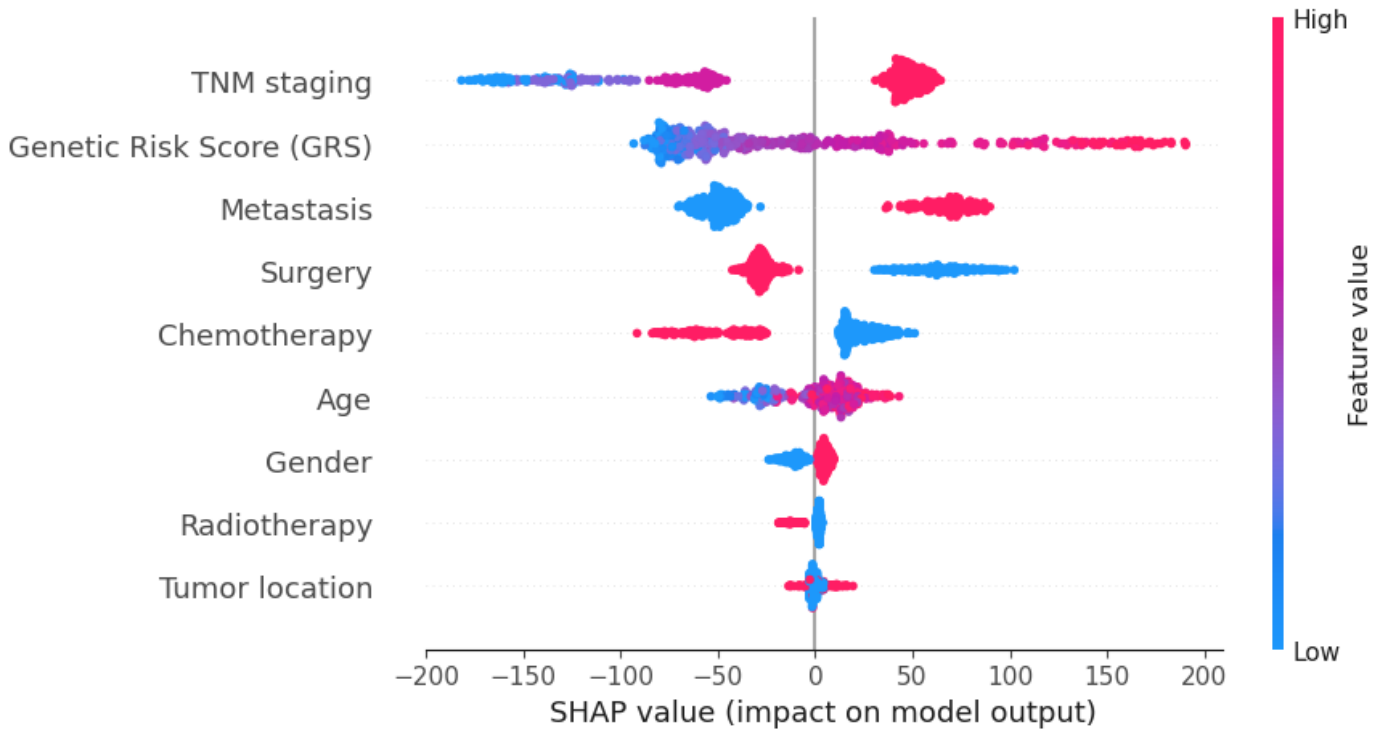
This graph represents the survival curve predicted by the model with a time period up to 10 years. The survival probabilities at 1.5, 3, 5 and 10 years are shown.




