MICROB-PREDICT

MICROBiome-based biomarkers to PREDICT decompensation of liver cirrhosis and treatment response

D8.6 Implementation of an economic evaluation model in a software package

Project acronym:	MICROB-PREDICT
Grant Agreement:	825694
Project Duration:	01 January 2019 – 31 March 2025 (75 months)
Version:	V1
Date:	31/03/2021
WP Leader:	Itziar de Lecuona (13 University of Barcelona)
Authors:	Hans Olav Melberg (17 University of Oslo)
Due date of deliverable	Month 27
Actual submission date	26/03/2021
Dissemination level	PU

Table of Contents

1. Executive Summary

We have developed an app that implements an economic evaluation – calculations of costs and benefits – of a diagnostic test for response to Albumin treatment for patients with liver cirrhosis.

2. Introduction

The model for the economic evaluation in the MICROB-PREDICT project has been implemented and is available at the following link: https://microbpredict.anvil.app/

The deliverable is the model implemented in the app, and not this report, but we have included this document to describe the implementation and show how to use the model and the software to conduct an economic evaluation.

3. Background

The MICROB-PREDICT project examines several interventions and diagnostic tools to improve the treatment of patients with liver cirrhosis.

One treatment for the patients with advanced decompensated cirrhosis is Albumin.[1] In the MICROB-PREDICT project (Work Package 7) a version of this type of treatment is tested where the patients receive an almost every week for 6 months. This type of treatment is believed to halt the progression to more serious conditions, but the treatment is costly and does not work equally well for all patients. Moreover, Albumin may have some serious side-effects.[2]

The MICROB-PREDICT project tries to identify biomarkers and develop diagnostic tests that better identifies patients that are likely to respond to Albumin treatment. The benefit of such a test would be to avoid giving an expensive treatment to patients who would not respond, or who would respond negatively. However, the test will also create costs. One question is then whether the increased predictive ability justifies the costs.

The economic evaluation depends on the sensitivity and the specificity of the diagnostic test, as well as the cost and effect of the Albumin treatment. At this point in the project the results about how effective Albumin treatment is (WP 7), are not yet available. However, it is possible to conduct an economic evaluation of different scenarios: Under different assumptions of the effects of the Albumin

[1] EASL Guidelines J Hep 2018, J Hepatol. 2018 Aug;69(2):406-460. doi: 10.1016/j.jhep.2018.03.024. Epub 2018 Apr 10.PMID: 29653741

[2] See, for instance, Lancet. 2018 Jun 16;391(10138):2417-2429. doi: 10.1016/S0140-6736(18)30840-7. Epub 2018 Jun 1.PMID: 29861076

treatment, how accurate does the diagnostic test have to be for it to be cost-effective? This is useful because the conclusion of such an analysis would help decide how many biomarkers need to be tested to get a test that is accurate enough to justify the costs.

4. Economic evaluation

There are three general types of economic evaluation: First, cost-effectiveness where the medical outcome is measured in natural units like 'reduction in blood-pressure' and similar medical units. Second, cost-utility, where the outcome is measured by the increase in life-expectancy of an intervention weighted by the health-related quality of the years gained (often, but not always, measured by quality adjusted life years, QALY). Finally, there is cost-benefit analysis where the outcome is measured in a monetary unit: The amount of euros gained from the intervention.

The economic evaluation in the MICROB-PREDICT project uses the cost-utility approach. This is most commonly used in the medical field because – unlike cost-effectiveness - it aggregates and includes more than a single outcome measure. Outside the medical field, cost-benefit is also more common, but in medical interventions it is often difficult – practically and normatively - to give a monetary value to increases in quality of life and life-expectancy. For this reason, we use the cost-utility approach.

In the cost-utility approach we measure the gains from the intervention by comparing the life expectancy and quality of life for patients in a system without the diagnostic test (the benchmark model) and the expected result with the diagnostic test (intervention model). In the implemented model we use the health care perspective to model cost (including the cost of the test, the cost of administering and analyzing it). A social perspective that also includes lost labour productivity is not possible since there is very little information about the degree to which the test and the treatment improves labour participation and productivity.

A comparison of the costs and QALYs gained in the benchmark model and the model with the new diagnostic test, gives the incremental cost-effectiveness ratio (ICER) i.e. the cost per QALY gained from adopting the new diagnostic test. If the costs per life-year gained are lower than a given threshold, the intervention is said to be cost effective. The threshold varies between different countries, and in some countries, there are no explicit thresholds. For instance, in the UK, an intervention that costs more than 30 000 pounds per life year gained (adjusted for quality), is considered not cost-effective and it will not be adopted. In Norway, for instance, the threshold is not as explicit, and it depends on the severity of the disease.

