Finding the Impact of Demand Variability in Youth Psychiatry by Workload Forecasting

Author:

A.J.M. Lohschelder	University of Twente

Supervisors:

Prof. Dr. Ir. E.W. Hans	University of Twente
Dr. M.D. Post-Hubers	University of Twente
Hans Attema	Company supervisor

Finding the Impact of Demand Variability in Youth Psychiatry by Workload Forecasting

Author	A.J.M. (Aletta) Lohschelder
Date	July 2023
Educational institution	University of Twente PO Box 217, 7500 AE Enschede Tel. +31(0)534899111
Educational programme	Bachelor of Industrial Engineering and Management
Supervisors **University of Twente**	 Prof. Dr. Ir. E.W. Hans Dr. M. D. Post-Hubers
Company supervisor	Hans Attema

Management Summary

Problem context

Company X, an organisation and academic centre for child- and youth psychiatry in the Netherlands has started an improvement programme to shorten the waiting lists of patients and create a less fluctuating workload for employees. The therapists at Company X are specialized in providing complex care and patients often receive personally adjusted therapy from multiple therapists simultaneously or sequentially.

Problem definition

The core problem of this research is the high level of demand variability in the treatment processes of psychiatric patients, which causes fluctuating workloads and long waiting lists. This research aims to answer the research question: *"What is the impact of youth mental healthcare demand variability on the required therapy capacity?"*.

Method

To solve the problem, we first present a quantitative and qualitative analysis of the current situation regarding demand variability at Company X. We then propose a forecasting model for a therapist's weekly workload based on the probability mass function of the Poisson binomial distribution. While this forecast gives an exact distribution of the weekly numbers of expected appointments, the runtime increases drastically as the caseload increases. Therefore, we apply the model on a small instance to show its potential for application in another context. In addition to this exact model, we also present a simulation-based approximate model using Monte Carlo simulation. This model forms the basis for a forecasting simulation model that we asses for application in Company X.

Results

The results of this research are the exact forecasting model and the simulation forecasting model. The exact forecasting model might be useful in other applications, but the simulation forecasting model is the most useful for Company X. We created a forecasting simulation tool in VBA for psychomotor therapists that outputs the forecasted workload in three ways: the number of weekly appointments, the probability distribution of the weekly workload in hours, and the probability distribution of the weekly workload compared with the available number of hours. Figure 1 displays the three outputs in three graphs on the tool's dashboard.

