A Double-Loop Patient-Oriented Learning Cycle for Therapy Decision-Making

Raphaël Ménard-Grenier

Thesis submitted to the University of Ottawa
in fulfillment of the requirements for the
Master of Science in Health Systems

Telfer School of Management
University of Ottawa

Table of Contents

Acknowledgment

I would like to say a special thank you to my supervisors, Lysanne L. and Antoine S. Their support, guidance and overall insights in their respective fields have made my journey an inspiring experience worth pursuing.

I would also like to thank my wonderful partner Amélie, for her patience and continuous support throughout the years.

Abstract

Therapy decision-making for patients with chronic diseases can be difficult. Such patients usually live with their illness(es) all their life, and therapies can only help them improve their condition by managing symptoms, not curing them. Patient-oriented approaches are common to caring for people with chronic conditions because patients' priorities become relevant means of prioritizing therapies in the absence of a cure. While such type of approach is shown to be effective, it does not leverage evidence on the success of given therapies to achieve specific similar patient goals in the past. Evidence-Based Medicine (EBM) is a concept that was introduced to the medical field in the early 90s to invalidate previously accepted tests and therapies and replace them with new, more powerful, more accurate, more efficacious, and safer ones. Unfortunately, despite the prevalence of patient-oriented approaches for patients with chronic diseases, data collected on patients is not systematically leveraged to support therapy decisions. Combining evidence-based decision-making and patient-oriented approaches could potentially further improve patient outcomes by leveraging the most up-to-date data to recommend and discuss therapy options for patients with chronic conditions.

The development and implementation of Learning Health Systems (LHS) is another solution to improving patient outcomes, one that the US Institute of Medicine strongly recommends. The development and implementation of a LHS to support therapy choice for patients with chronic conditions could improve related decisions by fostering continuous learning regarding which therapy may help better achieve which patient goals. However, a learning process that systematically leverages a relevant basis of evidence to support patient-oriented approaches has yet to be defined. As such, this study aims at articulating a learning process for therapy decision-making in the context of chronic conditions. The result is framework and a demonstration of its application using the Goal Attainment Scale (GAS) and synthetic data.

List of Tables

List of Figures

1. Introduction

Health spending in Canada represented 11.5% of Canada's gross domestic product (GDP) in 2019 [1]. Chronic diseases are a major economic burden in healthcare. In 2010, conditions such as chronic obstructive pulmonary disease (COPD), cardiovascular diseases, mental illnesses, musculoskeletal disorders, and diseases of the nervous system and sense organs accounted for about 40.5% of all direct health care spending on hospitals, doctors, and drugs in Canada [2]. In 2016-2017, COPD was at the top of Canada's most expensive health conditions, with an associated cost of $753.3M. Another chronic disease in the top 5 was dementia, with a cost of $404.0M [3].

Chronic conditions are defined by the Public Health Agency of Canada (PHAC) as “diseases that are persistent and generally slow in progression which can be treated but not cured” [3]. Chronic diseases include cancer, cardiovascular conditions, chronic respiratory disorders, diabetes, and neurological conditions, among others. In the US, the Center for Disease Control and Prevention (CDCP) defines chronic diseases broadly as “conditions that last one year or more and require ongoing medical attention, limit activities of daily living, or both” [4].

Therapy decision-making for patients with chronic diseases can be complex. Such patients usually live with their illness(es) all their life, and therapies can only help improve their condition by managing symptoms, not curing them [5]–[7]. As such, there are many dimensions to consider when choosing the right therapy for a particular patient. Among them, we have patient goals, comorbidities, therapy center characteristics, such as location, interactions between different drugs, secondary effects, and more.

Patient-oriented approaches are common to caring for people with chronic conditions because patient priorities become relevant means of prioritizing therapies in the absence of a cure. These approaches focus on the patient’s perspective to decide which therapy the patient should undergo. One predominant example of such an approach is the Person-centered care approach, which, in 2001, was identified by the Institute of Medicine as one of the six pillars of quality health care for the 21st century [8]. Person-centered care is defined as "providing care that is

respectful of and responsive to individual patient preferences, needs, and values, and ensuring that patient values guide all clinical decisions" [9, p.6]. Another approach that has been proposed, especially for patients who have coexisting conditions and/or chronic diseases, is goal-oriented care (GOC). GOC focuses a patient specific health goals. Symptoms' acuteness and physical functional status (e.g., mobility) are examples of dimensions for which patients need to define how they see themselves and what the therapy should help them achieve [9]. Many studies are already advocating for a shift from "problem-oriented" care to "goal-oriented" care, especially in cases where there is multimorbidity, mental illnesses or other chronic diseases. This shift seeks to align clinical outcomes with patients' personal goals [10]–[17]. Other similar approaches include patient-centered care, patient-oriented care, family-centered care, family-oriented care, and person-oriented care. They all take a patient's perspective when choosing a course of action and have in common patients' goals as a focus for therapy decisions [16]–[18], [19], [20], [21].

Two key pillars of patient-oriented approaches are aligning patient goals with therapies and a shared decision-making (I.e., treatment choice) model. Patient goals are required in patient-oriented approaches to defining therapy objectives and are usually identified through patient-physician encounters [20]. They also need to be reassessed periodically to ensure the effectiveness of the chosen therapy and address changes in the evolving life and health of patients [16]. The identification of goals and their reassessment is achieved through discussion between the patient and the physician during a medical encounter [20]. These discussions are also part of the shared decision-making model, which has been strongly advocated for improving patient outcomes by including patients in the therapy choice process. While such an approach has been shown to be effective [5], it does not leverage a rigorous basis of evidence about the success of given therapies to achieve specific patient goals in the past. In the early 90s, Sackett et al. introduced Evidence-Based Medicine (EBM) to the medical field as a way to integrate "individual clinical expertise with the best available external clinical evidence from systematic research" [22, p.71]. Since then, the goal of EBM has been to invalidate previously accepted tests and therapies and replace them with new, more powerful, more accurate, more efficacious, and safer ones [21]. Unfortunately, despite the prevalence of patient-oriented approaches for patients

with chronic diseases, data currently collected on patients is not systematically leveraged to support therapy decisions with evidence.

Combining evidence-based decision-making and patient-oriented approaches could potentially further improve patient outcomes by leveraging the most up-to-date data to recommend and discuss therapy options for patients with chronic conditions. As patient goals are constantly evolving [16], the choice of therapy needs to be reassessed to ensure alignment with the current patient goals. As such, the knowledge base used for evidence-based decision-making is also continuously evolving. While combining these approaches provides an opportunity to improve patient outcomes for chronic conditions, the resulting approach should be understood as an ongoing process of realignment between patient goals and therapy choice. Furthermore, integrating evidence-based decision-making into patient-oriented approaches requires finding the means to systematically leverage patient-oriented data to improve care, which to our knowledge has not been achieved yet.

The development and implementation of Learning Health Systems (LHS) is another approach to improving patient outcomes, one that is strongly recommended by the US Institute of Medicine [22]. A LHS involves the use of clinical and research data as well as advanced computing capabilities to transform current health systems into rapidly learning systems that can support the provision of quality and cost-effective health care and improve patient outcomes [23]. According to the Institute of Medicine, a LHS is a health system in which science, informatics, incentives, and culture are aligned for continuous improvement and innovation. Best practices are then seamlessly embedded in the care delivery process, and new knowledge is generated through the delivery experience to improve patient outcomes at a lower cost [22] [24]. A LHS has also been defined as an organizational system with rapid-learning capabilities that quickly adapts to past and new medical data and research information to support the effective delivery of therapies [25]. In 2012, a national LHS Summit was held to establish a consensus on 10 LHS core values. Among these core values, *Scientific Integrity* states that a LHS requires the most rigorous application of science to ensure the validity and credibility of findings [26]. This core value is aligned with the evidence-based medicine approach, one of the major paradigm shifts in medical decision-making that has been shown to improve care [27]. Indeed, a LHS relies on "*findings*" as

the basis of decision-making. Similarly, evidence-based medicine integrates individual clinical expertise with the most *up-to-date clinical evidence* from systematic research [21]. Another core value identified for a LHS was *Person-Focused,* a patient-oriented approach that engages individuals, families, groups, communities, and the general population in improving healthcare [26]. This core value is broader than the one described for patient-oriented care but still has the same goal: to engage patients in making decisions regarding their health.

At the center of a LHS, one or more learning cycle(s) continuously integrate, transform, and disseminate data to improve evaluation, management, and patient care [28], [29]. A generic learning cycle involves three processes. The first is the conversion of data to knowledge (D2K). This knowledge is then used to improve performance (K2P). Finally, the system performance is evaluated to generate new data which is fed back into the next iteration of the cycle (P2D) [26]. Consequently, a learning cycle requires the definition of which data has to be collected, how it will be analyzed, how knowledge will be generated, what changes will be triggered by this knowledge, and how these changes will be disseminated and implemented [28], [30].

Thus, the development and implementation of a LHS to support therapy choice for patients with chronic conditions could improve related decisions by fostering continuous learning regarding which therapy may help achieve which patient goals. Since such an approach implies having the ability to make predictions to guide decision-making, continuous learning is needed to adapt therapy choice decisions as patient goals evolve over time [31] as well as to improve the quality of the predictions. A well-functioning goal-oriented care learning cycle would thus capture new goals and/or changes in patient preferences, and it would adapt therapy consequently. The example below illustrates how such a learning cycle could support therapy decision-making:

A young patient comes to an hospital with his parents to be treated. He has multiple complex medical conditions (e.g., chronic diseases). The physician explains the different therapy options to the parents and their child, but each therapy option is associated with various secondary effects, outcomes, and impacts on quality of life. Thus, the parents and their child must choose the therapy best suited for their goals and preferences. To make this decision, typically, they would discuss the different options with the physician and compare their alignment with the

child's goals. A learning cycle could be implemented to analyze patient goals and preferences and recommend a ranking of therapies based on existing evidence. The physician could then discuss the different options with the patient and his family and jointly decide on a therapy option. This discussion would benefit from the physician's experience, evidence in the form of ranked therapy recommendations, and a shared decision-making process. The outcomes of the chosen therapy could then be measured and integrated into the learning cycle to better predict which therapy is best suited for which patient based on the new data generated from previous therapy decisions.

As such, a learning model could help propose therapies aligned with patient goals, preferences, and medical outcomes. The physician would then be able to discuss different therapy options with the family based on their goals, clinical evidence, and experience. While this approach should improve patient satisfaction with respect to therapy, the initial therapy choice would be based on the patient's and parents' current goals and preferences, which may change over time. Feeding data related to evolving goals and, as a result, potential dissatisfaction with therapy recommendations into the learning model should help physicians adapt the recommendations over time.

1.1. Key Terms and Acronyms

Refer to appendix A.

2. Research problem

Integrating patient goals and shared therapy decision-making for patients with a chronic condition(s) within a LHS requires developing a learning cycle focused on continuously improving the care provided to this population. As patient-oriented approaches are strongly recommended for such patients, leveraging patient-oriented data (e.g., goals, preferences, and more) to support therapy decision-making could improve patient outcomes. However, a learning cycle that systematically leverages a relevant basis of evidence to support patient-oriented approaches has yet to be defined. Rather, therapy choice decisions typically rely on medical expertise and experience without evidence as a standpoint [32]. In addition, while LHS are defined as patient-oriented [22] [24], in practice, patient goals are not systematically collected and analysed to inform therapy decision-making. Most existing LHSs [30], [33]–[37] focus on Patient Reported Outcome Measures (PROMs) rather than on patient goals and preferences, and the data collected was not directly used to improve therapy recommendations.

The three generic processes used to describe a generic learning cycle (i.e., D2K, K2P, P2D) provide guidance on the development of a learning cycle for such patient, but they do not specify what data to collect, how to analyse it, and how to feed back the decision-making results in order to continuously improve decision-making for this domain. Moreover, there are three levels of learning, also called “loops,” in a LHS. Single-loop learning entails the identification of discrepancies between the intended and the identified performance of the system. Double-loop learning encompasses questioning the existing processes and procedures of the system. Triple loop learning refers to people understanding the process by which they learn and then learn how to learn [38]. Each loop impacts the improvement differently. In the context of therapy recommendations, the first loop helps improve the alignment of patient preferences to therapy outcomes based on patient-reported data, while the second loop questions the data and analytical model used to provide the recommendation.