To estimate the cost per QALY gained from the diagnostic test, it is necessary to develop a model that allows us to estimate the expected costs and benefits in a system with and without the test. The models are needed since we usually only have data on patients for a limited time period, while the

costs and benefits may last a lifetime. The model allows us to use the data we have and based on this estimate the average lifetime costs and benefits for patients.

Several models can be used: Decision trees (where patients convert from one state to another and never return back), Markov models (where patients change forth and back between different states with given probabilities), Discrete Event Simulation (where patients jump to their next state after a randomly drawn time unit based on given probability distributions) and agent based simulations (where patients' characteristics affect the probability of the state the patient will be in in the next time period).

In the economic evaluation of a diagnostic test for the response to Albumin treatment, we have used a combined decision tree and Markov model approach. The justification is based on the fact that, first, there is a short-term decision, and treatment that has to be modeled (the test of whether the patient is likely to respond to Albumin and the result of the initial 6-month treatment depending on the sensitivity and specificity of the test and the proportion of patients that are true responders vs. non-responders to Albumin). Next, after the initial treatment, the model uses a Markov model to account for the fact that liver cirrhosis is a disease with changes and consequences over a lifetime. This makes it important to capture the effect of Albumin over the lifetime and not just the first six months. A Markov model allows us to set up the different states the patient can be in, to specify the costs and utilities associated with the different states, and the probabilities of going from one state to another – depending on the initial treatment outcome. Based on this, we can simulate what would happen to a group of patients and find the average lifetime costs and benefits.

5. Software implementation

To create and simulate the results of a Markov model, it is necessary to use software. There are different kinds of software that can be used: Excel, TreeAge and computer programming. In the MICROB-PREDICT project, we implemented and developed the model in a computer program (Python), because this enables a fast and flexible estimation. However, computer programs can be complex and one of the aims of MICROB-PREDICT is also to build a model that can easily be adopted to new information and new tests or interventions. For this reason, the model has been implemented in an online app that only requires the user to input the information as pure text, without having to know computer programming. The way this works will be described next.

6. A textual representation of a model

A Markov model consists of the different states a patient can be in (for instance healthy, sick, dead), the probability of going from each state to another state (also called the transmission matrix), and the costs and utilities of a patient in the different states. In the developed framework, this information is

provided using keywords. Here is an example of a textual representation of a model that can be inserted into the app:

```
name: example  (Markov)

state = healthy, sick, dead

cost
  healthy=0
  sick=1000
  dead=0

utility
  healthy=1
  sick=0.8
  dead=0

probability
  healthy=0.9,0.1,0
  sick=0.5,0,4, 0.1
  dead=0,0,1
```

In addition to the costs, utilities, and probabilities, it is useful to allow the specification of important information relevant to estimate the model, such as the discount rate and other important model parameters (specified under the keyword *info*). We also need to specify an initial distribution of the patient population (e.g. that they all start out as healthy), using the keyword *population*. It is also useful to include probabilities associated with age and general mortality (keyword *mortality*). Finally, in order to capture uncertainty, the parameters should be allowed to be drawn from probability distributions, as opposed to be constants. These can be specified and given a name (that can be used later in the model description) under the keyword *variable.* For instance, a simple textual model including all these features is presented below:

```
name: example  (Markov)

info
  discounting = 0.03

state = healthy, sick, dead
```

```
variable
  cost_sick = random.gamma(100,10)

cost
  healthy=0
  sick=cost_sick
  dead=0

utility
  healthy=1
  sick=0.8
  dead=0

probability
  healthy=0.9,0.1,0
  sick=0.5,0,4, 0.1
  dead=0,0,1

mortality
  0-10 = 0.01
  11-20 = 0.02
  20-40 = 0.03
  40-60 = 0.05
  60- = 0.08
```

The advantage of using an app and a textual representation of the model, is that is becomes much easier to change the model for users that are not experienced computer programmers. For instance, in order to examine the effect of an intervention, they may increase the probability of going from sick to healthy, but also the cost of treating a person when sick. When pressing the button 'simulate', one gets new estimates for average lifetime costs and benefits under the new assumptions.

In addition to the 'simulate' button, there is a button for making a visual representation of the text-based model structure (plot model creates a directed graph visualization), and a button to plot the number of individuals in the different states at each step in the simulation. The screenshot below shows how the model is implemented in the app (**Figure 1**), the available buttons, and an illustration of the output (after pressing the *plot model* button).

