

Using Design Thinking for Explainable AI: A Case Study Predicting the Start of the Palliative Phase in Patients with COPD or Heart Failure

Iris Heerlien¹, Jeroen Linssen¹, Lorenzo Gatti², Maya Sappelli³, Betsie van Gaal³, Richard Evering¹ and Noah Letwory⁴

¹Saxion University of Applied Sciences, M. H. Tromplaan 28, 7513 AB, Enschede, The Netherlands

²University of Twente, Drienerlolaan 5, 7522 NB, Enschede, The Netherlands

³HAN University of Applied Sciences, Kapittelweg 33, 6525 EN, Nijmegen, The Netherlands

⁴Ecare, Capitoool 11, 7521 PL, Enschede, The Netherlands

Abstract

The workload in the healthcare sector is increasing, requiring the need for innovative solutions. One such solution is for AI to assist in clinical decision-making by extracting information from patient's records. To ensure healthcare professionals stay in the lead, the reasoning of the AI should be transparent, creating the need for explainable AI (XAI). As this XAI representation should fit the users' needs and workflows, the user needs to be included in the design process. This research focuses on a case study using the Design Thinking method for generating an XAI representation for predicting the start of the palliative phase in patients with chronic obstructive pulmonary disease (COPD) or heart failure. This paper presents knowledge about and experiences with the design practices used, focusing on the ideation, prototype, and test phases. This contributes to the understanding of the needed design process to design XAI representations in the healthcare sector.

Keywords

Explainable AI, Design Thinking, Palliative Care, COPD, Heart Failure

1. Introduction

The workload in the healthcare sector is increasing, requiring the need for the adoption of innovative solutions [1, 2]. Several innovations are upcoming to reduce the workload, such as remote care reducing the need for hospital visits or eyedrip glasses used by the patient to drip their eyes without help of a healthcare professional [3]. Another innovation that could reduce the workload is the usage of Artificial Intelligence (AI). AI can be used to extract information from sources such as patient's records written by healthcare professionals, which can assist in e.g., clinical decision-making. However, it is crucial to ensure healthcare professionals maintain control and responsibility, and the risks of AI system errors leading to patient harm is minimized [4].

Explainable AI (XAI) is a solution to provide transparency in the AI reasoning [5]. To ensure the healthcare professional understands the AI reasoning and is able to use this in their daily job, the XAI representations should fit the needs and process of the healthcare professional.

However, a large part of the research performed in the area of XAI focuses on the technical perspective instead of the user perspective. A literature analysis showed that less than 1% of the publications validated their work with users comparing the literature on XAI [6]. Not validating findings and decisions with users could lead to XAI systems that the user cannot understand, resulting in wrong implications and eventually wrong decisions. To overcome this, the user should inform the XAI design and development cycle.

EXPLIMED 2025 - Second Workshop on Explainable Artificial Intelligence for the Medical Domain - 25–30 October 2025, Bologna, Italy

✉ i.r.heerlien@saxion.nl (I. Heerlien); j.m.linssen@saxion.nl (J. Linssen); l.gatti@utwente.nl (L. Gatti); maya.sappelli@han.nl (M. Sappelli); betsie.vangaal@han.nl (B. v. Gaal); r.m.h.evering@saxion.nl (R. Evering); noah.letwory@visma.com (N. Letwory)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

This study proposes the use of the Design Thinking methodology [7] to design the XAI representation. By an XAI representation we mean a visualization or narrative representing the reasoning behind an advice given by an AI. The Design Thinking methodology is inherently user-centered, which fits the goal of including users in the design and development cycle.

In this case study, we present insights about the used design practices for designing XAI representations in the healthcare sector. As this paper is part of a bigger research project in which the focus of other project partners is on the empathize and define phases, this paper focuses on the ideate, prototype, and test phases. The research question is as follows: ‘What should the design process of designing an explainable AI for the recognition of the palliative phase of patients with COPD or heart failure entail?’

This research contributes to understanding the needed design process to design XAI representations, and a start of creating a standardized process which can be followed by designers and developers when developing XAI representations in this sector.

1.1. Case

We conducted a case study to evaluate the use of the Design Thinking method for designing XAI representations, focusing on the timely recognition of the palliative phase for patients suffering from chronic conditions, such as chronic obstructive pulmonary disease (COPD) and heart failure. The target audience was healthcare professionals working in the home care sector in the Netherlands.

During the palliative phase, the focus of care shifts from curative treatments to symptom management and providing comfort. In patients with COPD and heart failure, this phase is typically identified when the patient’s expected remaining lifespan is approximately twelve months [8]. However, determining the onset of the palliative phase is particularly challenging, often resulting in delayed recognition and suboptimal care provision. Additionally, in the Netherlands approximately 600,000 individuals suffer from COPD, making it the sixth leading cause of mortality, with projections from the World Health Organization (WHO) indicating that it may become the third leading cause of death globally by 2030 [9]. Approximately 240,000 individuals in the Netherlands suffer from heart failure [10], highlighting the significant impact of these conditions on public health.

Healthcare professionals emphasize the importance of early recognition of the palliative phase, as late identification –sometimes occurring only shortly before death [11, 12] –adversely affects both quality of care and quality of life. Timely recognition enables healthcare professionals to initiate palliative care interventions and engage in Advanced Care Planning. This includes discussions with patients and their families regarding end-of-life care goals, treatment preferences, and support measures [13].

Several factors contribute to the difficulty of recognizing the palliative phase. Firstly, healthcare professionals primarily focus on curative treatment. Secondly, their assessments are often centered on isolated patient symptoms rather than holistic indicators of condition progression. Thirdly, there is limited awareness and utilization of tools designed to facilitate palliative phase recognition. Finally, the transition to the palliative phase is defined by a combination of subtle clinical indicators and is not a clear-cut criterion [9].

AI could be of use by generating alerts when recognizing these subtle changes in a patient’s condition over time. By incorporating XAI, the reasoning behind these alerts should provide the necessary insights to evaluate why this alert was generated and to evaluate whether this patient indeed has entered the palliative phase. This approach has the potential to increase the accuracy of palliative phase initiation, resulting in the ability to provide patient’s palliative care when needed. By including the users of the system the XAI will be part of, the representations will be more understandable and inform them in a suitable way.