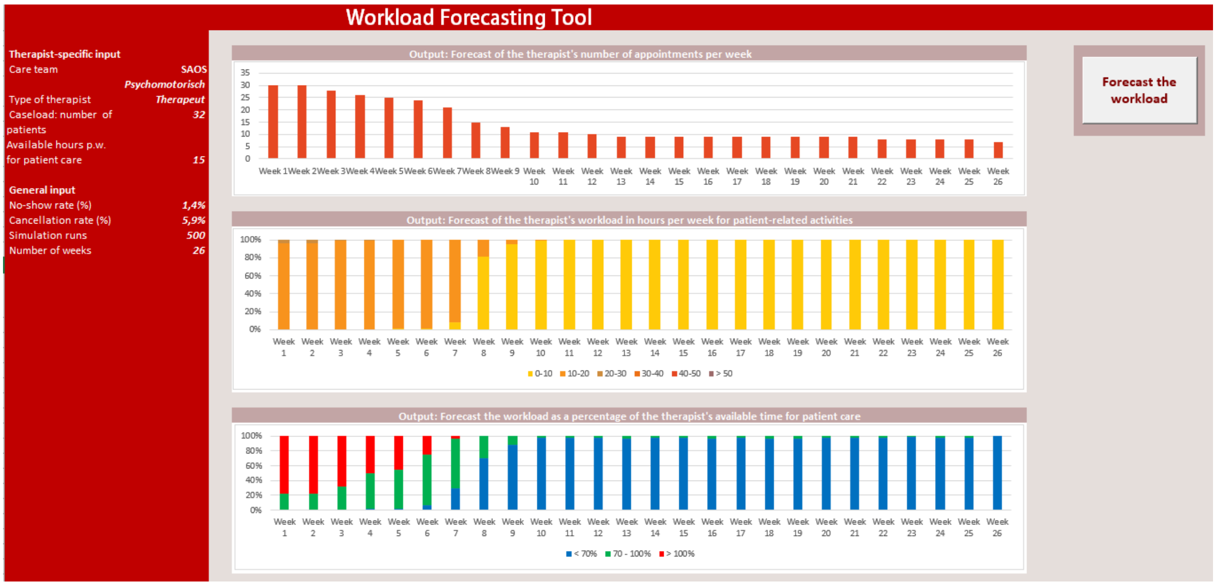

Figure 1: The dashboard of the workload forecasting tool.

The exact forecasting model precisely predicts the workload, while the simulation forecasting model closely approaches the workload as the number of simulation runs increases. To verify the forecasting simulation model, we compared the results of the exact model and simulation model. The models were compared for two samples of a therapist's caseload of 5 patients, with different starting conditions. We forecasted weekly workload for 12 weeks using the same input data for the two models and compared the results. Figure 2 shows the results of the first sample and Figure 3 of the second sample. Both figures show that the rounded results of the simulation model are similar to the results of the exact model, so the simulation correctly forecasts the workload for these two samples.

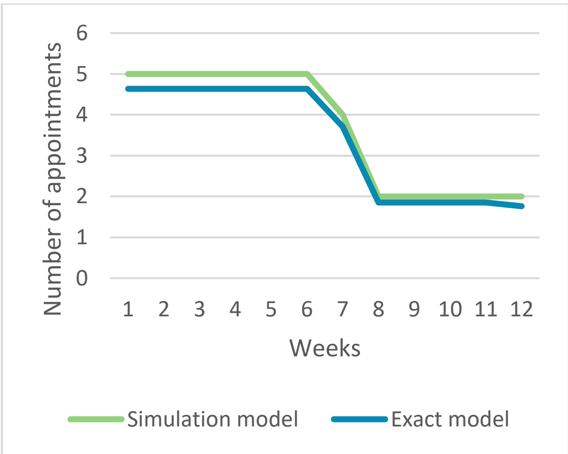

Figure 2: Comparison of the results of the exact model and the simulation model for the first testing sample.

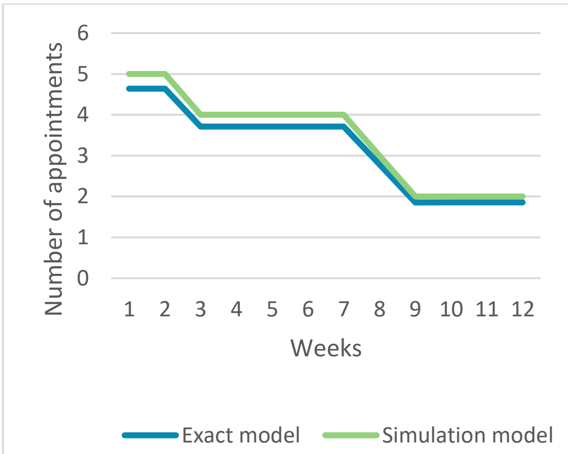

Figure 3: Comparison of the results of the exact model and the simulation model for the second testing sample.

Conclusion and discussion

This research was undertaken to evaluate the impact of the demand variability in youth mental healthcare on the required therapy capacity. We are the first to apply Monte Carlo simulation to forecast therapists' workload in a psychiatric context. Additionally, the core problem of this research is the high level of demand variability in the treatment processes of psychiatric patients, which causes fluctuating workloads, and long waiting lists. This is a problem that mental healthcare organisations throughout all of the Netherlands are currently facing. The research makes a practical contribution by creating a method to create a forecasting simulation model for Company X to implement to foresee when there is room in the workload for new intakes and when solutions are needed for therapists with an excessive weekly workload. The approach can be easily applied in other mental healthcare organisations.