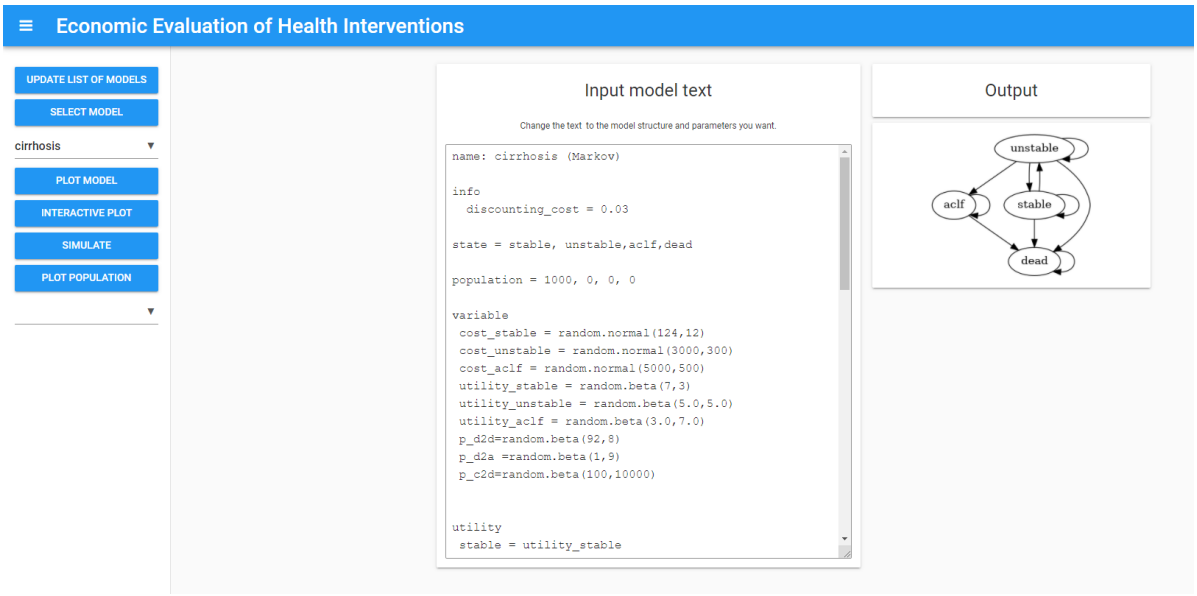

Figure 1: A screenshot of the MICROB-PREDICT cirrhosis model implemented in the software application.

7. Building a model for evaluating a test for response to Albumin

The starting point for building the specific economic evaluation model of a diagnostic test for the response to Albumin is **Figure 2**. The pathway described in the 'before MICROB-PREDICT' line represents the benchmark case. The key states in the model are: Stable cirrhosis, unstable cirrhosis, ACLF and dead. Each step in the cycle represents three months. The costs for an individual in a state (for three months), is based on information from cost-studies of cirrhosis patients. The transition probabilities are chosen to make the average length of stay in the different states equal to the information in the figure (10 years in cirrhosis, 2-3 years in decompensation, 3 months in ACLF.[3] The information about utility in different states is from an article published by Wells *et al*. [4]

To capture uncertainty, the parameters were drawn from probability distributions where the lower and upper values (95% confidence interval) were given by 20% below and above the mean. For probabilities and utilities, the beta distribution was used to make sure that the values ended up between 0 and 1. The costs were modelled using a gamma distribution to make sure that no values were below zero, and because this also allows for some very costly individuals.

[3] For information about the costs and the probabilities that we have used, see our publication: Asphaug L, Thiele M, Krag A, Melberg HO (2020) Cost-Effectiveness of Noninvasive Screening for Alcohol-Related Liver Fibrosis. Hepatology 71:2093–2104. https://doi.org/10.1002/hep.30979

[4] Wells CD, Murrill WB, Arguedas MR (2004) Comparison of Health-Related Quality of Life Preferences Between Physicians and Cirrhotic Patients: Implications for Cost–Utility Analyses in Chronic Liver Disease. Dig Dis Sci 49:453–458. https://doi.org/10.1023/B:DDAS.0000020502.46886.c1