2. Related work

In recent years, the use of AI in palliative care has been increasing, with promising results [14]. Several models have already been created to predict 2-year, 1-year, 6-months and 3-days mortality, survival estimation and 1-year frailty [15, 16, 17, 18]. Zhang et al. [19] have created a 1-year mortality prediction

for patients suffering from chronic conditions, resulting in a Receiver Operating Characteristic (ROC) curve of 0.73. The detection of palliative status has been done by Sandham et al. [20], resulting in ROC scores between 0.6 and 0.724.

In addition to these studies, which demonstrate that AI could provide valuable information in the healthcare sector, another line of research has focused on the role, usage and effects of XAI in the same domain.

However, Chen et al. [21] performed a systematic review to understand the inclusion of users in the design process and concluded that no study from their review of papers between 2012 and 2021 reported a formative user research to create XAI systems in medical image analysis. To overcome this, they introduce the INTRPRT guideline which states that users are incorporated in the steps 'formative research' and 'ideation', after which the input gathered during these steps are incorporated in the development phase and validated by users.

The lack of inclusion of end users improved after 2021. One example is by Blanes-Selva et al. [22], who used XAI in a clinical decision support system, resulting in a good user experience score and acceptable usability. They included users in the validation process of the XAI system by performing a task test followed by the System Usability Scale (SUS) and the User Experience Questionnaire - Short Version (UEQ-S) questionnaires. Shulha et al. [23] used the Design Thinking method to create XAI representations. They explored if using a design thinking approach to create a decision support tool based on XAI techniques would increase the clinical implementation. Multiple research activities per Design Thinking phase were performed, such as focus groups, a rapid review and a scoping review in the empathize phase, and paper prototype testing before the development of a working prototype in the prototype phase. A framework to explore clinician trust in AI was used in the ideate, prototype and test phases. LIME (local interpretable model-agnostic explanations) [24] was used to interpret the model output. The Design Thinking approach was seen as valuable. Additionally, they state that incorporating the chosen conceptual frameworks, the Non-adoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability (NASSS) framework [25] and the framework for clinician trust in machine learning [26], increased the robustness of the collaborative tool design.

Panigutti et al. [27] created a user interface for a clinical decision support system using XAI techniques. An iterative design approach was used in which users (healthcare providers) were asked to validate the prototype to understand the impact of explanations on the users' trust after which these insights were used to redesign the interface. An heuristic evaluation was performed comparing the two interfaces using the Nielsen and Norman Usability Heuristics [28], resulting in a preference for the new interface. They concluded that explanations increased the users' trust in the system. Since perceived usefulness is dependent on the correctness of the prediction by the algorithm, only correct suggestions were included in the evaluation.

Zhuang et al. [29] used XAI techniques in predicting mortality risk. They created an XAI model based on patient's records that predicted the 365-mortality risk for patients with advanced cancer. Shapley Additive Explanations (SHAP) [30] values were used to explain the model outputs to increase trust and adoption in a clinical setting. The domain experts were involved in the feature selection process, resulting in recognizable features in the visualization. The XAI visualization itself is not created and validated using a user-centered design method.

From this overview we learn that involving users in the design process of XAI solutions deliver positive results. However, there is no research in which the palliative status is detected, and a user-centered XAI is designed. In this research we combine those insights into a case study in which we aim for detecting the palliative phase while including users to create an XAI solution that fits them best.

We show concrete examples of how the methods for the Design Thinking phases ideate, prototype, and test can be applied when designing an XAI system for a medical system. We believe the learnings from our use case are useful to other practitioners that are investigating a similar task to improve their processes, and to researchers who are not familiar with design thinking to learn the kind of insights that it can lead to.

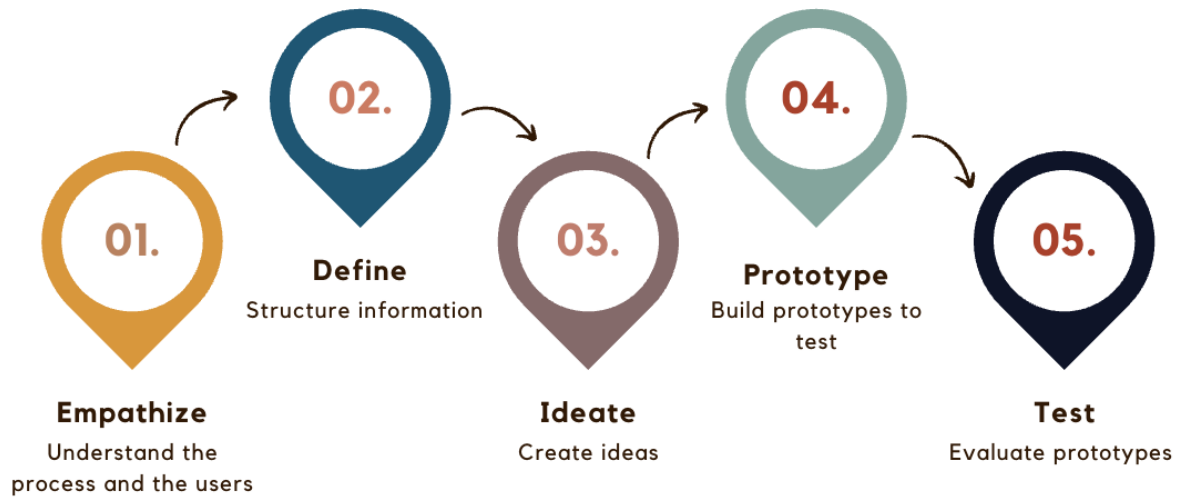


Figure 1: The Design Thinking methodology.

3. Method

This research uses the Design Thinking methodology (see Figure 1) to design the XAI in a user-centered way. The first phase of the Design Thinking method is the empathize phase, in which the task is to understand the users and their needs. During the second phase, the define phase, the information from the empathize phase is combined to define the user groups and workflows of which the XAI will be part.