We recommend Company X to promote a more uniform way of working among therapists of the same type, implementing standardized care pathways, and improving data tracking standards to enhance data quality and support data-driven decision-making. By implementing these recommendations and utilizing the forecasting simulation model, Company X can improve the stability of workloads, reduce waiting lists, and make data-driven based on quantitative insights about workload.

This thesis research has limitations due to the limited timeframe and data availability. Due to data availability and time constraints, we were not able to further validate the forecasting simulation model based on actual data and/or more data based on the exact model. Additionally, the absence of booking rules for all therapist types limited the development of a fully functional forecasting simulation model for all therapist types.

The forecasting simulation model can be further improved by extending the simulation model with features to schedule more than one appointment per week, reschedule cancelled appointments and distinguish between appointments in direct and indirect time. The simulation model's accuracy can be improved by including seasonality in the model.

Acknowledgements

I would like to express my sincere gratitude to everyone who has enabled me to successfully complete my bachelor thesis, which I proudly present to you here.

My gratitude goes out to everyone at the company who helped me to do this assignment. I am grateful to Hans, for all his time and explanations of how a psychiatric organisation works and his enthusiasm about the results of my thesis. Bas, for guiding me in structuring the research and fine-tuning the simulation model. Neeltje, for giving me access to all the data I needed, often within a day.

I also want to express my gratitude to my supervisors at the university for helping me in completing my first research. Erwin, for his enthusiasm, unwavering trust, and his assistance in structuring the model. Especially during the last weeks when I doubted whether I could complete the thesis in time. Mireille, for her interest and extensive feedback, which allowed me to see my research from a different perspective.

Lastly, I want to thank all my friends and family for their support and encouragement. Sharing my challenges and successes helped me to work hard and stay motivated.

Aletta Lohschelder

Enschede, August 2023

Table of Contents

1 Introduction

This chapter starts by introducing the research context by providing information about Company X in Section 1.1 and the company's improvement programme that this research contributes to in Section 1.2. Section 1.3 discusses the problem context, problem cluster, definition of the core problem, and the main research question. Section 1.4 explains the research methodology.

1.1 Introduction to Company X

Company X is an academic centre for youth psychiatry in the Netherlands. Patients are diagnosed and treated for autism, ADHD, anxiety, depression, obsessive-compulsive disorders, behavioural disorders, psychosis, and other psychiatric issues. Patients often have complex diagnoses consisting of multiple disorders, demanding care from multiple therapists. The company's therapists are highly educated and specialized and are thus given a high level of professional autonomy.

1.2 Introduction to the improvement programme

Company X started an improvement programme in collaboration with Rhythm. The improvement programme team consists of the company's employees and Rhythm's consultants. Rhythm is a spinoff by the research group CHOIR of the University of Twente and ORTEC. Rhythm uses academic knowledge to enhance logistics in healthcare organisations (Rhythm, n.d.). The three main goals of the improvement programme are to shorten access times for patients, to create a less fluctuating workload for employees, and to ensure healthy company operations. To achieve this, Integral Capacity Management (ICM) is being integrated into the organisation. ICM is a concept developed by the research group CHOIR (Centre for Healthcare Operations Improvement and Research) of the University of Twente. ICM is implemented in many healthcare organisations by Rhythm (CHOIR, n.d.). The implementation of ICM creates a better balance between demand and supply, because of the optimization of staff and resources (ORTEC, n.d.). One of the company's main aspects of the improvement programme is called '*ontzorgen,'* which is a term for the movement of the planning responsibility from therapists to planners. Thus, more available capacity of therapists can be spent on patient care instead of planning, while the capacity of therapists stays the same. This new planning procedure is currently being implemented in four care teams spread over two locations. This research is carried out at one of these two locations, from now on referred to as Location Y. The two care teams in Location Y that are involved with the improvement programme are SAOS and ADHD/ASS. The ADHD/ASS team is specialized in the treatment of ADHD and autism, and the SAOS team focuses on anxiety, depression, and other behavioural disorders.