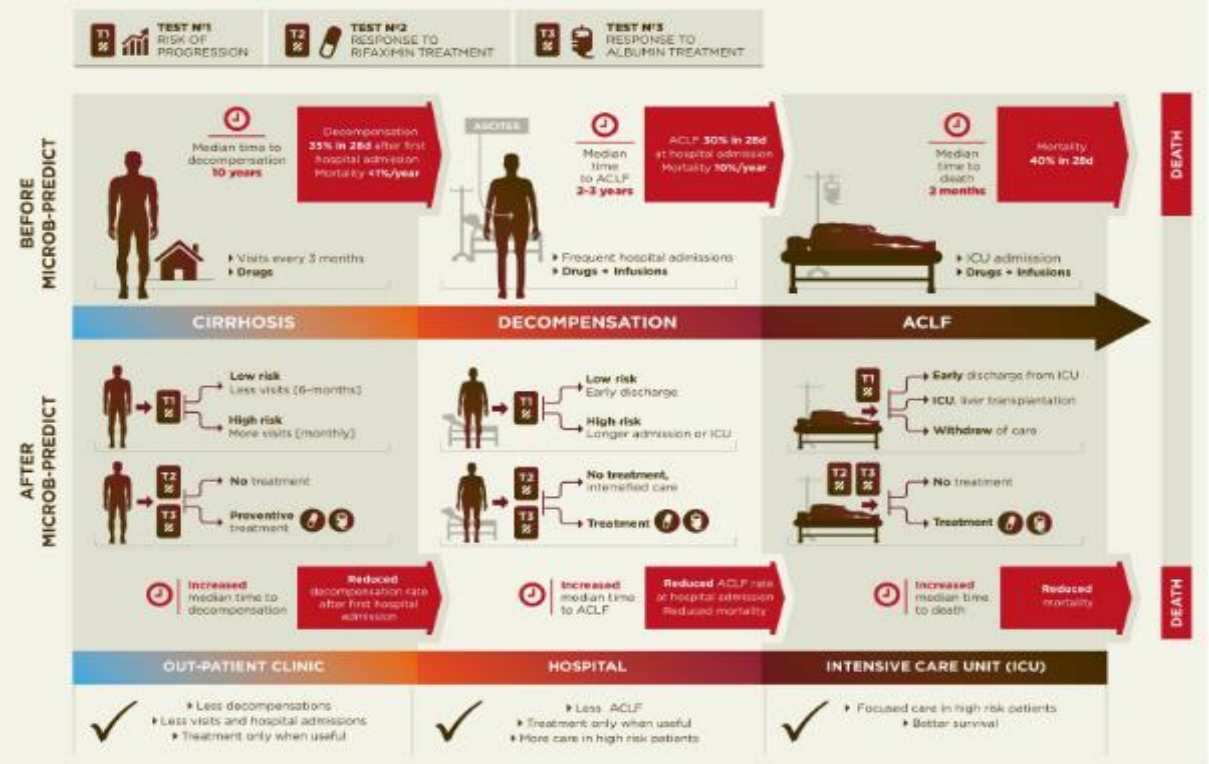

Figure 2: A simplified representation of the situation before and after a test for Albumin response in MICROB-PREDICT.

To model the diagnostic test, we introduce a decision tree before the Markov model which represents the test and the end-result. There are four possible cases: The test may show a patient that is predicted to respond well to Albumin or not. However, there are false positives and negatives, because depending on the sensitivity and specificity of the test, some individuals who are predicted to respond well, may respond poorly and some that are predicted to not respond may in fact be good responders. This creates the four possible end nodes in the decision tree:

1. Positive test and Albumin treatment for a responder
2. Positive test and Albumin treatment for a non-responder
3. Negative test and no Albumin treatment for a responder
4. Negative test and no Albumin treatment test for a non-responder

For each end node, there is a Markov model that creates the average lifetime cost for the type of situation described by the end-node in the decision tree. These models are variations on the benchmark model where we change the following parameters:

1. **Change in benchmark Markov model after End node 1: Positive test for a responder:** A positive test for responders leads to a Markov model with the use of Albumin (increase treatment cost parameter in the model), but also a smaller probability of transferring to a worse state.

2. **Change in benchmark Markov model after End node 2: Positive test for a non-responder.** Higher costs, but no change in the probability of transferring to a more serious condition.
3. **Change in benchmark Markov model after End node 3: Negative test for a responder:** No increase in cost, but also no change in transition probabilities.
4. **Change in benchmark Markov model after End node 4: Negative test for a non-responder:** A decrease in costs, no change in transition probabilities.

Based on these assumptions and models, the computer model estimates the average lifetime costs for 1000 individuals who enter the model.