These phases are crucial to identify who are the stakeholders of the system, their needs, and how they currently address them. In this work, however, we will only briefly touch upon them, as they were performed by different project partners and our research focused on the later stages of the Design Thinking method.

During the ideate phase, co-creation methods were used to create ideas. Based on this, a prototype was created and evaluated during the test phase.

3.1. Developed AI component

The XAI representation we set out to design was for an AI system that had already been developed as part of the research project. Specifically, the system is a classifier that was trained to predict whether the patient is entering the palliative phase, using the data in the Electronic Health Record (EHR). The database contains all information about a patient, from personal data to results of medical exams and reports of healthcare professionals.

The Random Forest classifier uses as features static data (such as the patient's age and gender) and dynamic signals coming from the reports written by the healthcare professionals when visiting the patients in the previous 30 days. The text of these medical reports is processed with bag-of-word techniques, based on the count of words pertaining to important "dimensions" (i.e. the physical, social, psychological and spiritual dimensions) and indicators of medical events (e.g. a visit to the doctor or being admitted to an intensive care facility), similar to what the Linguistic Inquiry of Word Count (LIWC) text analysis tool uses [31]. The lexicons for these dimensions and indicators were built by identifying keywords (e.g. 'pain', 'lonely', 'anger', 'isolation', 'fear') through a literature research, but also with interviews and focus groups with healthcare professionals and patients. The lexicons were then expanded by including synonyms and related forms of the original words. The final list of words was validated with healthcare professionals.

While the classifier can predict with a recall of 0.88 that a patient is entering the palliative phase, the actual decision should remain in the hand of medical professionals. This is where XAI comes in. Using

XAI techniques, the professional can evaluate why the classifier came up with the advice and is able to decide how to incorporate this advice in their decision making.

3.2. Empathize and Define phases

To understand the users and the process the XAI will be part of, two focus groups were held with two different teams from two different home care organizations in the Netherlands.¹ In total, eleven participants joined the focus groups. All participants were female, except for the general practitioner. During the first focus groups, four district nurses, of which one was specialized in palliative care and one was trained to be a specialist in palliative care, and one nurse with a minor in palliative care were present. During the second focusgroup, one general practitioner, two palliative nurses, and two registered nurses were present. The age range was not restricted, resulting in participants from an age range that covered the entire user population. The focus group was kicked off by an introductory presentation defining the goal and necessity of the representation and the way it is represented to them.

After this, the group was divided into five pairs after which they received a worksheet. One participant decided to work individually. The participants were first asked to draw or describe the visualizations and graphs they already know from the Electronic Health Record (EHR) system they use in their daily work, such as bar graphs and pie charts. To inspire them and to ensure the assignment was understood, two examples were given during the instruction phase: a bar chart and a decision tree. They were also asked to write down what they liked and disliked about the representations they know. One of the goals of this task was to get to know the healthcare professionals and what they are used to. Another goal was to encourage them to create a critical attitude.

3.3. Ideate phase

The same participants from the first focus groups in the Empathize phase were involved in the focus group of this phase. The task of the participants was to design their ideal process and representation of the reasoning of the AI. The participants were already acquainted with the goal of the AI, i.e., classifying patients entering the palliative phase, since they were present during earlier focus groups of the research project. This was enough information for them to understand what kind of advice the AI would give. The questions we asked the participants to incorporate in their designs were: what information do users need to understand the AI prediction? Who should it communicate to? What should the action of the user be following the representation? The sketches and descriptions that were created by them were discussed during the focus groups and used in the next phase by the research and development team to create a prototype.

3.4. Prototype phase

The information from the focus groups was discussed in the research team, consisting of representatives from the healthcare and data science sector, and the development team responsible for the EHR system to decide what was feasible for a high-fidelity prototype. Additionally, decisions were made about what the prototype should include from the research team's perspective to ensure the professional remains aware of the limitations of the algorithmic output.

A low-fidelity design was created and discussed with three healthcare professionals during a general meeting of the project. After this, a high-fidelity prototype was implemented which was evaluated during the test phase.

3.5. Test phase

The prototype was evaluated during focus groups with five district teams of two home care organizations, two of one organization and three of another organization. They used the prototype for a period of

¹All the participants involved signed a consent form for this and all other user interactions in this research.

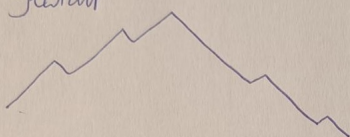
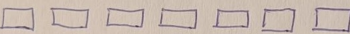

Stap 1	Stap 2
gewicht 	- helder - snel zichtbaar hoe 't verandert
defecatiepatroon visuele schaal 	- niet helder - vraagt meer aandacht om er conclusies aan te verbinden
keonts 	duidelijk
pols	

Figure 2: This is an example of an output of the Empathize phase. They state three examples of visualizations: a line chart, a rating scale and a bar chart. In the right column they state that the line chart is 'clear' and 'shows the trend very well', the scale is 'hard to read', and the bar chart is considered as 'clear'.

three months, during which the focus groups were held to understand if it fit the users and if they were able to use the tool during their daily job.

4. Results and Discussion

Information gathered during the focus groups during the empathize and ideate phases were used by the research team to create a prototype, which was evaluated by the healthcare professionals. The following sections explain the results and our discussions of these per Design Thinking phase.

4.1. Empathize and Define phases

When asked to draw the representations they know, healthcare professionals presented pie charts, bar charts, line charts, rating scales, speedometers, tables, smileys, traffic lights, scores, and percentages.

They stated that the *bar chart* was considered the most clear and easy to interpret, especially for identifying stability and peaks at a glance. The *line chart* was also seen as clear, particularly for displaying a single symptom, and is useful when combined with a bar chart for comparison. The *pie chart* was experienced as providing a lot of information but becomes unclear when more than three categories were included, making it less preferred. The *rating scale* and *scores* were perceived as unclear because they experienced that it requires more cognitive effort. The *speedometer* is visually clear but difficult to interpret. A *table* is useful when more detailed information on a single symptom is needed. *Smileys* were considered too simplistic and lacking nuance, and *percentages* were not preferred, though the reason is unclear. An example of the output of this assignment can be seen in Figure 2. Overall, the bar chart or line chart were experienced as most clear and easy to interpret. These results were taken into account during the Ideate phase.