## Global Explanation. Features contribution

This graph represents the global explainability of the model through the contribution that each variable has had in the constructed model. The position of the x-axis is determined by the Shapley value, and the position of the y-axis is determined by the variable, ordered from most to least important. Positive shap values indicate risk (poor prognosis), and negative values indicate the opposite. The color indicates the value of the variable, represented by blue for the

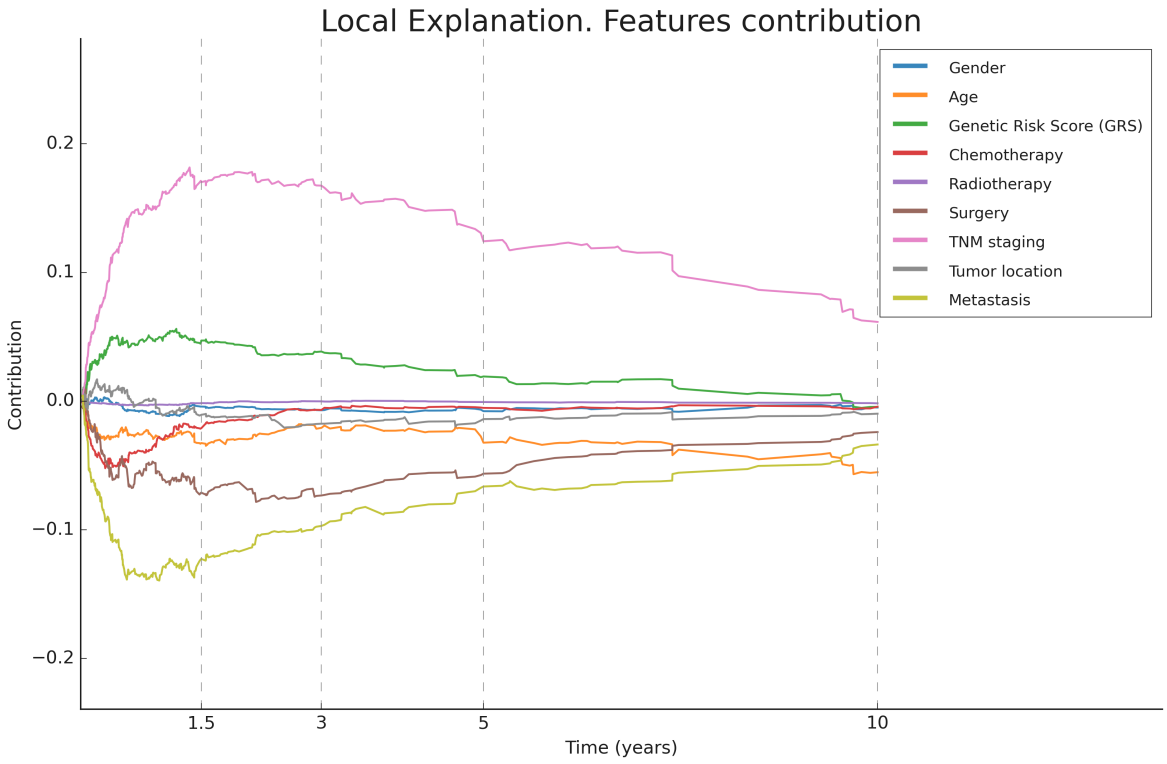
lowest values and red for the highest values. The variables of surgery, chemotherapy, and radiation therapy are binary variables where a value of 1 indicates that the patient has been treated, and 0 indicates the opposite. A value of 0 in the gender variable indicates female, and 1 indicates male.




The graph indicates that the TNM stage and genetics are the factors that have most globally influenced the model's decision, and, conversely, tumor location has had the least influence. Advanced TNM stages, high values of the genetic risk score, having metastasis at diagnosis, not having been treated with surgery or not having been treated with chemotherapy are factors with a negative impact on survival (poorer prognosis).

## Local Explanation. Features contribution

This graph represents the local explainability, i.e. the contribution of each variable that influenced the model's decision for the particular patient. Contributions are represented over time. Negative contributions indicate a higher risk (worse prognosis) and positive contributions the opposite.




Genetic Risk Score

4.48

Value from 0 to 10