1.3 Problem definition

As explained in Section 1.2, the improvement programme team is moving the planning responsibility from therapists to planners. During this process, it has become clear that the demand of patients is highly complex, which causes high demand variability. There are two sides to this variability. The first side of the variability is due to the complexity of patient's diagnoses. For more than half of the patients, the diagnosis consists of multiple disorders. Hence, patients often receive consecutive or simultaneous treatments from multiple therapists. The second side of the demand variability is caused by the low level of treatment standardisation. Company X has standardised care pathways, but therapists have a high level of autonomy and thus frequently modify these standard pathways, or create new pathways tailored to a patient's needs. Additionally, even appointments of the same type exhibit significant variability in terms of appointment duration and frequency. Therapists were inquired to determine whether this variability is solely caused by healthcare-related factors. It was found that healthcare-related reasons were often not the primary cause for variability. Therapists often do not see the benefits of a more uniform way of working, and have their own preferences for appointment duration, frequency and to what extent the standardized care pathways are utilised. This suggests that it is possible to implement a higher degree of therapy standardisation without compromising the quality of care. The high degree of demand variability at Company X makes it difficult to forecast demand, and it is thus more challenging to foresee opportunities and bottlenecks. The improvement programme aims to create more uniformity in the treatment processes of therapists without compromising on the quality of care.

Another factor driving the need for a more standardized way of working amongst therapists is the limited scalability of the current planning process. Taking over the planning responsibility from therapists will be a highly time-consuming and complex task for planners, if the planning procedure remains as it is. Personal preferences can still be considered to a certain degree, but the planning of the majority of appointments should go according to a set of booking rules. These booking rules are rules created by therapists about appointment duration and frequency per appointment type. While this will standardize the therapy processes, it is essential to maintain the therapist's professional autonomy. This is guaranteed by the remaining possibility to deviate from appointment norms. Nevertheless, therapists are encourages to follow the norms unless there is a valid healthcare-related reason to deviate rather than the other way around, which will create uniformity between therapists.

If therapists create a set of booking rules per type of therapist per care team that is accepted and used by all therapists, the level of variability in the therapy processes at Company X decreases. Forecasts about future patient demand can then be created to gain more insight into the expected workload and capacity utilization. This will enable improved planning abilities and improved response to bottlenecks, as these are not as unexpected anymore. Overall, it represents another step forward in the direction of Integral Capacity Management.

As mentioned, two of the objectives of the improvement programme are to minimise waiting times and create a less fluctuating workload for employees. The negative effects of waiting times and a high workload can be considerable. Substantial quality of life costs can occur if patients experience long waiting times, such as the obvious extended time in poor health, but also the anxiety of waiting, health possibly deteriorating even further, and losing confidence in the hospital or therapist (Vanberkel, 2011). Fluctuations in employees' workloads can have significant effects on employees' well-being (Wood et al., 2013). In general, there should be enough staff to be able to effectively handle workload fluctuations. This is crucial for maintaining efficiency, effectiveness, and safety (Gabbay & Bukchin, 2009).

The problem cluster in Figure 4 visualises the causes and consequences of the long waiting lists and fluctuating workloads. By going back into this causal chain, the causes of these problems become clear. As mentioned, patients at Company X have complex diagnoses and thus receive a personally adjusted treatment, often by multiple therapists. This complexity makes it difficult to predict the demand of psychiatric patients. Therefore, Company X lacks the knowledge of which capacity shortages are most significantly leading to longer waiting lists and where therapists structurally have time left. The complexity of the care is not a solvable problem, since therapists' ability to provide high-quality complex therapy is precisely where Company X excels. However, the limited amount of uniformity in therapy processes is also a cause of numerous problems in the organisation, including long waiting times, and fluctuating workload. This is a solvable problem, so we formulate the core problem of this research as follows:

"There is a high level of demand variability in the treatment processes of psychiatric patients, which causes fluctuating workloads, and long waiting lists."