The way the healthcare professionals envisioned the general process is shown in Figure 3. They expect that a notification indicating that the AI detected a patient entering the palliative phase will appear in the patients list in the electronic health record system. By clicking on this notification, they expect a pop up to be shown including the dashboard or a link to the dashboard. This dashboard should

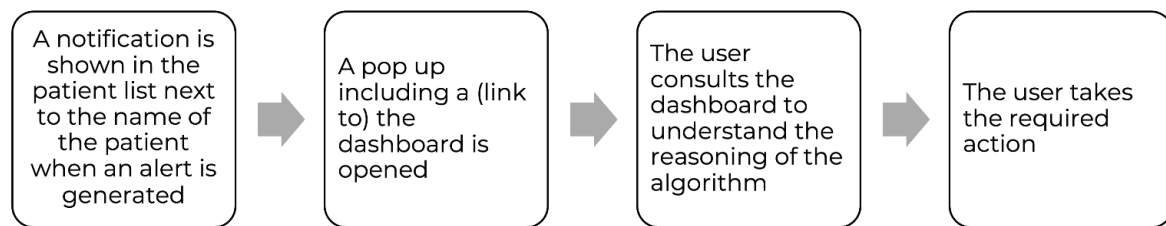


Figure 3: The envisioned process when a patient is seen as palliative by the system.

show the XAI representation, allowing the user to understand the reasoning of the AI and therefore *why* someone is likely entering the palliative phase, based on the patient data. The district nurse can then assess this, use it alongside their knowledge and experience with the patient, and decide whether they agree or disagree before taking the necessary action.

This envisioned process was seen differently by the participants and discussed during the focus groups. The biggest questions were what the action of the user should be, and who should be included in this process and thus be able to see this dashboard. People from one healthcare organization believed, for example, that the family of the patient should be able to see it as well, since they think the family should be involved; the other healthcare organization, instead, opposes this to protect the family. Also, the action following the representations differed per group. Some participants believed the general practitioner should be signaled directly, while others believed the nurses specialized in palliative care should be included first for a check before the general practitioner was involved.

This process is used in the Ideate phase to brainstorm about a solution fitting this way of working. The action of the user stays unclear from these focus groups. Therefore, one option is chosen in the Prototype phase which is evaluated. Based on this, changes could be made in an iteration for refinement.

Learnings

During this phase, the participants were asked to draw the visualizations they know and their opinion about this. This was experienced as a valuable method as this triggered the participants. In our case, this phase received less attention as this has already been covered by other research activities of the same research project. In the general case, however, the focus should be on understanding the users and the current process, e.g., by persona creation, empathy mapping, and creating a user journey.

4.2. Ideate phase

The representations that the participants came up with when asking to create their ideal representation were different. However, they all focus on a dashboard format. Eight different options were created, of which one is shown in Figure 4. Two representations showed the general information at the top of the dashboard. Three representations presented a bar chart to show a top 5, one representation presented a bar chart or a pie chart, and two representations did not present a specific graph. The AI is trained using words categorized in four dimensions, the physical, social, psychological and spiritual dimension. Three representations showed the dimensions that was triggered on, three representations showed the words used instead of the dimensions, and two representations were unclear in this. Five representations showed a drill-down to create another level with more detailed information. Two representations presented a level to show when the words were used in the patient's records and another level to show the progress of the words over time. Two representations used this drill down to create a level of words next to a level showing a graph on a dimension level. One representation presented a level to show the progress of the words used over time.

The actions that should be suggested by the system differ as well. One suggested action was to talk to the patient, general practitioner and the nurse specialized in palliative care. Another suggestion was

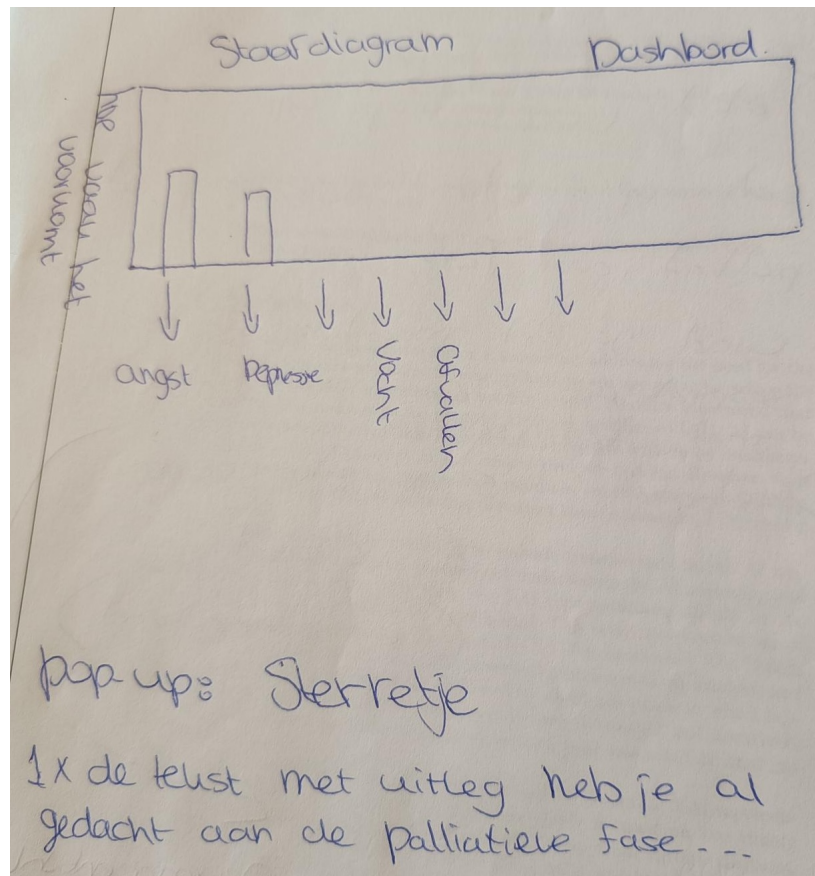


Figure 4: One example of the output of the Ideate phase. In this example, the participant designed a bar chart showing the most important words and how many times they occur in the reports. They want to be informed with a star next to the name of the patient in the EHR system. When one clicks on this star, a pop-up is shown. The action they want to be able to follow from this dashboard is by asking 'have you already thought about the palliative phase?'