Consequently, one way to improve shared therapy decision-making for patients with chronic diseases is to leverage therapy-related data stored in Electronic Health Records (EHR) and other databases. Moreover, to address the needs of such patients and support therapy decisions from a goal-oriented perspective, the patient's perspective must be included in the basis of evidence.

However, to the best of our knowledge, "goal-oriented data" is not systematically recorded in medical databases nor linked to therapy choice and outcomes. This may be due to the lack of knowledge on which patient data should be collected to support patient-oriented therapy decision-making and how they should be analyzed.

3. Research question

The research question driving this study is: "How can a double-loop learning cycle that supports therapy choice for chronic conditions be articulated?"

The answer to this question will be provided through the achievement of the following two research objectives:

RO1. Develop a framework that describes a double-loop learning cycle in the context of therapy choice for patients with chronic conditions.

RO2. Demonstrate the application of the framework.

RO1 focuses on the development of a framework which can be used as a basis for the implementation of a LHS for therapy decision-making. The framework will illustrate the learning flow from every patient translated into better therapy decision-making.

To achieve RO2, a synthetic dataset will be used to demonstrate the applicability of the proposed framework.

4. Methodology

The Design Science Research Methodology (DSRM) [39]–[41] was used to guide this research. Design Science Research (DSR) is "a research paradigm in which a designer answers questions relevant to human problems via the creation of innovative artifacts, thereby contributing new knowledge to the body of scientific evidence." [39] It is especially well suited for this thesis as it aims at developing a framework (the artifact) to answer questions related to therapy decision-making. A core principle of DSR is that "knowledge and understanding of a design problem and its solution are acquired in the building and application of an artifact" [39].

Design Science Research Methodology (DSRM) comprises of five (5) activities (see Figure 4.1):

1. **Identify problem and motivate:** In this activity, the researcher the problem and its context to build a background to show the importance of the problem.
2. **Define objectives of solution:** In this activity, the researcher aims at defining, based on existing knowledge or theory, what the artifact should achieve to solve the problem. The research question and objective(s) are then defined.
3. **Design & Development:** In this activity, the researcher aims at creating the artifact through iterations of design and development.
4. **Demonstration:** In this activity, the researcher should demonstrate that the artifact solves the problem initially defined.
5. **Evaluation:** In this activity, the researcher evaluates performance and iterate back to the previous activities as necessary.

For this thesis, the first two activities were completed through a comprehensive literature review on LHS, learning cycles and outcome measures for chronic conditions. The review provided the relevant information required to understand the context surrounding the problem at hand and its importance. The research question and objectives, as stated in Section 3 section, reflect the acquired knowledge for therapy decision-making for patients with chronic conditions. The *Design & Development* and *Demonstration* are described in the sections below. The evaluation is beyond the scope of this thesis.

Figure 4.1 describes how the DSRM was applied for this thesis.

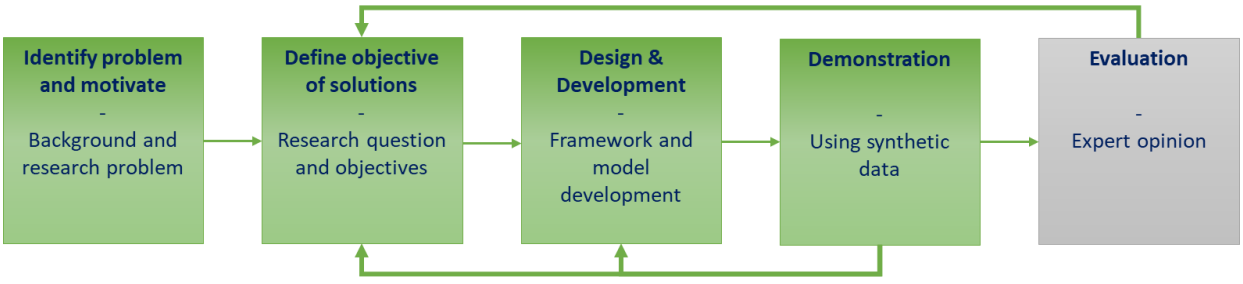

Figure 4.1. Design Science Research Methodology

4.1. Design and development

4.1.1. Data requirements

Given the context of care for which the framework was developed, our first step was to identify the patient-oriented outcome measures used for chronic diseases. We needed an outcome measure that allowed us to link patient goals and preferences to clinical outcomes. We identified, in the literature, three predominant outcome measures used in this context, namely the Goal Attainment Scale (GAS), the Canadian Occupational Performance Measurement (COPM), and the Patient-Reported Outcomes Measures (PROMs).

The GAS was introduced in 1968 by Kiresuk and Sherman [42] to assess therapy outcomes in mental health. Since its initial release, it has been adapted for other areas such as elderly care, chronic pain, cognitive rehabilitation, amputee rehabilitation, geriatric care and learning disabilities among others [43], [44]. The GAS is particularly well suited for measuring outcomes in the context of chronic diseases where trade-offs between therapy outcomes are common in therapy choice. It is a method to score the extent to which patient goals are achieved after undergoing a given therapy. Each patient defines their own outcome measures, which are associated with specific goals. The outcome measures are then scored using a standardised numerical approach. Outcome measures are compared to "success" criteria established "a priori" between the physician and the patient before therapy. The physician helps patients set realistic expectations and create scaled goals for their outcome measures. Scaled goals are defined as goals described with sufficient detail to allow the measurement of their achievement [45]. SMART goals (Specific, Measurable, Attainable, Relevant and Time-Bound) can also be used [43]. Each goal is rated on a 5-point scale, each point corresponding to a different level of achievement.

If the patient achieves the expected outcome level, this is scored at 0.

If the patient achieves a better-than-expected outcome, this is scored at:

+1 (somewhat better) or

+2 (much better)

If the patient achieves a worse-than-expected outcome, this is scored at:

-1 (somewhat worse) or

-2 (much worse)

The levels of achievement of the different goals can also be weighted relative to their importance (from the patient's perspective) and probability of success (from the physician's perspective).

An overall goal attainment score can then be calculated using the following formula:

$$\text{Overall GAS score} = 50 + \frac{10\ \sum(w_i x_i)}{[(1-\rho)\sum w_i{}^2 + \rho(\sum(w_i)^2]^{1/2}},$$

Where w_i is the weight assigned to the *i*th goal (if equal weights, $w_i = 1$ for all i), x_i is the numerical level of achievement (between –2 and + 2), and ρ is the expected correlation of the goal scales.

For practical purposes, according to Kirusek and Sherman [42], ρ is typically close to 0.3, so the equation simplifies to:

$$\text{Overall GAS score} = 50 + \frac{10\ \sum(w_i x_i)}{\sqrt{(0.7\sum w_i{}^2 + 0.3(\sum(w_i)^2)}}$$

The resulting score is then the sum-product between attainment levels and the relative weights of the different goals, transformed into a standardized measure called "T-score" [42], [45]. Thus, by construction and over a sufficiently large number of patients, one would expect the overall GAS score to be normally distributed with a mean of 50 which is the threshold for "success" and a standard deviation of 10 [42]. The score provides an overall measure of change, or more specifically, an overall measure of goal achievement resulting from undergoing a specific therapy. It is important to note that the GAS is limited by its use of ordinal non-linear scores, which can compromise the interpretation of change [46]–[48]. In an ordinal scale (non-linear), the marginal categories (e.g., 1 to 5) are limited to the pre-defined categories and the respondent estimates the value closest to his perspective of the score, whereas a linear scale offers more range in scoring allowing for a more precise measurement of level. The ordinal scale then allows for over or under-interpreting of change due to the limited categories of level.

The second outcome measure identified in the literature was the Canadian Occupational Performance Measure (COPM). The COPM is a published instrument with an approach similar to that of the GAS. It takes a patient-oriented approach to identify performance problems, concerns, and issues from a patient perspective. The COPM is based on the results obtained from a semi-structured interview and was designed by occupational therapists taking into account three aspects: self-care (personal care, functional mobility and community management), productivity (paid/unpaid work, household management and play/school) and leisure (quiet recreation, active recreation and socialization) [49]. The COPM follows a five-step approach. In Step 1, the patient and the therapist identify problems in daily activities. In Step 2, the patient is asked to rate the importance of each problem to determine the five most important. These five problems are then rated in Step 3 on a scale of 1-10 with respect to two dimensions: performance (ability to perform the activity) and satisfaction (on their performance). A baseline level (between 1-100) is then determined for each dimension by multiplying the dimension score and the importance rating. The summation of each dimension score divided by the number of rated activities gives an overall score that can be used for comparison purposes across time. The therapist and the patient then discuss the goal of the therapy (e.g., restore mobility = improve performance score and/or satisfaction score or maintain/prevent change in performance score or satisfaction score for mobility). To understand the problem, the therapist may need to assess performance to identify those components contributing to the patient's difficulties in performing an activity (e.g., physical endurance, range of motion, etc.) These evaluations help the therapist in identifying the cause of dysfunction and plan an appropriate intervention in order to achieve the patient's goals. Step 4 is a re-assessment of the problem after the patient completes therapy. The patient is asked again to rate performance and satisfaction for the same five problems identified in Step 2. These new scores are then multiplied by the initial importance rating, summed, divided by the number of rated activities, and finally compared to the baseline scores to measure change. Finally, Step 5 is a follow-up step where the patient and the care provider discuss therapy continuation. To that end, they use a new COPM form and repeat Step 1 to identify any remaining problems and decide together on the next steps [49], [50]. Despite its ease of use, the COPM limits the range of goals to three categories (productivity, leisure, and self-care). Consequently, it is not suitable for a patient who may want to define goals outside of these three categories. This results in a limited scope of application for this outcome measure [49].

The third outcome measure identified in the literature is the Patient-Reported Outcome Measures (PROMs). PROMs are measurements of health status from a patient's perspective and involve, among other things, symptoms, level of function, perception of health, and health-related quality of life [51]. The number of measures included in PROMs is in the thousands and are still expanding, with new ones added every day [10]. PROMs measure patients' preferences, health, well-being, and behavior to inform clinical care decisions and to support research on healthcare interventions efficacy in care delivery [52]. The type of data (e.g., preferences, symptoms, level of function, etc.) usually dictates how it can be collected. For example, when dealing with patient preferences, it has been recommended the use preference elicitation methods such as Conjoint Analysis (CA), Discrete Choice Experiments (DCE) and Analytic Hierarchy Process (AHP), among others [53]. For other types of PROMs, standardized self-administered questionnaires can be used to collect data. For example, the MD Anderson Symptom Inventory (MDASI) is used to measure symptoms-related PROMs, and the EQ-5D questionnaire is used to measure health-related quality of life-related PROMs [53]. Despite their benefits, PROMs data aggregation and their integration into quality improvement still remain unclear. There is a lack of standards regarding which measures should be selected, collected, interpreted, reported, and merged into clinical databases to ensure validity for clinical care and clinical decision-making [10], [53]. While PROMs offer a wide variety of categories, this type of outcome measure does not offer a systematic method of collecting data. From a goal-oriented care perspective, PROMs are very well suited to measure subjective goals as this approach focuses primarily on patient perceptions and feelings about themselves. However, PROMs are not appropriate for patients who are unlikely to improve or are expected to decline, as it is the case for a patient with chronic conditions. In addition, not every possible patient goal has a validated PROM, and the interpretation of change or the meaning of a PROM can be difficult to understand from a patient perspective. [31]

Choice of Outcome Measure

The decision of which outcome measure to use could be based on a number of criteria. For example, the outcome measure of choice should:

1. Allow measuring the impact of a given therapy on achieving patient goals in the context being considered. This criterion ensures that the outcome measure can be adapted to chronic conditions where goals may vary from decreased pain to improved mobility or to retained autonomy. The latter implies that the outcome measure needs to allow for the measurement of patients expected to decline or patients that want to prevent change;
2. Provide a standardized utilization protocol on which a learning cycle for therapy recommendations can be built. The protocol will serve as the basis for the framework. This criterion ensures the alignment of the framework with the process by which patients are treated;
3. Be useable without specific user training. This criterion ensures that users will not be required to attend a particular user training session or require a medical degree.
4. The use of the outcome measure must be accessible for free.