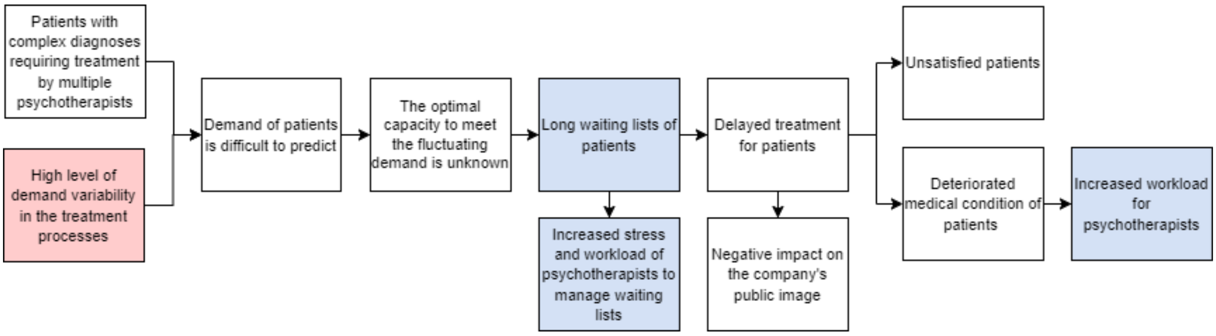

Figure 4: Problem cluster with the core problem marked in red and the main improvement objectives of the improvement programme marked in blue.

In order to measure the performance of the variables in the core problem, it is important to measure the norm and reality of the core problem against each other. When the improvement programme started in October 2022, the access times and workload fluctuation were measured. The norm is to treat 100% of patients within 14 weeks after application and have an average weekly workload between 70-100% of available hours for all therapists during every week. At the start, only 23% of patients were treated within 14 weeks and the weekly workload of just 47% of all therapists met the norm of having a workload in the range of 70-100% of the therapist's total available hours. The starting performance of these two variables is visualized in Figure 5 and Figure 6.

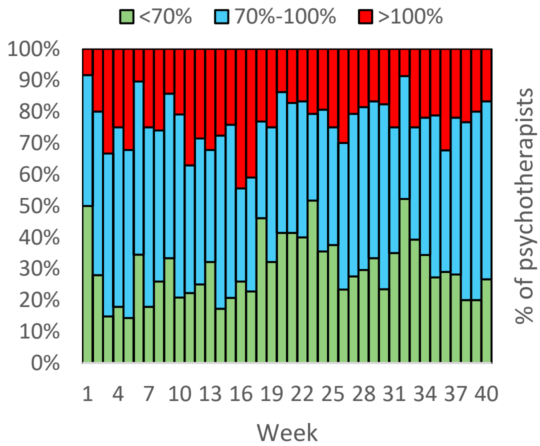

Figure 5: Average weekly workload of therapists from January to October 2022 (Company X, 2022).

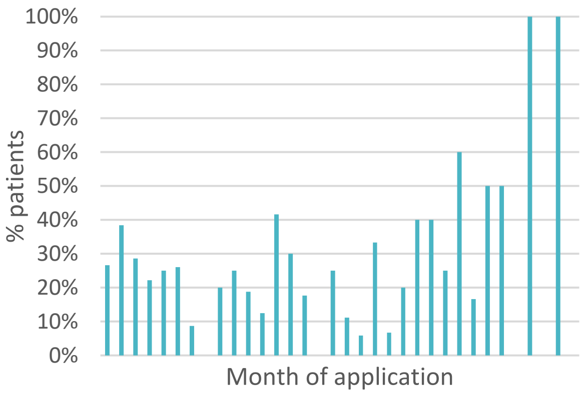

Figure 6: Access times per month of application between January 2020 and October 2022 (Company X, 2022).

To find a solution to the core problem, we pose the following main research question:

"What is the impact of youth mental healthcare demand variability on the required therapy capacity?"

By answering the research question, this research contributes to one of the company's goals in line with Integral Capacity Management, which is 'to improve integrated decision-making by creating more insight into the expected care demand and the required capacity'. The research is scoped to a research population consisting of all patients treated by psychomotor therapists in the SAOS care team at Company X at Location Y. We construct a generally applicable forecasting model for youth mental healthcare and then apply the model to a small research population, to show the model´s application. The thesis is scoped as such to ensure a possible completion within the limited timeframe of ten weeks.

1.4 Research methodology

This subchapter describes the deliverables in Section 1.4.1, the research questions in Section 1.4.2 and the research design in 1.4.3.