a yes/no question. The suggestion to first talk to the specialized nurse and after this with the general practitioner was seen twice. Another suggestion was to directly talk to the general practitioner and send an email in case of questions to the specialized nurse. The suggestion to show a text was seen twice, once to show the text 'have you already thought about the palliative phase?' and once to show the text 'based on the data above, this patient could possibly be marked as palliative. The advice is to start a conversation about this with the patient. For questions, send an e-mail to the nurses specialized in palliative care.' This input is the start of the prototype and test phases. As the action is not entirely clear, one option is chosen to evaluate. Based on this, changes could be made in a refinement loop.

Learnings

During the ideate phase the participants were asked to sketch their preferred visualizations and process. This was done in one assignment. In hindsight, we would advise to separate this by first looking at the process, then what information is needed, and lastly how this should be communicated. We expect this would give more in-depth results and also shows what users want to understand about the reasoning of an AI system instead of only how it should look like. Although this is expected to be part of the chosen representation as this is what we asked the participants to take into account, making it more explicit would help in the discussion and the creation of the prototype.

Additionally, an introductory presentation was shown with two example representations to inspire them. The decision was made to use standard representations to not steer them too much. However, one of the representations shown during the presentation, a bar chart, was often used by the participants

when sketching their preferred visualizations. It is difficult to find a trade off between not inspiring them at all, possibly resulting in a misunderstanding of the assignment, and inspiring them too much, as could have happened here. It is also hard to measure this, as it is unclear what has happened when the representations were not shown. Additional research is needed to improve on this.

Next to this, the representations sketched were merely dashboards. Although this may seem the most logical way and the use of SHAP and Local Interpretable Model-agnostic Explanations (LIME) representations are also seen often in related research [23, 29], it could also have been because of the steering in the introductory presentation. This should be taken into account when performing additional research. Separating the different stages of which information is necessary and how to represent this is expected to help in this as well.

Another takeaway is that the conversation is valuable. Starting by asking to sketch the process and visualization helps everyone think for themselves, which is a good basis for the conversation. The conversation shows the points that are most important for the participants and shows the differences and reasoning behind the choices made. This is insightful for the researcher and shows the things that should get extra attention in the later phases of the Design Thinking cycle.

4.3. Prototype phase

The results from the focus groups were discussed in the research and development team to see what was feasible. From the ideation phase we learned an improved workflow and representation. Based on these results, a low-fidelity design was created (see Figure 5).

One of the focus points in these discussions was how to inform the healthcare professionals in a way that the professional remains aware of the limitations of the algorithmic decision. The goal is to generate a 'digital colleague' that gives them the insights from the data, but ensuring they will make the decision themselves. Therefore, the decision was made to not link the representation back to where in the patient's records the words were that triggered the algorithm, to make sure patient's records that were as important but did not include those words were not skipped by the user when looking back.

In this design, a bar chart is used as this was experienced as the most clear representation. Additionally, a drill-down creating another level with more detailed information was seen as helpful. This is included in the design. The asked for information is placed on top of the design and one type of action is placed below the design. This design was discussed with three healthcare professionals during a general meeting of the project. They reacted enthusiastic and saw the information that they gave during the focus group back in the design. Due to technical boundaries it was not possible to implement the intended workflow in the prototype which was used in the test phase of the system. Therefore, the focus of this prototype was solely on the visualizations.

For the evaluation of the prototype, the decision was made to show the words and focus areas which were most important for the prediction, as detected by SHAP [30] (Figure 6). This technique shows the features, in this case the words, that had the most impact on the final advice. At another tab, the healthcare professional could read more values than only the top 5 (Figure 7). Important to note is that also the absence of words and focus areas could be seen as an indicator (e.g., the absence of the word 'happy' could be an indicator), showing these in the representation as well.

Learnings

In the prototype phase we focused not only on what the healthcare professionals needed, but also on how to make sure that the representations convey the right information to make it impossible to conclude wrong things from it. For example, we decided to not show where in the reports the words that triggered the algorithm are located, to encourage them to base the decision on their experience as well. This avoids over-reliance on the evidence provided by the algorithm and the risk of neglecting their intuition and the reports that could paint a different picture. The decision not to present where in the report these words are located was against the wish of the healthcare professionals, but we believe that, especially in areas such as healthcare, decision-makers should be in the lead and given tools to



Figure 5: The low-fidelity design based on the information of the focus group. The top shows the general information of the patient, followed by an action (“Watch out! Based on the data below, this patient could be considered palliative. It’s recommended to discuss this with them. If you have questions, send an email to ‘kerntaakhouders@organisatie.nl’”) and a graph on the level of the healthcare domains (physical, spiritual, psychological). In the lower part of the dashboard, a bar chart is shown stating which words in these domains triggered the algorithm most (“tiredness”, “weight loss”, “arm pain”). Next to the bar chart, a line graph shows the development over time of the mentions of the selected word (“weight loss”) in the reports.

improve their reasoning instead of automatically trusting the decision of the algorithm. Future work should incorporate this aspect even more to understand how to include it in the applications without disturbing the workflow. This will also make it possible to evaluate the different workflows mentioned during the ideate phase.

4.4. Test phase

The prototype was evaluated in five teams providing care in different neighborhoods. Focus groups were held to understand what went well, what went wrong, and how it could be improved. During those focus group the entire tool was evaluated; the results (AI), and the dashboard (XAI).

The prototype evaluated was a dashboard which was not integrated directly in the EHR system itself, but accessible via a link in the system. It included data from real patients to make it as real as possible for the healthcare professionals using it. As the dashboard is not integrated in the system, solely the representations were evaluated and not the workflow as it is not representative for this.