Among the three outcome measures identified in the literature and described above, the Goal Attainment Scale (GAS) is the only one that meets the above four criteria. It is especially well suited for situations in which there exist trade-offs between therapies and patient goals, as it is the case with chronic diseases. Prioritization of therapy is then based on patient preferences for the trade-offs. The GAS was created to assess the impact of a therapy on a patient's goal. It also has a standardized protocol, including an interactive spreadsheet tool acting as a “cheat sheet” that can be used for help with the long calculations that come with its use [43], [44]. This spreadsheet can be especially useful in developing a semi-automated process to analyze data and iteratively recommend therapies. The spreadsheet facilitates the use of GAS by automating the score calculation and providing a template for our model. Additionally, it is free to use and does not require specific user training [49].

The Canadian Occupational Performance Measure (COPM) does allow to measure the impact of therapy on patient goals using two dimensions, performance, and satisfaction. By comparing the impact of therapy (i.e., the difference between pre-and post-therapy scores) for each problem identified by the patient, it is easy to understand how a therapy impacts the achievement of patient goals. Additionally, there is a standardized 5-step approach to be followed on which a

learning cycle can be built upon to rate and recommend therapies. However, the COPM is a copyrighted outcome measure approach and, consequently, could not be used in this thesis.

Finally, the Patient-Reported Outcome Measures (PROMs) could be used to measure the impact of a therapy on patient goals by using the preference elicitation methods suggested previously. While PROMs can be used without specific training, thousands of PROMs are available for various diseases. PROMs are separated into two categories: disease-specific and generic. The generic PROMs generally focus on aspects such as self-care and mobility, which may not always be relevant for some chronic diseases. The disease-specific PROMs are the opposite. They may be too specific to one condition to be suitable for chronic diseases in general. Using multiple PROMs would be complex, particularly when describing the development of the proposed framework and demonstrating its use. As we aim to develop a generic framework for patients with chronic diseases, the outcome measure of choice must be flexible enough to be applicable to most chronic diseases.

The choice of outcome measure greatly impacts the data requirements associated with the proposed framework. It dictates the data needed as well as the model outcomes. For GAS, the data inputs are simply patient goals, patient preferences, and therapy options. The latter are required to link their impact to patient goals. Thus, the data requirements associated with the use of GAS are very clear.

The GAS provides two outputs. First, the T-score offers a measure of the overall goal achievement post-therapy. Second, each goal has its own individual score to assess the success of therapy. These two outputs are used in the next section to evaluate the impact of therapy on patient goals and rate which therapy is best suited for a patient, as well as to develop two feedback loops to improve therapy recommendation. The first loop, based on the individual goal scores, will update the therapy profiles, while the second loop will identify when patients are systematically not satisfied by the model recommendation

The data requirements for our framework include patient goals, patient preferences and therapy options. To recommend the right therapy for a patient, a link between therapy, patient goals, and

outcomes must be established. To link therapy to patient goals, subject matter experts are needed to identify which therapy is suited for which goal. Consequently, data on patient goals and therapy options need to be collected and stored in a database for further analysis. Categories of goals would need to be defined to limit the number of goals to a reasonable number. For example, similar goals could be grouped together as the same therapies are likely to be recommended. Therapy options also need to be available in a database for the model to match patient preferences to the right therapies. As such, subject matter experts are needed to identify which therapy can help achieve which goal. The development of these databases is outside of the scope of this study but still is required to implement the framework and the model described later in practice.

4.1.2 Learning model development using GAS

We used the GAS protocol to establish a framework for patients' journey, from visiting their care provider to undergoing therapy and providing feedback on its impacts. This framework was then used to identify the critical points where the feedback loops would be implemented to improve therapy recommendations. Developing a learning cycle in which the GAS can be used was key. A learning cycle within a Learning Health System comprises three different learning levels: single-loop, double-loop, and triple-loop learning [54]. Single-loop learning entails the identification of the gap between the intended and the identified performance of the system. Double-loop learning involves a broader scope of learning. It often requires audit and evaluation of performance to implement more extensive changes at the organisational level. The definition of new organizational goals and directions or different processes through which the organization can attain its goal, are example of double-loop learning. Lastly, triple-loop learning is about "learning to learn better." It refers to the capacity of the system to learn about "learning" and improve upon that learning [38]. This study will focus on the first two loops as the third loop is usually achieved through people within organizations.

To align therapy recommendations with patient goals and preferences, we also needed to match patient goals with therapies. To this end, we used a scale from 1 to 3 for the initial rating of each goal (we use the name "attributes" of therapy instead of goals for now on) to develop therapy profiles comparable to the patient's rating. The importance of each attribute, as defined by the

patient, becomes the "patient profile," and therapy profiles can be ranked based on the fit to the patient profiles. The care provider can then use this ranking to recommend the therapy that is best suited for the patient. The therapy profiles were generated randomly for this study, but subject matter experts would be necessary for a real-world application to develop accurate initial therapy profiles. While subject matter experts developing the initial therapy options for patient goals and the categorization of goals does not support an evidence-based decision-making approach, it is the proposed model that embeds the most up-to-date data into therapy recommendation by indicating when an update of the therapy options and goal categorization is required. The model will then integrate patient data into the basis of evidence for therapy recommendations by providing an accurate picture of patient outcomes in regard to the therapy received.

The model was developed using the GAS protocol and formulae. We added a learning function to analyze the patient data and the outputs of GAS (overall GAS score and individual attribute scores). The learning function comprises two loops. The first updates the therapy profiles when there is discrepancy in individual attribute scores. The second identifies when the model fails systematically and needs to be reviewed completely (i.e., when the choice of outcome measure and inputs, for example, need to be revisited).

For the first learning loop, the variance in individual attribute scores was used to identify when an attribute of a therapy profile needed to be updated to better reflect the impact of the therapy on the attribute. We developed simple rules to update the therapy profiles based on the average variance for each attribute, but subject matter experts should set the thresholds for implementation. For example, an attribute with a positive average variance greater than the threshold suggests that the therapy profile should score higher for that attribute. The first learning loop can be automated by repeating iteratively this process when sufficient new data is available.

For the second learning loop, we used the overall T-score obtained from the overall GAS formulae presented earlier. This measure of change indicates overall satisfaction toward the therapy. This second loop is important for our model as it ensures continuous improvement. The

first learning loop has some limitations regarding quality improvement; it is highly dependent on the data fed to the model. Hence, if the data used to recommend therapy is insufficient or inadequate, it will fail to improve recommendations, and this failure will be reflected in the overall satisfaction level (T-score). The mean of the T-score (50) is the threshold that defines the success or improvement of a goal. For our model, the second learning loop only indicates when the model fails systematically (T-score consistently below 50). The second loop encompasses questioning the suitability of the model being used and its inputs; it includes revising goals and preferences and changing outcome measures if necessary. Hence, it cannot be automated within the model, but it serves as an indicator that manual revision is required.

4.2. Demonstration

For the demonstration, we used a synthetic dataset and a spreadsheet to automate the first learning loop and update the therapy profiles using a synthetic dataset. The dataset was generated randomly based on the identified data requirements and did not represent actual patient data. Fifty-one patients were simulated with the following data: the therapy the patient received, their specific goals, their rating of importance for each specific goal (from 1 to 3), and the achieved level post-therapy (from -2 to 2). We aimed at demonstrating how the first loop could learn from new data by consciously making the variance of specific attributes higher or lower in different simulations experiments to ensure that therapy profiles were updated following the established rules below. The simulations were done iteratively to test every pre-established rule and ensure that the therapy profiles were updated accordingly. The rules employed are arbitrary but could be adapted to specific contexts or defined using data-driven methods in a real-life implementation.

Rules:

- If the average variance is > 2.5, the weight of the attribute is increased by 1;
- If the average variance is > 1 and <= 2.5, the weight of the attribute remains the same;
- If the average variance is > -0.5 and <=1, the weight of the attributes is reduced by 1; and
- If the average variance is <= -0.5, the weight is reduced by 2.

In summary (see Figure 4.1), the first step was to define the data requirements for our framework through the identification of an appropriate outcome measure (GAS) in the literature. The outcome measure facilitated the identification of *Data Requirements* by providing a standardized protocol, which included inputs and outputs for its use. It also provided guidance on what types of data were required (patient goals, patient preferences, and therapy options). The second step was the *Model development* using the outcome measure previously identified. We started by developing a generic learning framework and then employed the GAS protocol to capture a patient's journey from visiting their care provider to undergoing therapy and providing feedback on its impacts. Therapy profiles were then developed based on the GAS inputs to allow for comparison between patients' preferences and therapy profiles and rank therapy options. We then used a GAS spreadsheet to develop the first learning loop involving the average variance in attributes (difference between achieved score and baseline) and pre-established rules to update the therapy profiles. For the second learning loop, the overall GAS score was used to identify when the model fails systematically (overall GAS score below 50), indicating the need for goals and attributes to be revisited for a better goal categorisation or to better capture goals dimensions. The choice of outcome measure should also be revisited if goals and attributes cannot be improved over time.

5. Results

This section is divided in subsections. The first subsection defines the framework developed using the Goal Attainment Scale (GAS) and describes the generic process behind its use. The second subsection demonstrates its application using synthetic data.

5.1. Framework

As previously mentioned, we aimed at articulating a learning cycle that supports therapy choice for patients with chronic conditions. As such, our first objective was to define a framework that defines a double-loop learning cycle in this context. Figure 5.1 illustrates the proposed framework on which we based our demonstration. It shows how patient goals and preferences align with therapy choice and how recommendations can be improved over time. The framework has six layers. The first is the Goals layer, where we identify a patient's different goals. The second layer identifies the different Attributes (dimensions) that describe each goal. Various dimensions need to be considered when working toward achieving a specific goal. Those dimensions need to be prioritized based on patient preferences (e.g., a high level of mobility is preferred to a low level of pain). The third layer, Preferences, defines the importance given to a dimension of a goal or a specific attainment level from the patient's perspective such as the importance given to a dimension of a goal or a specific attainment level (e.g., mobility is currently a 2 out of 5, to goal is to achieve a 4 out of 5). Although, preferences still need to be discussed with the physician to ensure realistic expectations, the use of SMART goals or SCALED goals as in the GAS can circumvent the issue of unrealistic aspirational goals. SMART goals are Specific, Measurable, Attainable, Relevant and Time-Bound goals [43] while scaled goals are defined as goals described with sufficient detail to allow the measurement of the achievement [45]. The fourth layer is the mathematical model used to support treatment choice recommendations.

Such a model needs to collect and analyse patient data and allow for therapy ranking based on the attributes previously identified. For the demonstration, we used the Goal Attainment Scale with an added learning function. Still, any other outcome measures relevant to the context of chronic care could be used within this framework. The fifth layer is the Therapy Recommendation. The output could be a rating or a ranking of therapies. Finally, the sixth layer

involves computing the performance measures to update and improve recommendations. Multiple performance measures could be included here depending on the mathematical model used. From the Goal Attainment Scale, we consider individual attribute scores as well as the overall GAS score, which provides a measure of how well the therapy is aligned with patient preferences to develop the first and second learning loops.