1.4.1 Deliverables

The following deliverables will be provided to answer the main research question.

- A general forecasting model that predicts a therapist's weekly workload.
- A applied forecasting model that predicts the weekly workload for a psychomotor therapist.
- A recommendation for Company X on how to implement the forecasting model into the organisation for all types of therapists and how to improve their approach to demand variability.

1.4.2 Research questions

The aim of the research is as follows:

"This research *aims to find the impact of youth mental healthcare demand variability on the required therapy capacity."*

To accomplish this goal, we follow these research questions:

1. What are the characteristics of the demand and capacity at Company X? (Chapter 2)
    i. What is the general patient journey?
    ii. How are the standardized care pathways constructed?
    iii. What is the total available capacity for direct and indirect patient-related activities?
    iv. What insights about demand variability can be derived from the data visualization?
2. How to create a forecasting model to predict a therapist's workload? (Chapter 3)
    i. What forecasting models can be used to predict patient demand in healthcare?
3. How to construct a forecasting simulation model to predict a therapist´s weekly workload? (Chapter 4)
    i. What is the impact of demand variability on a therapist´s workload according to the forecasting simulation model?

1.4.3 Research design

We use a mixed-methods approach is used for answering all research questions. We use qualitative data gained during meetings and conversations with employees from company X, and a quantitative data analysis to answer question 1. We answer questions 2 and 3 by conducting a qualitative literature study and constructing a model. The mixed-methods approach creates a more comprehensive and reliable understanding of the core problem and ensures that the research questions are answered effectively. The research designs have different advantages, which makes them complementary (Lund, 2012). Lund also states that this creates a higher validity of the research design compared to using a single strategy.

We examine the effect of demand variability on the required therapy capacity, so this is a explanatory research, since it examines the type of relationship between variables. This is researched by creating an exact model in Chapter 3 and a simulation model in Chapter 4. Additionally, there is a descriptive part of this research, including the literature studies.

This research aims to find a general method to create a forecasting model to predict a therapist's weekly workload and is applicable in youth psychiatry. We create a simulation model that forecasts a psychomotor therapist's weekly workload to illustrate how the simulation model can be expanded to a complete model for all types of therapists. The simulation modelling method is applicable to the company's other care teams, other locations and to other psychiatric care organisations in general. Psychiatric care organisations all over the Netherlands face the same issues as company X concerning long waiting lists and fluctuating workloads (Ministerie van Volksgezondheid, Welzijn en Sport, 2023). The research thus makes a scientific contribution by creating a forecasting method that is beneficial for other organisations with the same problem. Moreover, the research contributes to science by combining Monte Carlo simulation with workload forecasting in a psychiatric context, which has not been researched in previous studies. Searching on Scopus yielded significantly more results for optimization in hospitals compared to optimization in the field of psychiatry. This shows that there has been limited research conducted on optimization in therapy processes in psychiatry, despite its potential for significant improvements. This underscores the scientific relevance of this research. The practical relevance for Company X lies in the implementation of the forecasting simulation model and expanding it to be used for all types of therapists.

2 Characteristics of demand and capacity

This chapter examines the therapy processes and capacity characteristics at Company X, in order to answer the first research question:

> *"What are the characteristics of the demand and capacity at Company X?"*

Section 2.1 visualizes the patient journey and Section 2.2 illustrates the utilization and structuring of care pathways. Section 2.3 examines the therapy capacity and Section 2.4 performs a data analysis concerning demand variability in therapy processes in general and specifically for psychomotor therapists.

2.1 Patient journey

We visualized the patient journey in a Business Process Diagram in Bizagi using internal company data with the widely accepted Business Process Modelling Notation (BMPN). Business Process Diagrams are based on flowcharting techniques tailored especially for graphically visualizing business process operations (White, 2004). In this research, BPMN is used to create a Business Process Diagram of the patient journey in Figure 7, to create insight into the standard process patients follow from registration to deregistration, regardless of their diagnoses.