Their opinions were that they understood why the dashboard, the XAI component, is useful. They understood that being able to read why someone was signaled as entering the palliative phase will help them in agreeing with this or not. In addition, they stated that it could help them in the communication with the general practitioner, and could even help them to understand on which topics they should report more to be able to generate better data for the AI. They also stated that this tool will improve the



Figure 6: The first page of the prototype of the XAI representation, showing the top five features for the advice created by the algorithm for this patient. The feature importances are calculated using SHAP. On the x-axis one can read the words that the algorithm triggered most on. From left to right: ‘elastic stocking’, ‘gender’, ‘diabetes’, ‘oral care’, ‘eye drops’. On the y-axis is the score showing the impact of the word.

care itself as the palliative phase could be recognized on time.

However, there were some limitations mentioned. The results showed in the prototype felt confusing to them, making it hard to understand and trust the system. For example, if someone did not use support stockings, this was shown as a sign that someone would enter the palliative phase.

Besides, there was no drill down functionality in the prototype. This was stated in the ideate sessions and was also something they were missing in the prototype. The second page (Figure 7) shows more features than only the top five, but the users experienced it as hard to read.

A requirement that came up during the sessions was that they also wanted to know when someone was not classified as palliative. It was expected this would be helpful to also understand when someone is not entering the palliative phase based on the data, while they would expect it from their perspective.

Finally, they stated it would be useful to see in what patient’s records the words were used most often. The research team decided not to implement this, as was explained in Section 4.3.

Learnings

The prototype tested in the test phase was not ideal; due to some technical boundaries, the workflow and representation evaluated was not as it was supposed to be. This was experienced as disturbing, making it hard to evaluate the usefulness of the type of XAI representation and the workflow the XAI is part of. We learned from this that a lower fidelity prototype is expected to be more useful in testing than a prototype of a higher fidelity but with limitations. In the next steps of this research, the focus

Uitgevoerde rapportages		Top 5		Uitkomsten		Exportoverzicht	
ClientID	Q	Datum_verversing_model	Q	Domein	Q	Score	Q
35306		17-2-2025		Sociale dimensie		0,02	
35306		24-2-2025		Geslacht		0,02	
35306		10-3-2025		Elastische kous		0,01	
35306		10-3-2025		Geslacht		0,01	
35306		17-2-2025		Elastische kous		0,01	
35306		24-2-2025		Elastische kous		0,01	
35306		24-2-2025		#Aandachtsgebieden		0,01	
35306		24-2-2025		Diabetes		0,01	
35306		17-2-2025		Fysieke dimensie		0,01	
35306		10-3-2025		Diabetes		0,01	
35306		10-3-2025		Risico inventarisatie		0,01	
35306		17-2-2025		Diabetes		0,01	
35306		10-3-2025		#Aandachtsgebieden		0,01	
35306		10-3-2025		Oogdruppelen		0,01	
35306		24-2-2025		Mondzorg		0,00	
35306		10-3-2025		Mondzorg		0,00	
35306		17-2-2025		Mondzorg		0,00	
35306		10-3-2025		Dementie		0,00	
35306		24-2-2025		Oogdruppelen		0,00	
35306		17-2-2025		Oogdruppelen		0,00	
35306		10-3-2025		Wonden		0,00	
35306		24-2-2025		Stomazorg		0,00	
35306		10-3-2025		Stomazorg		0,00	

Figure 7: The prototype of the XAI representation, showing the list of features ordered on importance that triggered the algorithm. The columns from left to right show the client-id, the last date the model is retrained, the words that triggered the algorithm, and the score. The words from top to bottom: ‘social dimension’, ‘gender’, ‘elastic stocking’, ‘number of focus areas’, ‘diabetes’, ‘risk inventory’, ‘diabetes’, ‘number of focus areas’, ‘eye drops’, three times ‘oral care’, ‘dementia’, two times ‘eye drops’, ‘wounds’, two times ‘stoma care’.

should be on taking a step back and using the low-fidelity prototype in testing or use paper prototyping as has been done by [23], before implementing it.

During this phase, the usability of the tool is evaluated. However, it is not evaluated whether the users can interpret the results correctly. Previous work [32] already identified the challenges that data scientists can have when interpreting the output of explainable tools like SHAP, but this should be expanded with an evaluation centered on the healthcare professionals.

The participants stated that using and evaluating the tool makes them more aware of how to report such that better data is created for the AI. Although this seems a positive side effect, it has implications. At first, the data the AI is fed changes, which will change the outcome of the algorithm. This should be accounted for when maintaining the algorithm. In addition, it changes the way of working of the healthcare professionals. Future work should show if this has a negative implication on the provided care. A second iteration is not included in this research. Therefore, we do not draw conclusions about the impact and level of satisfaction. Work on an improved version is to be included in future research.

4.5. General learnings

Overall, the methods used in the different steps of the Design Thinking method were experienced as valuable when designing XAI for end users. The goals of using the Design Thinking method to involve the users in XAI representation design are to being able to communicate the reasoning of the AI to the users in a way that fits their way of working and preferences and to ensure the representations are interpreted correctly by the users. This will enhance the chances the system will be used by them.

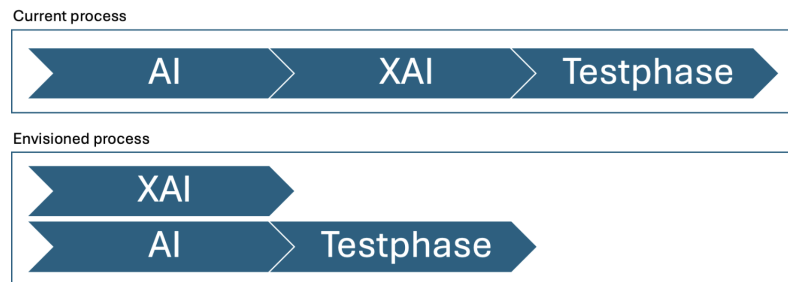


Figure 8: The current and envisioned process combining AI and XAI design. By combining the AI and XAI process, the Empathize to Prototype phases of the Design Thinking method could inspire the AI development. Also the other way around in which the possibilities from an AI perspective could inspire the methods used in the XAI process. The Test phase of the Design Thinking process could evaluate the entire application.