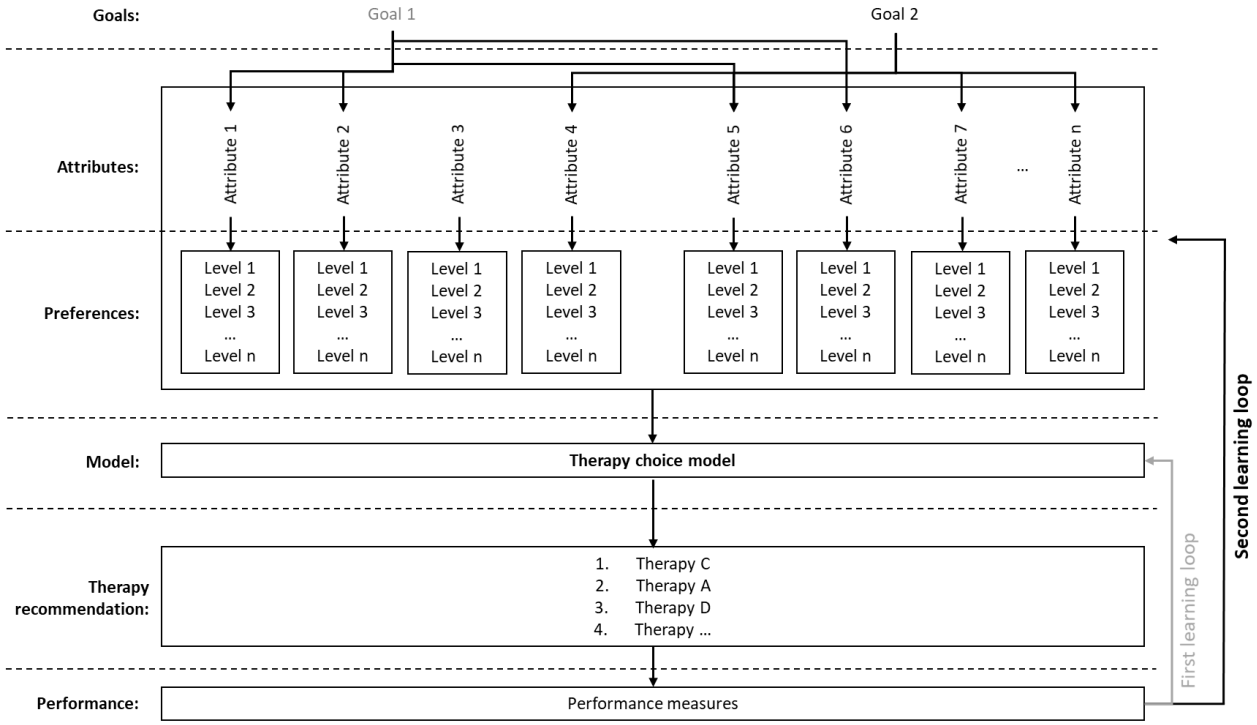

Figure 5.1: Generic learning cycle for therapy decision-making

Using the Goal Attainment Scale (GAS), we developed a process map to illustrate how a double-loop learning cycle can be implemented in the context of patients with chronic conditions. Figure 5.2 illustrates the flow of learning from every patient encounter with the goal of improving decision-making about treatment choice. The process shown in Figure 5.2 comprises two different learning loops, each leading to a better recommendation of which treatment is best suited for which patient goals. The first learning loop, in green, shows the steps required to update an attribute weight, while the second learning loop, in red, refines the definition of goal dimensions.

The interactions between data, patient and physician, and the system can be seen across the lanes shown in the process model in Figure 5.2. The process starts with a patient with one or more chronic conditions coming to a clinic. A discussion with a physician allows to identify an overall goal, such as being able to work or attend an important family event. The overall goal is then matched with a corresponding category in a *patient goals database*. Next, the system pulls out different therapy options with their corresponding attributes from a *therapy options and attributes database.* The process model assumes that these databases are available and accessible to the physician, and that data is continuously added and revised. Each therapy option will have its own "profile" that will provide information on the importance of each attribute or its impact on the patient's goal.

As an example of the process described above, consider the two therapies A and B in Table 1 with the following attributes: pain and convenience. The convenience attribute for therapy A has a level of 3, while for therapy B it is 1, with 3 being high and 1 being low. Then, therapy A would be rated with a higher weight for convenience. Therapy B may have a lower pain level during and after being delivered and would then be rated with a higher level of importance for pain than therapy A. In this example, the model will then pull the list of therapy options with their attributes, as shown in Table 1, and the physician will ask the patient to score the importance (from 1 to 3) of each attribute. Based on the patient importance scores, the model will create a ranking of therapy options to be used by the physician to guide the patient in deciding which therapy option to choose. Two therapies may offer the same or very similar outcomes. In this situation, the physician's expertise would play a key role in helping the patient with their therapy choice. He might also consider other variables which are not taken into consideration in the model (i.e., which are not attributes). The patient will then undergo the chosen therapy, and effects will be measured using a Goal Attainment Scale (GAS) and recorded in a "patient GAS" database for future analysis. The measurement of achievement is as follows:

If the patient achieves the expected outcome level, this is scored as 0.

If the patient achieves a better-than-expected outcome, this is scored at:

+1 (somewhat better) or

+2 (much better)

If the patient achieves a worse-than-expected outcome, this is scored at:

-1 (somewhat worse) or

-2 (much worse)

After a predefined number of total entries in the patient GAS database, for example a total of 50, we should have sufficient data to identify any discrepancies in average attributes variance. The attribute variance is defined as the difference between the achieved score (between -2 and 2) and the baseline (usually -1).

	A	B
Pain	1	3
Convenience	3	1

Table 5.1: Examples of therapy profiles

Thresholds must be defined by subject matter experts based on their expertise and clinical literature to develop ranges of variance, which will be used to adjust the therapy profiles later during the process. For example, thresholds could be set to ± 30% of the initial profile for the first learning cycle of this process. If no discrepancy in the variance is identified, the model will then look at the overall goal attainment T-scores.

The Overall GAS T-score can be obtained using the formula below:

$$\text{Overall GAS score} = 50 + \frac{10\ \sum(w_i x_i)}{[(1-\rho)\sum w_i^2 + \rho(\sum(w_i)^2]^{1/2}},$$

where w_i is the weight assigned to the *i*th goal (if equal weights, $w_i = 1$ for all i), x_i is the numerical level of achievement (between –2 and + 2), and ρ is the expected correlation of the goal scales and is typically close to 0.3 [45].

This score is a measure of change where a score ≥50 represents a successful therapy [43]–[45], [49]. The T-score is used to identify when the model fails systematically to provide quality recommendations. In this situation, we need to update the attributes used in the model, the mathematical function, or the granularity of the goals or their attributes to better predict which therapy is best suited for which patient goals. This update cannot be automated as it would require the manual intervention of experts. Our model assumes that if the overall goal attainment

T-scores are consistently low and there are no variance discrepancies, then the model does not capture all relevant attributes, or the current goals are not tied to the correct attributes. Then, using the current attributes, the model cannot improve its prediction with the data it is currently fed. We then need to revise the data required to predict which therapy is best suited for which patient. In this case, experts will revisit which therapy(ies) is(are) well suited for which goal. This will then ensure that the attributes capture the different dimensions of each goal. In short, if the model cannot learn from current data and offers poor predictions, it means that the data used to predict does not capture the dimensions of a goal and treatment profiles cannot be updated anymore; they need to be revised to identify new or different attributes, add, or remove levels, etc.

There might be situations where therapy is profiled correctly but does not reach a T-score of 50 or more. If the model consistently recommends a therapy that has a T-score below, for example, the mean (50) and no discrepancy is apparent in the therapy profiles, then this indicates that something different may be causing the model to fail. As such, the second loop, shown in red in Figure 5.2, questions the model's input, mathematical models used, the granularity of goals and attributes, and more to improve recommendations. The second loop is not automated as the first loop is. It requires a revision of the model and its inputs. One or more qualified people must do this by reviewing the therapy profiles and attributes, looking for better modelling options in situations where therapy profiles cannot be improved anymore.

5.2. Demonstration

The second objective of this study is to demonstrate the applicability of the framework described in Section 5.1. To this end, we used the Goal Attainment Scale spreadsheet developed by King's College of London [55]. A blank version of the tool is shown in Figure 5.3. The tool embeds the underlying calculation of the GAS and, as such, is practical for the purpose of the demonstration. To use this tool, a physician, and a patient first need to identify a patient's overall goal translated into several specific goals (SMART or scaled goals). Then, a description of these goals baseline (or current) level of achievement is specified. The baseline is usually -1 and is the current state of that specific goal. For example, if we use increased mobility as the patient goal, walking by themselves could be the specific goal, while the baseline would be the ability to walk a few steps (1 to 3). The patient is then required to "weigh" the importance of each specific goal on a scale of

1 to 3, with one being "important" and three being "extremely important." The physician also rates the probability of achieving each goal on a scale of 1 to 3, where 1 is "doubtful" and 3 is "probable." Once therapy is completed, the patient would be asked to score the achievement of each specific goal on a scale of -2 to 2, where -2 is "worse" and two is "a lot more achieved." Based on this information, the spreadsheet is then used to calculate the variance of each specific goal. At the bottom of the spreadsheet (see Figure 5.3), the final T-score and the change would then appear under "Achieved score" in the GAS calculation section.

For demonstration purposes, we assume that the probability of achievement equals 3 (GAS uses a scale of 1 of 3 for the probability of achievement, where 1 is “doubtful” and 3 is “probable”) for all specific goals. As we are trying to demonstrate the relationship between therapy performance and patient goals using synthetic data, the probability of accomplishment of each specific goal is not strictly necessary. In the process map in Figure 5.1, the spreadsheet would be integrated into the model lane as it is part of the automation of the first learning loop.

The first learning loop described previously will use the variance obtained from the difference between the achieved score (between -2 and 2) and the baseline (-1) to update the therapy profiles based on patient results. Initially, a patient's goal will fall under a category of goals. That category of goals will have a list of therapies with their "profile" (attributes and weighting) tied to it. For example, a patient is provided with the therapy options listed in Table 5.2. Four therapies are available to achieve the patient's goal, each with a different impact on the dimensions of that goal. In this case, five SMART goals or attributes are measured. Depending on the importance of these dimensions, one therapy may be more appropriate based on the patient's preferences than the others. For example, we can see that therapy 1 (T1) addresses the location dimension, but it does not help the pain nor the convenience. In contrast, therapy 2 (T2) is very good at handling the convenience dimension and good with pain.