8. Results

The model has been implemented in a computer program. The final calculations will be made when the clinical trial is finished and when we have estimates of the effect of Albumin. Here, we will report how the implemented model is designed to derive answers. Moreover, we demonstrate how the implementation will be used to explore the question of how accurate a test must be before it is cost-effective. This is a useful result, because it helps guide the research on biomarkers in its aim to determine how extensive the test must be (i.e. how many biomarkers are needed), before the test is accurate enough to be cost-effective.

By clicking 'plot population' the model will calculate the estimated number of individuals in the different states at different points in time from a starting cohort of 1000 individuals with cirrhosis in the benchmark model of 'before MICROB-PREDICT' (**Figure 3**).

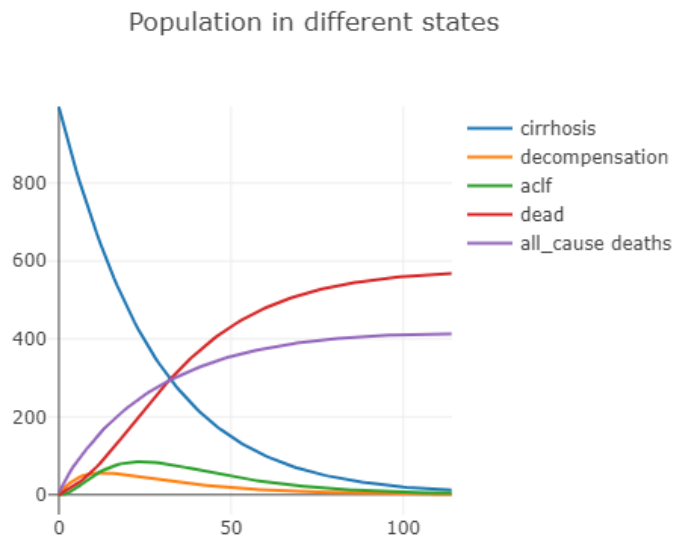

Figure 3: Number of individuals (y-axis) in different states at different months (x-axis) after being diagnosed with cirrhosis. This is implemented in the app automatically and appears after pressing the 'plot population' button.

Based on the number of people in the different states, and the costs/utility of these states, we can calculate the total lifetime cost and utility for all the individuals in the simulation. Taking the average of these, in turn, gives the overall expected cost and utility for a cirrhosis patient in the benchmark case. We can do this for different values (drawn from the probability distributions that are specified in the model). The results of 1000 such draws and calculations of expected costs and utilities from these draws are illustrated in **Figure 4**. This figure visualizes the degree of uncertainty in the conclusion.

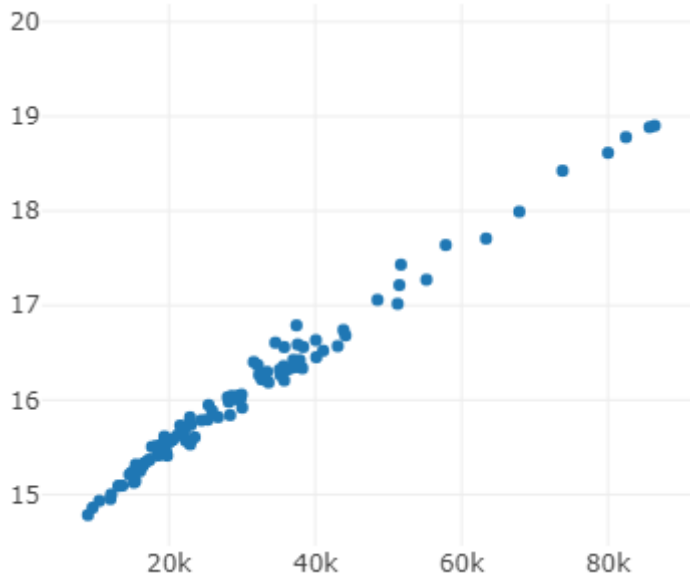

Figure 4: The costs (x-axis) and quality adjusted life years produced (y-axis) for different parameter values in the MICROB-PREDICT cirrhosis model (benchmark). Implemented by pressing 'Simulate' in the app.

This process is repeated for all the end nodes in the decision tree, which, in turn, allows us to estimate the expected cost and quality adjusted life years associated with each end node, the incremental cost-effectiveness ratio when comparing the results with a diagnostic test (of varying sensitivity and specificity) with the benchmark of no such test and the uncertainty/confidence interval by calculating the results for different parameter values.

9. Acknowledgement and Disclaimer

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825694.

This report reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.