Additionally, one can evaluate thoroughly if the XAI system is used next to their own knowledge and experiences instead of taking over the advice without a critical attitude. This will make sure that the professional stays in the lead and the responsibility is not shifted to the AI.

The goal of performing the Design Thinking method in this project is to create a system that fits the user's workflow and informs the user in a way that is understood correctly by them. Although a start is made in the focus groups and prototype sessions, an iteration cycle is needed to meet this goal. However, although the AI development is in a further stage, the results from the iteration cycle to create this correct system could also affect the AI algorithm and design. To optimize the design and development cycle of the AI and XAI part, we propose a shift in the process, as explained in the next section.

4.6. Envisioned XAI design process

In the current case study, the XAI design process initiated only after the AI was already developed. We advise for similar future projects to start the design phase together with the AI development cycle as shown in Figure 8. We hypothesize this will increase the quality of the design phase, as there is time to iteratively design and evaluate the XAI representation thoroughly separately from the AI, before using it in the overall test phase. This should overcome the limitations in this research, as a refinement loop is possible which leads to understanding the users better and validating the design and expected workflow multiple times in different stages of fidelity before implementing it. In addition, this way both processes can also inform each other, e.g., the XAI design process could steer the methods used in the AI development cycle. This could for example overcome the issue of the healthcare professionals experiencing confusing information in the dashboard during the overall test phase (see Section 4.4 for more details on this). After evaluating the XAI and AI separately, both parts could be combined in a final test phase testing the entire system.

5. Conclusion

In this research we present a case study in which we use the Design Thinking methodology to create a user-centered XAI solution for detecting the palliative phase. A first iteration was performed in which the users were included in the empathize, ideate and test phases. The knowledge about the users and workflows, the users' knowledge and preferences were experienced as valuable in designing and validating the interface.

To answer the research question 'What should the design process of designing an explainable AI for the recognition of the palliative phase of patients with COPD or heart failure entail?', we observed that co-designing the XAI solution by involving the users in the Empathize, Ideate, and Test steps of the Design Thinking approach was valuable. Our results show that starting with understanding the process and learning which visualizations are known and used by the end-users in the Empathize step,

followed by a brainstorm session to visualize the ideal process and representation, helps in defining the prototype. Testing the prototype with end-users helps in understanding what should be improved. These methods gave useful results although some improvements could be made. The most important optimization is starting the XAI design process together with the AI development process to enhance the communication between the two processes. This way, there is more time to iteratively develop the XAI part, and the XAI and AI development will inform each other, resulting in a more complete prototype to evaluate. This will eventually lead to improved XAI designs which will enhance the usage of these solutions in the healthcare sector.

6. Acknowledgments

This publication is part of the project ‘Technology marks the palliative phase’ (Telemap)² which is financed by a RAAK Public subsidy.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

- [1] Ministerie van Volksgezondheid, Inspectie Gezondheidszorg en Jeugd, Personeelstekorten in de zorg., 2023. URL: <https://www.igj.nl/onderwerpen/personeelstekort>, accessed on 23-03-2025.
- [2] Dienst Rijksoverheid, Integraal Zorgakkoord: ‘Samen werken aan gezonde zorg’, 2022. URL: <https://www.rijksoverheid.nl/documenten/rapporten/2022/09/16/integraal-zorgakkoord-samen-werken-aan-gezonde-zorg>, accessed on 12-09-2025.
- [3] Coöperatie VGZ, 2022. URL: <https://www.cooperatievgz.nl/cooperatie-vgz/zorg/personeelstekort-zorg>, accessed on 01-02-2025.
- [4] W. Nicholson Price II, Risks and remedies for artificial intelligence in health care, 2019. URL: <https://www.brookings.edu/articles/risks-and-remedies-for-artificial-intelligence-in-health-care/>, accessed on 12-09-2025.
- [5] Z. Sadeghi, R. Alizadehsani, M. A. CIFCI, S. Kausar, R. Rehman, P. Mahanta, P. K. Bora, A. Almasri, R. S. Alkhawaldeh, S. Hussain, B. Alatas, A. Shoeibi, H. Moosaei, M. Hladík, S. Nahavandi, P. M. Pardalos, A review of Explainable Artificial Intelligence in healthcare, *Computers and Electrical Engineering* 118 (2024). doi:10.1016/j.compeleceng.2024.109370.
- [6] A. Suh, I. Hurley, N. Smith, H. C. Siu, Fewer than 1% of explainable AI papers validate explainability with humans, 2025. *arXiv:2503.16507*.
- [7] Nielsen and Norman Group, Design thinking 101, 2025. URL: <https://www.nngroup.com/articles/design-thinking/>, accessed on 01-02-2025.
- [8] Palliatieve Zorg Nederland, Kwaliteitskader palliatieve zorg Nederland, 2017. URL: <https://palliaweb.nl/zorgpraktijk/kwaliteitskader-palliatieve-zorg-nederland>, accessed on 12-09-2025.
- [9] Longfonds, COPD, 2025. URL: <https://www.longfonds.nl/longziekten/copd>, accessed on 01-02-2025.
- [10] Hartstichting, Cijfers hart- en vaatziekten, 2025. URL: <https://www.hartstichting.nl/hart-en-vaatziekten/feiten-en-cijfers-hart-en-vaatziekten>, accessed on 01-02-2025.
- [11] A. Francke, M. Meurs, A. van der Plas, H. Voss, Inventarisatie van advance care planning. ZonMw-projecten, methoden, uitkomsten en geleerde lessen over gebruik, implementatie en borging, NIVEL (2020). URL: <https://www.nivel.nl/nl/publicatie/inventarisatie-van-advance-care-planning-zonmw-projecten-methoden-uitkomsten-en-geleerde>, accessed on 12-09-2025.