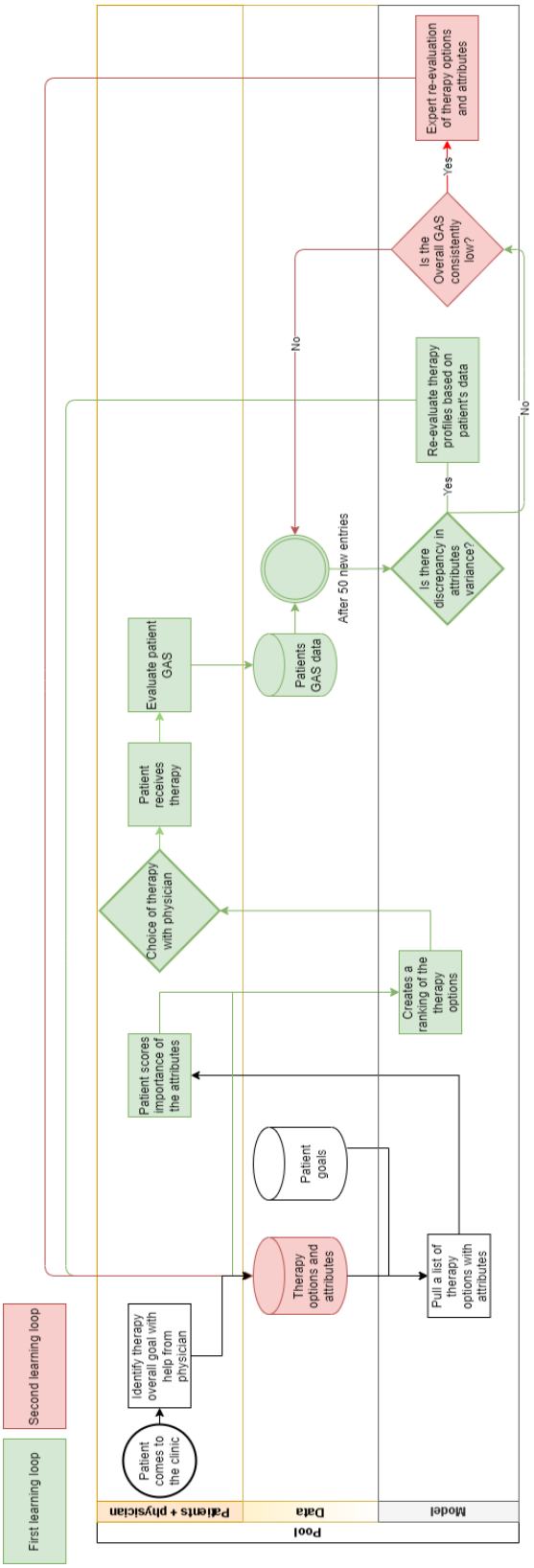

Figure 5.2: Application of the framework for therapy recommendation using GAS

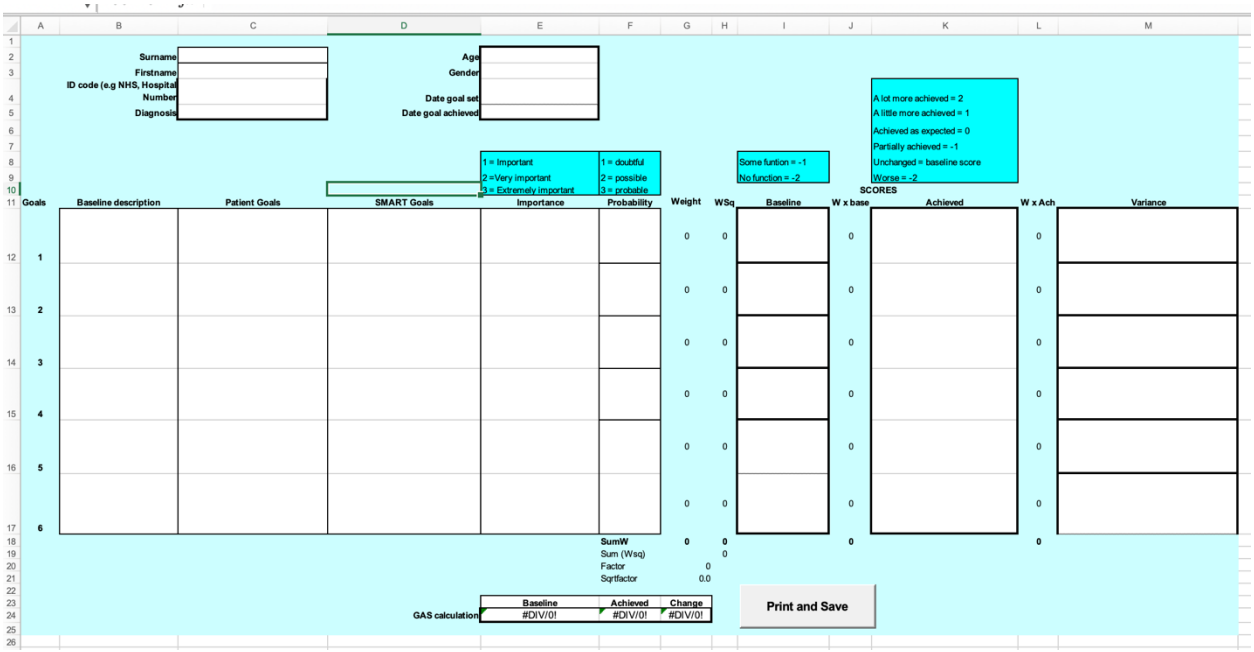

Figure 5.3. Goal Attainment Scale interactive spreadsheet (http://www.kcl.ac.uk/lsm/research/divisions/cicelysaunders/resources/tools/gas.aspx)

	T1 Profile	T2 Profile	T3 Profile	T4 Profile	Patient profile
Mobility	2	1	1	1	1
Pain	1	2	2	3	2
Location	3	2	3	1	3
Convenience	1	3	2	2	2
Duration	2	1	1	3	1

Table 5.2: Initial therapy profiles

With the therapy list at hand, the patient must rate the importance of the different attributes to create a "Patient Profile," which will then be used together with the therapy profiles to identify the best therapy. For example, a patient could create their own profile by rating the five attributes (mobility, pain, location, convenience, and duration) from Table 5.2. The results can be found in Table 5.3. and are as follows: mobility and duration are of low importance (importance scored at 1), pain and convenience are of medium importance (importance scored at 2), and location are of high importance (importance scored at 3). Multiplying the patient and therapy profiles will generate a score for each therapy. The higher the score, the better the fit. In Table 5.3, each attribute level for each therapy profile has been multiplied by the corresponding patient profile

level. The total score is provided in the last column. For example, the T1 score for mobility is two because the T1 profile for mobility has a level of importance (or weight) of 2, and the patient-rated importance for mobility is 1 (2 x 1 = 2). In the end, the sum of the attribute scores for each therapy (in yellow in Table 5.3) is used to rank the treatment. In this case, therapy 3 (T3) has the highest score (19), so it appears to be better suited to achieving the patient's goals.

	T1 score	T2 score	T3 score	T4 score
Mobility	2	1	1	1
Pain	2	4	4	6
Location	9	6	9	3
Convenience	2	6	4	4
Duration	2	1	1	3
Total	17	18	19	17

Therapy	Rank
T3	1
T2	2
T1	3
T4	3

Table 5.3. Therapy scores and ranking

The physician can then recommend a therapy using the ranking provided by the tool or use the results to further discuss the different therapy options with the patient before choosing one. Once the choice is made, the patient will undergo therapy and, afterward, an achievement score for each attribute will be collected. The variance is then calculated by subtracting the baseline (-1) to the achieved score (between -2 and 2) and used to update the profile of the therapy the patient chose. The scores are then stored in a "GAS database" to be utilized later during the next iteration of the learning cycle. After an expert's defined number of entries, the average variance for each therapy attribute is computed and used to update the corresponding therapy profile.

For this proof of concept, simple rules have been created to update the profiles:

- If the average variance is > 2.5, the weight of the attribute is increased by 1;
- If the average variance is > 1 and <= 2.5, the weight of the attribute remains the same;
- If the average variance is > -0.5 and <=1, the weight of the attributes is reduced by 1; and
- If the average variance is <= -0.5, the weight is reduced by 2.

We assume that an average variance higher than 2.5 means that the therapy has a better impact on the previously identified attribute (goal dimension) and needs to be updated. If the average variance is between 1 and 2.5 (inclusively), it suggests that the impact of the therapy on the attribute is correctly weighted, and it is what is expected from the therapy. If the average variance is lower or equal to 1, the therapy profile scores are lower for that attribute than expected and need to be updated. Using these rules, we can identify discrepancies in the variance and adjust therapy profiles accordingly. The variance for each attribute will fall in one of the different intervals and be updated according to the abovementioned rules. The same rules will be used to update the therapy profiles every 50 entries in total using the new weighting for the various attributes. This cycle will happen continuously. This is the first learning loop in this learning cycle where we learn to better predict the impact of treatments from the data. For example, in Table 5.4, we can see the rules to update the therapy profile and the average variance for each attribute and therapy using synthetic data. The variances for each attribute were automatically calculated using the GAS spreadsheet (See Figure 5.3) by subtracting the achieved level for a given attribute, with the baseline. It quantifies the difference between current level of achievement compared to the baseline which was the state prior to undergoing the therapy. The average variances are the sum of variances for a given attribute, divided by the number of variances. The example shown in Table 5.4. used a total of 50 entries but we also tried with 50 entries per therapies and did not see significant change.

Average variance

	T1	T2	T3	T4
Mobility	0.00	1.00	1.57	1.47
Pain	1.42	3.00	1.07	1.59
Location	1.08	0.86	1.50	1.18
Convenience	1.58	1.43	1.93	2.06
Duration	0.92	3.43	2.14	2.24

Table 5.4. Average variance for each attribute and therapy

In Table 5.4, T2 is an excellent example of a therapy that needs a profile update. The average variances for pain and duration are higher than 2.5. Consequently, +1 needs to be added to the levels/weights of these attributes in the therapy profile. Also, the location attribute is underperforming, with an average variance of just under 1 (0.86). Following the established rules, the attribute weight would need to be reduced by one. Table 5.5 shows the updated therapy

profiles based on the average variances in Table 3. The cells highlighted in green had their attribute weights increased, and those highlighted in red had their attribute weights decreased, depending on patients' ratings of accomplishment.

	T1	T2	T3	T4
Mobility	1	1	1	1
Pain	1	3	2	3
Location	3	1	3	1
Convenience	1	3	2	2
Duration	1	2	1	3

Table 5.5. Updated therapy profiles

Following an iteration of the first learning cycle, if the overall goal attainment scores are consistently low, there is the possibility that the model is failing to capture the necessary attributes required to align the therapy with the goal. In theory, provided that we can improve the weights of the therapy profile attributes indefinitely, we could improve the overall GAS score continuously. This is assuming that the attributes capture all dimensions of a goal correctly. However, if an attribute is missing or not relevant for particular therapy, the first learning cycle would not be able to deal with it. Here is where the second learning cycle plays a key role. It helps improve the capability of the model to better capture goal dimensions. This could be achieved by reviewing the categorization of each goal, the attributes associated with each goal, the levels of each attribute, or by identifying other therapies associated with a goal. A similar process as the first learning cycle could be used in this situation. First, a threshold could be defined to identify a low overall GAS score. Second, after (for example) 50 new entries in total, we could examine how well the model performs by comparing the overall GAS scores with this threshold or a time threshold (e.g., every 6 months) could also serve as indicator for review. Then, if most of the scores are lower than the threshold, it would indicate that the model fails systematically to make quality recommendations. In Figure 5.2, this is described as "every 50 new entries in total, if there is no discrepancy in attributes variance, the system will examine the overall GAS scores." If the overall GAS scores are found consistently low, or they have a lot of variability among individuals, a review of goal categories, therapy options, attributes and granularity will be required.

To summarize, the model developed based on the Goal Attainment Scale provides a ranking of therapies and a double learning loop to continuously improve this ranking. The first loop focuses on learning from the data collected post-therapy and seeks to update the therapy profiles based on patient experience. The second learning loop signals the need to revise the choice of outcome measure, attributes, or the attributes' granularity based on the overall T-scores, which indicate patient satisfaction with therapy. A rebuild is necessary when the model systematically fails to recommend a therapy that satisfies the patient. Outcome measures, attributes, and goals categorizations may be revisited to improve recommendations. The double-loop learning cycle allows for the model to be improved in a continuous manner by learning from the data and questioning the process by which it learns to determine if it could learn better.

6. Discussion

The research conducted as part of this thesis aimed at improving evidence-based decision-making in the context of goal-oriented care by operationalizing a double-loop learning cycle within a patient-centered Learning Health System to help determine which therapy is best suited for achieving specific patient goals. To this end, a proof-of-concept model was developed using the Goal Attainment Scale [45] and synthetic data. This model demonstrates how a double-loop learning cycle can be used to recommend therapy options aligned with patient goals and preferences. A demonstration of the use of the model was performed in the context of patients with chronic conditions for whom clinical outcomes of therapies are usually known, but multiple therapy options are typically available. These types of patients are more complex when deciding which therapy is best suited for them as the goal is not to cure, but to manage symptoms, making this particular context well-suited for the proposed model that aligns patient goals to therapy outcomes. This study shows how a Learning Health System (LHS) involving a double-loop learning cycle can improve therapy decision-making from a patient-oriented perspective. The double-loop learning cycle sits at the scientific dimension of a LHS architecture [54], where the process by which the LHS continuously improves is defined. The double-loop learning cycle describes how data is aggregated, analyzed, and used to generate new knowledge to be used in the following "learning" iteration. It also contributes to the LHS field by defining a framework that can be used as the basis for developing and implementing a full-scale LHS. While a full LHS integrates multiple dimensions (social, technical, ethical and scientific), learning cycles are part of the scientific dimension [54], and their definition is still a central piece required to enable the sustainability of a LHS. The learning cycle allows the health system to build on new data iteratively and continuously improve its function. In our case, the learning cycle could support therapy decision-making.

6.1. Situating the framework within Learning Health Systems

The learning theory from which the double-loop learning originates was first proposed by Chris Argyris in 1976 [56]. Single-loop learning is defined as correcting "behaviour" based on "errors" without changing the underlying reason for the error [57]. In a LHS, the first loop centers around learning from the data, where learning is defined as the detection and correction of errors [54],

[56]. In the case of the framework proposed in this research, the first loop involves updating therapy recommendations based on currently available data. The proposed framework uses patient data (preferences and goals) and clinical outcomes (measures of change in health status, functionality or quality of life resulting from care) to align therapy outcomes with patient expectations. Therapies are first profiled based on their clinical outcomes. Then, patient preferences are used to identify therapies that align with the goals of each specific patient. Patient-reported outcomes are continuously used to improve the therapy profiles by detecting misalignments and correcting errors using new data. This ensures that subsequent recommendations better align patient preferences and therapy outcomes. While, in theory, a higher volume of data may improve the quality of the recommendations, the first learning loop is based on specific events without interpreting them in their context, limiting the learning to what is known, leaving no room for questioning if the current process of learning is still effective [56]. It is also limited by the data it receives, which might not be sufficient to capture the full spectrum of attributes that define a goal.

To address the limitations of the first learning loop, the task of questioning processes and procedures is considered a second learning loop which involves reviewing the "how" and "why" we use specific data or get a particular result [57]. The second learning loop centers around improving the process by which the system learns [54], leading to changes at the core of the learning cycle. In this research, a second learning loop is required when the model consistently fails to accurately predict which therapy is best suited for which patient. The data requirements may be revised, and the modelling methodology changed for a better one. The data requirements may be revisited, adding, removing, or modifying which data is used and/or a new model may be developed based on a different approach, better suited for the specific context of use. Modifications to the data requirements will, in turn, impact which data is collected and stored in databases and will require subject matter experts to identify and apply the modification(s) required and lead the development of the new model. The second learning loop is then essential to maintaining the ability of the system to improve recommendations when the data analysis results are unsatisfactory. The second loop adds a mechanism to review the current state of the first learning loop and improve the model based on iterative feedback.

A recent study by Shaoibi et al. [58] proposed a Bayesian collaborative filtering algorithm to recommend treatments aligned with patient preferences and support shared decision-making. The algorithm combines patient preferences pre- and post-treatment using conjoint analysis, clustering, and patient satisfaction. The aim of the study is very close to ours, but the authors stopped at the first learning loop, where they used patient data to improve the quality of the recommendations. Our framework adds a mechanism, the second learning loop, to review how recommendations are made and ensure continuous improvement based on data. The latter, by improving the data requirements and choice of model behind the decisions. However, the algorithm proposed by Shaoibi et al. could be used within our framework, replacing the goal attainment scale, to create a more sophisticated first learning cycle.

The framework proposed in this research could be considered a clinical decision support LHS if implemented as a system or added to an existing one. Clinical decision support LHS are defined as electronic systems that support clinical decision-making by providing patient-specific assessments or recommendations to clinicians [59]. Our framework falls under the recommendation aspect of a clinical decision support LHS. Although, if it were implemented as a system, additional aspects would need to be considered, such as its interface and how the results are presented to users so they can access and interpret the recommendations.

Numerous LHS have been conceptualized or developed to include patient-reported data such as Patient-Reported Outcomes Measures (PROMs). Still, they have not explicitly aimed at therapy decision-making from a goal-oriented care perspective. Prominent examples include the *TRANSFoRm* initiative in Europe [33] to improve clinical research using patient (PROMs) and clinical data, and the *Learn From Every Patient* [30] LHS that combines data stored in EHRs and translational research to reduce health expenditure and develop a basis of evidence for research to improve care for children with cerebral palsy. Another example is *PEDSnet*, a clinical data research network that supports the development of a LHS at the national level. *PEDSnet* aims to develop a body of evidence of adequate size to produce generalizable knowledge for multiple pediatric conditions rather than for specific diseases [36]. A prototype using *PEDSnet* databases called *ImproveCareNow* was developed for quality improvement for children with Crohn's disease and ulcerative colitis, also called inflammatory bowel disease. *ImproveCareNow*

provides recommendations such as appropriate medication dosing and laboratory evaluations but does not directly address the alignment of therapy with patient preferences [37]. The framework proposed in this thesis is then part of a movement to incorporate patient-oriented data into the basis of evidence used for LHSs. Our framework was developed with therapy decision-making with a goal-oriented care perspective in mind. It complements the current knowledge on how LHS can be operationalized at the scientific level [54] by demonstrating how patient-oriented data can be used in a double-loop learning cycle for therapy recommendations.

Decision-making is a combination of information and preferences [59]. When deciding on a suitable therapy for a patient, the clinician's experience, knowledge, and preferences are considered [60], [61]. Quality of care was also shown to improve when patients are included in clinical decision-making, taking a patient-oriented approach [8], [9], [31], [61]. Additionally, quality of care can improve from physicians having access to relevant evidence to support their decisions and decision aid providing recommendations based on the best available evidence. This is in line with Evidence-Based Medicine, which is strongly recommended by the Institute of Medicine (IOM) [8], [22]. The framework proposed in this research could support these aspects by aligning patient goals with therapy decision-making and leveraging patient-reported data to improve therapy recommendations. While the physician has the final decision on what to recommend, their recommendation could be better supported by the proposed learning cycle.

Patient-centeredness has already received some attention from the LHS field. A few initiatives have taken a patient-centered perspective to develop a LHS framework, especially for mental health and chronic diseases [35]–[37], [58], [62]. However, there is still little information on operationalizing the learning loops, and most studies can be situated within the first level of learning. As mentioned above, we propose the addition of a second learning loop to further improve patient-centered learning within a LHS while also demonstrating the operationalization of both learning loops. The second learning loop ensures that the goal dimensions are correctly captured and fed into the first learning cycle to make quality therapy recommendations. It is in line with Evidence-based medicine (EBM), which is one of the guiding principles of a LHS. It forces the health system to invalidate previously accepted tests and treatments and replace them with new, more powerful, more accurate, more effective, and safer ones [21]. The definition of

EBM was initially focused on the clinical perspective, excluding any patient-reported data. Still, over the years, many new studies have used EBM to create a body of evidence on patient-reported data such as PROMs [30], [33], [35]–[37], [62]. However, to the best of our knowledge, no study has focused on patient goals and preferences as proposed in this research. This thesis thus contributes to the body of work on patient-centered LHS and the integration of EBM within a LHS by proposing a framework that integrates patient-centered data into the development of a basis of evidence in a manner that continuously supports improved personalized therapy decision-making.

6.2. Implications for practice

Users are essential to the implementation of the proposed framework. A recent study highlighted the fact that "doctors don't go to university to be told, by a computer, what to do" [52, p.4]. They rely on their own experience and knowledge to recommend what they think is the right therapy for different patients. Still, the concept of evidence-based medicine is increasingly advancing in health decision-making [63], pushing health practitioners to hone their decisions using evidence instead of intuition. Our framework can be thought of as the basis for the development of a clinical decision support system (CDSS) that is more acceptable from the physicians' point of view rather than a tool that would make decisions for them. It is important to note that the therapy recommendation model based on the framework only provides recommendations and that physicians still have the final say as to what therapy to propose to each patient. The goal a clinical decision-making LHS should not be to replace health providers but to help them make decision when the number of factors to be considered well exceeds the clinicians' ability to weigh every factor systematically [59]. As such, the proposed framework should be implemented as a decision aid for physicians to refer to during complex therapy decision-making.

Trust in the recommendations is another factor to consider when implementing such framework. Care providers need to trust the recommendations sufficiently to use the model and make decisions based on the evidence it provides. While trust is important, over-relying on the recommendations is also not the best approach as the recommendations still operates under uncertainty. Physicians' trust for recommendations may exceed the model's capabilities leading to errors. The model uses the data available to recommend the best treatment option (first loop).

If the data available is not sufficient or is not categorized correctly beforehand, it will fail to provide quality recommendations, but it would still be the care provider's responsibility to make the final decision on what is best for the patient [64], [65]. The opposite is also true. Physicians may not or rarely rely on the model's recommendations, which would affect the speed by which the model learns from new data as it would receive limited new input. In addition, from an Evidence-Based Medicine perspective, under-relying on the recommendations prevents the physician from using the most up-to-date evidence on current best practices [66]. When applying the framework to a real-world setting, the chosen outcome measure should be easy to understand or even known by the care providers to help build trust as users will know how it works and how to use it proficiently. It would also help them understand the reliability of the recommendations and make their therapy choice accordingly, preventing over or under-relying on the recommendations [67]. This proof of concept uses Goal Attainment Scaling (GAS), a measurement tool well-known by physicians working with patients with chronic conditions, making it easy to understand and use, hence easier to trust.

This framework sits at the scientific level of a LHS, where the purpose of a LHS and the activities that requires learning and improvement are identified [54]. In this study, the activity that requires learning is therapy choice. More specifically, two-levels of learning (loops) were identified in this study. The first level entails learning from the data and requires the implementation of databases where patient-reported data, as well as therapy outcomes, are stored for analysis. Moreover, adequate software and hardware are needed to access and analyse the data and generate recommendations. For the second learning loop, there is a need to identify a team of people with the required expertise to construct and revise treatment profiles.

Moreover, if the proposed framework is to be used, many stakeholders will need to work together to develop a functioning LHS. As such, data collection and data sharing could also be challenging; the data collected by the health organization would need to be complete and standardized and then shared across sites if there were more than one. Mechanisms to collect and share data must be put in place while following the organization's data regulations related to the security of sensitive information. While the previous considerations are all outside of the scope of this study, we acknowledge their importance when moving forward.

The application of the proposed framework was demonstrated in the context of chronic conditions using a Goal Attainment Scale. However, other variables may need to be considered for different contexts of care. Its implementation requires specific data based on the attributes and goals identified for the model (GAS or other) to work effectively. Goals and attributes typically depend on the context of care. They may need to be adapted to the site of care for seamless integration into the organization's health system. Identifying the required data is an immense undertaking. It involves mapping goals related to the context of care and critical attributes that describe the goals and capture patient needs and preferences. Adopting universal goals and attributes would limit patient-centeredness unless they are considered within the specific context of care [9], [31]. For the proof of concept described in this thesis, we used generic attributes relevant to the context of chronic diseases. While this context helped illustrate its application, the implementation of the proposed framework would need more thorough research on goals and attributes relevant to the particular context of care being considered, which will need to be validated by subject matter experts.

The presence of numerous preference-based approaches in health implies that the patient himself or the family responsible for the patient are innately important as they are central to the decision process [17], [19], [31]. Some key assumptions were made for the demonstration. Firstly, we assumed that physicians are rational and unbiased when defining goals with their patients. We also assumed that goal setting is achievable, and goals do not change in the short term. If the model used to demonstrate the framework were to be implemented, particular attention should be paid by physicians to ensure that patient goals are scaled appropriately (I.e., described with sufficient level of detail to allow the measurement of their achievement) or follow the SMART framework (Specific, Measurable, Attainable, Relevant and Time-Specified) [9], [31], [45], [55]. While this task is part of the process being considered, it is outside of the scope of this study. Another assumption made for the proof of concept is that every patient's goals and attributes can be categorized independently. However, there may be some associations between goals, especially from a patient perspective. The proposed second learning loop addresses this issue over time by revising goals and attributes when needed (e.g., adding goals, changing the granularity of others, etc.) and ensuring that the collected patient data capture the different

dimensions of each goal more accurately. Every iteration of the second learning loop will provide a more generalizable categorization where there are fewer associations between goals from a patient perspective and goals are more independent.

As this study takes a goal-oriented care approach to support shared decision-making, patient goals and preferences are central to accurately predicting which suitable therapy will be favoured by a patient. A key challenge is to operationalize the measurement of patient preferences while keeping the burden on the patient to a minimum and still improving patient-reported outcomes (e.g., patient satisfaction, quality of life). To minimize patient burden, we used the Goal Attainment Scaling (GAS), which requires very limited patient input pre-and post-treatment. Other approaches could have been used instead of GAS, such as PROMs, with a similar result, but GAS is more oriented towards the achievement of personal goals and less toward health status, which reduces the volume of information required from the patients [31]. In our framework, patients need to identify their primary goal(s), rank their respective attributes, receive a recommendation, and then report on the same attributes after undergoing the therapy. The patient's involvement is still there but limited to answering questions, which can be done during their first visit with their physician. Many preference-based approaches (e.g., conjoint analysis – CA and discrete choice experiments - DCE) require an extensive survey to profile patients before a recommendation is made and do not allow for an easy way to measure patient satisfaction afterward. As a result, an additional measure needs to be added, contributing to the complexity of the process the patient is a part of. In GAS, the trade-off measure is straightforward; it compares the same attributes pre- and post-therapy, while CA and DCE do not offer this possibility. An additional measure is then required for post-therapy satisfaction. Further research would be needed to implement the additional post-treatment measurement to allow the use of DCE or CA.