²<https://www.saxion.edu/business-and-research/research/smart-industry/ambient-intelligence/raak-publiek-telemap>

- [12] S. J. J. Claessen, M. A. Echteld, A. L. Francke, L. Van den Block, G. A. Donker, L. Deliens, Important treatment aims at the end of life: a nationwide study among GPs, *British Journal of General Practice* 62 (2012) e121–e126. doi:10.3399/bjgp12X625184.
- [13] J. A. C. Rietjens, R. L. Sudore, M. Connolly, J. J. van Delden, M. A. Drickamer, M. Droger, A. van der Heide, D. K. Heyland, D. Houttekier, D. J. A. Janssen, L. Orsi, S. Payne, J. Seymour, R. J. Jox, I. J. Korfage, Definition and recommendations for advance care planning: an international consensus supported by the european association for palliative care, *The Lancet Oncology* 18 (2017) e543–e551. doi:10.1016/S1470-2045(17)30582-X.
- [14] J. Bork-Zalewska, An overview of the role of artificial intelligence in palliative care: a quasi-systematic review, *Palliative Medicine in Practice* (2024) 2545–1359. doi:10.5603/pmp.103020.
- [15] V. Blanes-Selva, A. Doñate-Martínez, G. Linklater, J. Garcés-Ferrer, J. M. García-Gómez, Responsive and minimalist app based on explainable AI to assess palliative care needs during bedside consultations on older patients, *Sustainability* 13 (2021). doi:10.3390/su13179844.
- [16] V. Blanes-Selva, A. Doñate-Martínez, G. Linklater, J. M. García-Gómez, Complementary frailty and mortality prediction models on older patients as a tool for assessing palliative care needs, *Health Informatics Journal* 28 (2022). doi:10.1177/14604582221092592.
- [17] L. Wang, L. Sha, J. R. Lakin, J. Bynum, D. W. Bates, P. Hong, L. Zhou, Development and validation of a deep learning algorithm for mortality prediction in selecting patients with dementia for earlier palliative care interventions, *JAMA Network Open* 2 (2019). doi:10.1001/jamanetworkopen.2019.6972.
- [18] M. Mori, T. Yamaguchi, I. Maeda, Y. Hatano, T. Yamaguchi, K. Imai, A. Kikuchi, Y. Matsuda, K. Suzuki, S. Tsuneto, D. Hui, T. Morita, EASED collaborators, Diagnostic models for impending death in terminally ill cancer patients: a multicenter cohort study, *Cancer Medicine* 10 (2021) 7988–7995.
- [19] H. Zhang, Y. Li, W. McConnell, Predicting potential palliative care beneficiaries for health plans: a generalized machine learning pipeline, *Journal of Biomedical Informatics* 123 (2021). doi:10.1016/j.jbi.2021.103922.
- [20] M. H. Sandham, E. A. Hedgecock, R. J. Siegert, A. Narayanan, M. B. Hocaoglu, I. J. Higginson, Intelligent palliative care based on patient-reported outcome measures, *Journal of Pain and Symptom Management* 63 (2022) 747–757.
- [21] H. Chen, C. Gomez, C.-M. Huang, M. Unberath, Explainable medical imaging AI needs human-centered design: guidelines and evidence from a systematic review, *npj Digital Medicine* 5 (2022). doi:10.1038/s41746-022-00699-2.
- [22] V. Blanes-Selva, S. Asensio-Cuesta, A. Doñate-Martínez, F. Pereira Mesquita, J. M. García-Gómez, User-centred design of a clinical decision support system for palliative care: Insights from health-care professionals, *Digital Health* 9 (2023). doi:10.1177/20552076221150735.
- [23] M. Shulha, J. Hovdebo, V. D’Souza, F. Thibault, R. Harmouche, Integrating explainable machine learning in clinical decision support systems: Study involving a modified design thinking approach, *JMIR Formative Research* 8 (2024). doi:10.2196/50475.
- [24] M. Tulio Ribeiro, S. Singh, C. Guestrin, “Why should I trust you?": Explaining the predictions of any classifier, in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*, 2016, p. 1135–1144. doi:10.1145/2939672.2939778.
- [25] T. Greenhalgh, J. Wherton, C. Papoutsi, J. Lynch, G. Hughes, C. A’Court, S. Hinder, N. Fahy, R. Procter, S. Shaw, Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies, *Journal of Medical Internet Research* 19 (2017). doi:10.2196/jmir.8775.
- [26] S. Tonekaboni, S. Joshi, M. McCradden, A. Goldenberg, What clinicians want: Contextualizing explainable machine learning for clinical end use, in: *Proceedings of the 4th Machine Learning for Healthcare Conference (MLHC 2019)*, volume 106, 2019, pp. 359–380.
- [27] C. Panigutti, A. Beretta, D. Fadda, F. Giannotti, D. Pedreschi, A. Perotti, S. Rinzivillo, Co-design of human-centered, explainable AI for clinical decision support, *ACM Transactions on Interactive Intelligent Systems* 13 (2023). doi:10.1145/3587271.

- [28] Nielsen and Norman Group, 10 usability heuristics for user interface design, 2024. URL: <https://www.nngroup.com/articles/ten-usability-heuristics/>, accessed on 01-02-2025.
- [29] Q. Zhuang, A. Y. Zhang, R. S. T. Y. Cong, G. M. Yang, P. S. H. Neo, D. S. Tan, M. L. Chua, I. B. Tan, F. Y. Wong, M. Eng Hock Ong, S. Shao Wei Lam, N. Liu, Towards proactive palliative care in oncology: developing an explainable EHR-based machine learning model for mortality risk prediction, *BMC Palliative Care* 23 (2024). doi:10.1186/s12904-024-01457-9.
- [30] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*, 2017, p. 4765–4774.
- [31] Y. R. Tausczik, J. W. Pennebaker, The psychological meaning of words: LIWC and computerized text analysis methods, *Journal of Language and Social Psychology* 29 (2010) 24–54. doi:10.1177/0261927X09351676.
- [32] H. Kaur, H. Nori, S. Jenkins, R. Caruana, H. Wallach, J. Wortman Vaughan, Interpreting interpretability: Understanding data scientists' use of interpretability tools for machine learning, in: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*, 2020, p. 1–14. doi:10.1145/3313831.3376219.