Furthermore, the proposed framework assumes a perfect alignment between a therapy preferred by a patient and supporting medical evidence. Yet, a preferred therapy might have the weakest evidence in terms of its efficacy for. In this research however, the results of the model are meant to support physicians' decision-making, not replace it. Hence, we assume that such situations

would be addressed through the physician's expertise and experience, which would prevent them from recommending a therapy with sub-par evidence.

6.3. Methodological considerations for the model

The use of the GAS for our model requires a limited amount of information from the patient, which has the advantage of keeping the burden on the patient to a minimum. However, the GAS has a number of limitations. First, it uses ordinal non-linear scores in a mathematical formula, which may compromise the interpretation of score changes [47], [48], [68]. A recent study [46] examined the robustness of GAS using Rasch analysis with respect to the use of a more standardized (linear) approach. In other words, it studied the impact of using an interval scale instead of an ordinal scale. The authors mention the fact that GAS assumes equal-interval data while the non-linearity of the ordinal scale suggests that non-equal intervals are more appropriate. A change in score from 3 to 4 may not be equal to a change from 4 to 5. Consequently, GAS tends to underestimate the change for extreme score values and overestimate it close to the mean value [46].

Since GAS uses an ordinal scale, it is also affected by the bias of ordinality. In an ordinal (non-linear) scale, the marginal categories (e.g., 1 to 5) are estimated. The score chosen is then an estimation, whereas a linear scale offers more range in scoring. The ordinal scale then allows for over- and under-interpretation of change. Increasing the range of the ordinal scale (e.g., from a scale of 1 to 5 to a scale from 1 to 10) may help avoid this issue. [46].

Another issue with GAS is that attributes require unidimensionality to allow the computation of sum-product of weights and attainment across attributes [46], [69]. Since attributes vary from patient to patient, in practice, the approach would require proof of unidimensionality of the attributes used. We address this issue by defining goals (in our case, attributes) using a standardized bank of goals and attributes. While the latter addresses this GAS limitation, it also creates a new challenge: defining the right granularity of goals and attributes. We assume that we can categorize patient goals in a way that captures every variation and allows us to predict the therapy that best "fits" a patient, which is nevertheless a challenging endeavour. The second learning loop in the proposed framework ensures the continuous improvement of the process by

which the model captures goals and preferences, including the categorization of goals and attributes. Tennant (2007) also recommended categorization as a solution to the critique that goal-oriented approaches are individualized but cannot be generalized. This “item banking” allows for generalizability in group comparison by using individualized measurement based on mathematical principles [46]. While this research is exploratory in nature, it is the first step towards future work. The use of GAS within a LHS is novel and creates a steppingstone to be used in the future development of patient-centered LHS, an emergent field of research.

7. Conclusion

This research aimed to articulate a double-loop learning cycle for therapy recommendations in line with patient goals and preferences in the context of chronic diseases. We first proposed a framework that defines how to operationalise a double-loop learning cycle for therapy decision-making. We then developed a model based on the Goal Attainment Scale (GAS) outcome measure where two learning loops allow for continuous learning on therapy results with respect to patient goals. Each goal is modelled using a set of attributes (dimensions) that need to be considered when deciding which therapy is best suited for which patient. The first loop improves therapy matching based on patient feedback on attributes. The second loop provides a mechanism to identify when the model systematically fails to provide a positive overall therapy satisfaction level and, consequently, needs to be revisited. For example, the choice of outcome measure may change, the goal categorization may be updated, and additional attributes may be required. We then demonstrated the applicability of the model using synthetic data.

7.1. Contribution

This study contributes to the growing body of research on Learning Health Systems (LHS) by providing a framework that could be used as the basis for implementing a full-scale LHS in the context of therapy decision-making for chronic conditions. Unlike other studies that consider only one learning loop, this thesis describes how a double-loop learning cycle can be operationalized to support therapy recommendations. In addition, the use of GAS within a double-loop learning cycle illustrates how patient-oriented outcome measures can be used to support therapy decision-making. It serves as a demonstration for future research to integrate EBM within patient-oriented care using a learning cycle. If the framework was implemented in a health system, it could benefit patients and physicians who must determine which therapy would be best for achieving specific patient goals. The proposed framework could also improve shared decision-making by providing a basis of evidence on which therapy is best suited for which patient, and the resulting recommendations could be used for discussion between patients and care providers. This study also integrates evidence-based medicine into a patient-oriented approach, something that has been strongly recommended by the Institute of Medicine [8], [22], [24], [61].

7.2. Limitations

In addition to the limitations of the Goal Attainment Scale mentioned before, several other limitations need to be acknowledged for this study. Firstly, we have applied the framework only to synthetic data. While we attempted to generate realistic data, it is practically impossible to capture the full complexity of a real-world setting. For example, the synthetic dataset is missing the probability of achievement of goals which the physician should provide on a case-by-case basis. In the future, empirical data should be used to test the applicability of the model in the context of chronic conditions. Secondly, a few assumptions were made for the sake of simplicity of this proof-of-concept. A small number of goals is probably far from realistic as every patient will define their own personal goals differently, although we assume that we can categorize them.

Further research should be conducted on categorizing goals before implementing the framework. Moreover, when using the Goal Attainment Scale, we assume the same probability of success for the different attributes while the physician would need to rate these probabilities in addition to consider the importance the patient gives to each attribute. Combining both measures (importance and probability of success) would allow for better overall satisfaction measurement. A low probability of success is usually reflected in the overall T-score value, preventing the T-score from sinking below the threshold when the chosen therapy poorly addresses a particular dimension. We also assume that the thresholds used to update therapy profiles are adequate. However, the use of formal statistical analysis should be explored to define the thresholds depending on variability. For example, after several iterations of the learning cycle, the therapy profiles will approach the actual scores. The intervals between thresholds will be too wide to identify the need for an update. As such, new thresholds should be identified, defining smaller intervals, and updating the therapy profiles more precisely.

Additionally, collecting preferential information from patient can be a difficult task. While therapy profiles provide the main dimensions impacted by the therapy, these dimensions may not be the most important for the patient. Thus, further research on data-driven approaches to defining therapy profiles is required to assess if this is assumption is correct. Another solution could be to have a validated list of dimensions for therapies which is updated based on experts'

re-evaluation and by collecting feedback from the patient on the important dimensions to consider when deciding on a therapy in their view. A pre-defined period of time could be used to gather such data or “comments” to improve the list of dimensions. For examples, patients could be asked: Are there any other important dimensions that should be considered in the choice of therapy? Or Were there any aspect of your life that were positively or negatively impacted by the therapy which are not included in the current model? This would ensure that patient input is taken into consideration when building therapy profiles aligned with actual attribute dimensions.

References

[1] Canadian Institute for Health Information. (n.d.). *Health spending.* Retrieved July 3, 2020, from https://www.cihi.ca/en/health-spending

[2] The Public Health Agency of Canada. (2017). *Economic Burden of Illness in Canada, 2010.* Retrieved July 3, 2020, from https://www.canada.ca/en/public-health/services/publications/science-research-data/economic-burden-illness-canada-2010.html#cost

[3] The Public Health Agency of Canada. (2013). *Chronic Diseases.* Retrieve December 5, 2021, from https://cbpp-pcpe.phac-aspc.gc.ca/chronic-diseases/

[4] National Center for Chronic Disease Prevention and Health Promotion. (2022). *About Chronic Diseases*. Retrieved December 5, 2021 from https://www.cdc.gov/chronicdisease/about/index.htm

[5] Charles, C., Gafni, A., & Whelan, T. (1997). Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). *Social Science & Medicine, 44*(5), 681-692. 10.1016/S0277-9536(96)00221-3.

[6] Van Cleave, J., Gortmaker, S. L., & Perrin, J. M. (2010). Dynamics of obesity and chronic health conditions among children and youth. *Jama, 303*(7), 623-630. 10.1001/jama.2010.104

[7] Cohen, E., Kuo, D. Z., Agrawal, R., Berry, J. G., Bhagat, S. K., Simon, T. D., & Srivastava, R. (2011). Children with medical complexity: an emerging population for clinical and research initiatives. *Pediatrics, 127*(3), 529-538. https://doi.org/10.1542/peds.2010-0910

[8] Baker, A. (2001). *Crossing the quality chasm: a new health system for the 21st century (*Vol. 323, No. 7322, p. 1192). British Medical Journal Publishing Group. Retrieved July 20, 2020, from http://www.ncbi.nlm.nih.gov/books/NBK222274/

[9] Reuben, D. B., & Tinetti, M. E. (2012). Goal-oriented patient care--an alternative health outcomes paradigm. *The New England Journal of Medicine, 366*(9), 777. 10.1056/NEJMp1113631

[10] Black, N., Burke, L., Forrest, C. B., Ravens Sieberer, U. H., Ahmed, S., Valderas, J. M., ... & Alonso, J. (2016). Patient-reported outcomes: pathways to better health, better services, and better societies. *Quality of Life Research*, 25(5), 1103-1112. 10.1007/s11136-015-1168-3.

[11] De Maeseneer, J., & Boeckxstaens, P. (2011). Care for noncommunicable diseases (NCDs): time for a paradigm-shift. *World Hospitals and Health Services: the Official Journal of the International Hospital Federation, 47*(4), 30-33.

[12] Kramer, M. H. H., Bauer, W., Dicker, D., Durusu-Tanriover, M., Ferreira, F., Rigby, S. P., ... & van Hulsteijn, J. H. (2014). The changing face of internal medicine: patient centred care. *European Journal of Internal Medicine, 25*(2), 125-127. 10.1016/j.ejim.2013.11.013.

[13] Kuluski, K., & Guilcher, S. J. (2019). Toward a person-centred learning health system: understanding value from the perspectives of patients and caregivers. *Healthcare Papers, 18*(4), 36-46. 10.12927/hcpap.2019.26030.

[14] Metselaar, J., van Yperen, T. A., van den Bergh, P. M., & Knorth, E. J. (2015). Needs-led child and youth care: Main characteristics and evidence on outcomes. *Children and Youth Services Review, 58*, 60-70. 10.1016/j.childyouth.2015.09.005.

[15] Noble, L. M., & Douglas, B. C. (2004). What users and relatives want from mental health services. *Current Opinion in Psychiatry*, *17*(4), 289-296. 10.1097/01.yco.0000133832.42167.76.
[16] American Geriatrics Society Expert Panel on Person-Centered Care, Brummel-Smith, K., Butler, D., Frieder, M., Gibbs, N., Henry, M., ... & Vladeck, B. C. (2016). Person-centered care: A definition and essential elements. *Journal of the American Geriatrics Society*, *64*(1), 15-18. 10.1111/jgs.13866.
[17] Tinetti, M. E., & Fried, T. (2004). The end of the disease era. *The American Journal of Medicine*, *116*(3), 179-185. 10.1016/j.amjmed.2003.09.031.
[18] Foster, K., Maybery, D., Reupert, A., Gladstone, B., Grant, A., Ruud, T., ... & Kowalenko, N. (2016). Family-focused practice in mental health care: An integrative review. *Child & Youth Services*, *37*(2), 129-155. 10.1080/0145935X.2016.1104048.
[19] Tanenbaum, S. J. (2015). What is patient-centered care? A typology of models and missions. *Health Care Analysis*, *23*(3), 272-287. 10.1007/s10728-013-0257-0.
[20] Charles, C., Gafni, A., & Whelan, T. (1999). Decision-making in the physician–patient encounter: revisiting the shared treatment decision-making model. *Social Science & Medicine*, *49*(5), 651-661. 10.1016/S0277-9536(99)00145-8.
[21] Sackett, D. L., Rosenberg, W. M., Gray, J. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *BMJ*, *312*(7023), 71-72. 10.1136/bmj.312.7023.71.
[22] APA, Olsen, L., Aisner, D., & McGinnis, J. M. (2007). *The learning healthcare system: workshop summary*. (McGinnis, J. M., Aisner, D., & Olsen, L., Eds.). National Academies Press (US). Retrieved November 1, 2020, from https://www.ncbi.nlm.nih.gov/books/NBK53481/
[23] Lessard, L., Grudniewicz, A., Sauré, A., Szczotka, A., King, J., & Fung-Kee-Fung, M. (2019). Type and use of digital technology in learning health systems: a scoping review protocol. *BMJ open*, *9*(5), 10.1136/bmjopen-2018-026204.
[24] Institute of Medicine (US) and National Academy of Engineering (US) Roundtable on Value & Science-Driven, *Roundtable on Value & Science-Driven Health Care*. (McGinnis, J. M., Olsen, L., Goolsby, W. A., & Grossmann, C., Eds.). National Academies Press (US), 2011. Retrieved November 8, 2020 from https://www.ncbi.nlm.nih.gov/books/NBK61957/
[25] Etheredge, L. M. (2014). Rapid learning: a breakthrough agenda. *Health Affairs*, *33*(7), 1155-1162. 10.1377/hlthaff.2014.0043.
[26] Friedman, C. P., Rubin, J. C., & Sullivan, K. J. (2017). Toward an information infrastructure for global health improvement. *Yearbook of Medical Informatics*, *26*(01), 16-23. 10.15265/IY-2017-004.
[27] Barratt, A. (2008). Evidence based medicine and shared decision making: the challenge of getting both evidence and preferences into health care. *Patient Education and Counseling*, *73*(3), 407-412. 10.1016/j.pec.2008.07.054.
[28] Flynn, A. J., Patton, J., & Platt, J. (2015, January). Tell it like it seems: challenges identifying potential requirements of a learning health system. In *2015 48th Hawaii International Conference on System Sciences* (pp. 3158-3167). IEEE. 10.1109/HICSS.2015.381.
[29] Grossmann, C., Goolsby, W. A., Olsen, L., & McGinnis, J. M. (2011). Engineering a learning healthcare system. *Washington (DC): Institute of Medicine (US) and National*

Academy of Engineering (US) Roundtable on Value & Science-Driven Health Care. 10.17226/12213.
[30] Lowes, L. P., Noritz, G. H., Newmeyer, A., Embi, P. J., Yin, H., Smoyer, W. E., ... & Miller, M. (2017). Learn From Every Patient': implementation and early results of a learning health system. *Developmental Medicine & Child Neurology*, *59*(2), 183-191. 10.1111/dmcn.13227.
[31] Reuben, D. B., & Jennings, L. A. (2019). Putting goal-oriented patient care into practice. *Journal of the American Geriatrics Society*, *67*(7), 1342-1344. 10.1111/jgs.15885.
[32] Buchanan, F., Cohen, E., Milo-Manson, G., & Shachak, A. (2020). What makes difficult decisions so difficult?: An activity theory analysis of decision making for physicians treating children with medical complexity. *Patient Education and Counseling*, *103*(11), 2260-2268. 10.1016/j.pec.2020.04.027.
[33] Delaney, B. C., Curcin, V., Andreasson, A., Arvanitis, T. N., Bastiaens, H., Corrigan, D., ... & Wagner, P. (2015). Translational medicine and patient safety in Europe: TRANSFoRm-architecture for the Learning Health System in Europe. *BioMed Research International*, 961526. 10.1155/2015/961526.
[34] Margolis, P. A., Peterson, L. E., & Seid, M. (2013). Collaborative Chronic Care Networks (C3Ns) to transform chronic illness care. *Pediatrics*, *131*(Supplement_4), S219-S223. 10.1542/peds.2012-3786J.
[35] Amin, W., Tsui, F., Borromeo, C., Chuang, C. H., Espino, J. U., Ford, D., ... & Weiner, M. (2014). PaTH: towards a learning health system in the Mid-Atlantic region. *Journal of the American Medical Informatics Association*, *21*(4), 633-636. 10.1136/amiajnl-2014-002759.
[36] Forrest, C. B., Margolis, P. A., Bailey, L. C., Marsolo, K., Del Beccaro, M. A., Finkelstein, J. A., ... & Kahn, M. G. (2014). PEDSnet: a national pediatric learning health system. *Journal of the American Medical Informatics Association*, *21*(4), 602-606. 10.1136/amiajnl-2014-002743.
[37] Forrest, C. B., Margolis, P., Seid, M., & Colletti, R. B. (2014). PEDSnet: how a prototype pediatric learning health system is being expanded into a national network. *Health Affairs*, *33*(7), 1171-1177. 10.1377/hlthaff.2014.0127.
[38] Sheaff, R., & Pilgrim, D. (2006). Can learning organizations survive in the newer NHS?. *Implementation science*, *1*(1), 1-11. 10.1186/1748-5908-1-27.
[39] Hevner, A., & Chatterjee, S. (2010). Design science research in information systems. In *Design research in information systems* (pp. 9-22). Springer, Boston, MA. 10.1007/978-1-4419-5653-8.
[40] Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 75-105. 10.2307/25148625.
[41] Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*(3), 45-77. 10.2753/MIS0742-1222240302.
[42] Kiresuk, T. J., & Sherman, R. E. (1968). Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. *Community Mental Health Journal*, *4*(6), 443-453. 10.1007/BF01530764.
[43] Turner-Stokes, L. (2009). Goal attainment scaling (GAS) in rehabilitation: a practical guide. *Clinical Rehabilitation*, *23*(4), 362-370. 10.1177/0269215508101742.

[44] Toto, P. E., Skidmore, E. R., Terhorst, L., Rosen, J., & Weiner, D. K. (2015). Goal attainment scaling (GAS) in geriatric primary care: a feasibility study. *Archives of Gerontology and Geriatrics*, *60*(1), 16-21. 10.1016/j.archger.2014.10.022.
[45] T. J. Kiresuk, A. Smith, and J. E. Cardillo, Eds. (2014). *Goal Attainment Scaling: Applications, Theory, and Measurement*. New York: Psychology Press. 10.4324/9781315801933.
[46] Tennant, A. (2007). Goal attainment scaling: current methodological challenges. *Disability and Rehabilitation*, *29*(20-21), 1583-1588. 10.1080/09638280701618828.
[47] Merbitz, C., Morris, J., & Grip, J. C. (1989). Ordinal scales and foundations of misinference. *Archives of Physical Medicine and Rehabilitation*, *70*(4), 308-312. 10.5555/uri:pii:0003999389901512.
[48] Svensson, E. (2001). Guidelines to statistical evaluation of data from rating scales and questionnaires. *Journal of Rehabilitation Medicine*, *33*(1), 47-48. 10.1080/165019701300006542.
[49] Cusick, A., McIntyre, S., Novak, I., Lannin, N., & Lowe, K. (2006). A comparison of goal attainment scaling and the Canadian Occupational Performance Measure for paediatric rehabilitation research. *Pediatric Rehabilitation*, *9*(2), 149-157. 10.1080/13638490500235581.
[50] Law, M., Baptiste, S., McColl, M., Opzoomer, A., Polatajko, H., & Pollock, N. (1990). The Canadian occupational performance measure: an outcome measure for occupational therapy. *Canadian Journal of Occupational Therapy*, *57*(2), 82-87. 10.1177/000841749005700207.
[51] Valderas, J. M., & Alonso, J. (2008). Patient reported outcome measures: a model-based classification system for research and clinical practice. *Quality of Life Research*, *17*(9), 1125-1135. 10.1007/s11136-008-9396-4.
[52] Valderas, J., Kotzeva, A., Espallargues, M., Guyatt, G., Ferrans, C. E., Halyard, M. Y., ... & Alonso, J. (2008). The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Quality of Life Research*, *17*(2), 179-193. 10.1007/s11136-007-9295-0.
[53] Klose, K., Kreimeier, S., Tangermann, U., Aumann, I., & Damm, K. (2016). Patient-and person-reports on healthcare: preferences, outcomes, experiences, and satisfaction–an essay. *Health Economics Review*, *6*(1), 1-11. 10.1186/s13561-016-0094-6.
[54] Lessard, L., Michalowski, W., Fung-Kee-Fung, M., Jones, L., & Grudniewicz, A. (2017). Architectural frameworks: defining the structures for implementing learning health systems. *Implementation Science*, *12*(1), 1-11. 10.1186/s13012-017-0607-7.
[55] Malec, J. F. (1999). Goal attainment scaling in rehabilitation. *Neuropsychological Rehabilitation*, *9*(3-4), 253-275. https://doi.org/10.1080/096020199389365
[56] Argyris, C. (1976). Single-loop and double-loop models in research on decision making. *Administrative Science Quarterly*, 363-375. 10.2307/2391848.
[57] Jaaron, A. A., & Backhouse, C. J. (2017). Operationalising "double-loop" learning in service organisations: a systems approach for creating knowledge. *Systemic Practice and Action Research*, *30*(4), 317-337. 10.1007/s11213-016-9397-0.
[58] Shaoibi, A., Neelon, B., & Lenert, L. A. (2020). Shared decision making: from decision science to data science. *Medical Decision Making*, *40*(3), 254-265. 10.1177/0272989X20903267.

[59] Foley, T. J., & Vale, L. (2017). What role for learning health systems in quality improvement within healthcare providers?. *Learning Health Systems*, *1*(4), e10025. https://doi.org/10.1002/lrh2.10025.
[60] Wennberg, J. E., O'Connor, A. M., Collins, E. D., & Weinstein, J. N. (2007). Extending the P4P agenda, part 1: how Medicare can improve patient decision making and reduce unnecessary care. *Health Affairs*, *26*(6), 1564-1574. 10.1377/hlthaff.26.6.1564.
[61] McGinnis, J. M., Saunders, R. S., & Olsen, L. (Eds.). (2011). *Patients Charting the Course: Citizen Engagement and the Learning Health System: Workshop Summary*. National Academies Press. http://hdl.handle.net/1903/23882
[62] Kamal, A. H., Kirkland, K. B., Meier, D. E., Morgan, T. S., Nelson, E. C., & Pantilat, S. Z. (2018). A person-centered, registry-based learning health system for palliative care: a path to coproducing better outcomes, experience, value, and science. *Journal of Palliative Medicine*, *21*(S2), S-61. 10.1089/jpm.2017.0354.
[63] Montori, V. M., & Guyatt, G. H. (2008). Progress in evidence-based medicine. *Jama*, *300*(15), 1814-1816. 10.1001/jama.300.15.1814.
[64] Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*(2), 230-253. 10.1518/001872097778543886.
[65] Madhavan, P., & Wiegmann, D. A. (2007). Similarities and differences between human–human and human–automation trust: an integrative review. *Theoretical Issues in Ergonomics Science*, *8*(4), 277-301. 10.1080/14639220500337708.
[66] Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*(1), 50-80. 10.1518/hfes.46.1.50_30392.
[67] Bussone, A., Stumpf, S., & O'Sullivan, D. (2015, October). The role of explanations on trust and reliance in clinical decision support systems. In *2015 international conference on healthcare informatics* (pp. 160-169). IEEE. doi: 10.1109/ICHI.2015.26.
[68] Stucki, G., Daltroy, L., Katz, J. N., Johannesson, M., & Liang, M. H. (1996). Interpretation of change scores in ordinal clinical scales and health status measures: the whole may not equal the sum of the parts. *Journal of Clinical Epidemiology*, *49*(7), 711-717. 10.1016/0895-4356(96)00016-9. nj
[69] Wright, B. & Stone, M. (1979). *Best test design*. MESA Press: Chicago, IL Retrieved from: https://research.acer.edu.au/measurement/1

Appendix A: Key Terms and Acronyms

Analytic Hierarchy Process (AHP)

Canadian Occupational Performance Measurement (COPM)

Center for Disease Control and Prevention (CDCP)

Chronic Obstructive Pulmonary Disease (COPD)

Clinical Decision Support System (CDSS)

Conjoint Analysis (CA)

Discrete Choice Experiments (DCE)

Design Science Research Methodology (DSRM)

Design Science Research Methodology (DSR)

Electronic Health Records (EHR)

Evidence-Based Medicine (EBM)

Goal Attainment Scale (GAS)

Goal-Oriented Care (GOC)

Gross Domestic Product (GDP)

Learning Health Systems (LHS)

MD Anderson Symptom Inventory (MDASI)

Patient-Reported Outcomes Measures (PROMs)

Public Health Agency of Canada (PHAC)

Specific, Measurable, Attainable, Relevant and Time-Specified (SMART)