The pool in a Business Process Diagram represents the organisation, and the lanes the participants, so in this case, the patient and therapist. For the sake of simplicity, the general term therapist is chosen instead of defining more specific functions, since those specifications differ per patient and process. Furthermore, there are three types of flow objects: squares, diamonds, and circles. The activities are represented by squares, decisions by yellow diamonds and events by circles (White, 2004). The process starts and ends with a registration and deregistration event. The process then follows a standard sequence of activities and encounters decisions that determine the exactly followed process. This results in the Business Process Diagram of the patient journey in Figure 7.

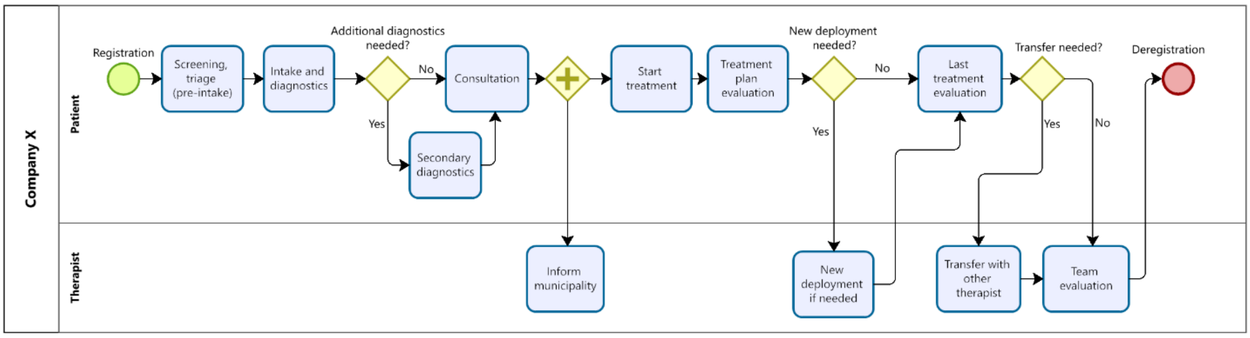

Figure 7: BPMN diagram in Bizagi representing the general patient journey as agreed to with Company X Gelderland.

2.2 Care Pathways

A care pathway is a sequence of modules designed by Company X's experts for a specific patient group. All care pathways that are part of SAOS and ADHD at location Y are listed in Table 1. The care pathways are created and verified using scientific psychiatric research. A patient's treatment programme can consist of multiple care pathways, especially if there are multiple diagnoses, as is often the case for Company X's patients. A care pathway consists of a goal, an end moment, consultation moments, a measurement of results and the therapy content. The therapy content consists of the activities that are offered to the patient, in what frequency and by which types of therapists. Standard care pathways contribute to the increased quality and effectiveness of healthcare. Using care pathways decreases variability between therapists and creates a cycle of constant measurement and improvement. Moreover, patients are provided with a better insight into their own treatment programme and it thus increases predictability and the ability to plan ahead.

Care pathways SAOS	**Care pathways ADHD/ASS**
Anxiety disorders	ADHD
Eating disorders	ASS
Behavioural disorders	
Personality disorders	
Psychotic disorders	
Mood disorders	
Somatoform disorders	

Table 1: Overview of all care pathways that are part of SAOS and ADHD/ASS.

Currently, both standard and generic care pathways are used by therapists. Standard pathways consist of a series of standard and optional modules. A generic care pathway, however, is a sequence of modules, tailored for a specific patient by a therapist. For all patients for which the care pathway was tracked between 2019-2022, generic pathways were used 48 percent of the time. Therapists are often using generic care pathways, as opposed to the standard care pathways created and validated by scientific research. The improvement programme team at company X proposes to return to essentially following the care pathways, unless the patient's needs do not fit in one of the many care pathways. The team suggests using generic pathways only 20 percent of the time, instead of the current 48 percent. Hereby the professional autonomy of therapists can be maintained, while creating a more uniform way of working.

A treatment programme consists of one or more care pathways, each characterised by a different sequence of standard and optional modules. Each module, in turn, encompasses one or more activities. An activity is one appointment. Figure 5 visualizes and clarifies this structure.

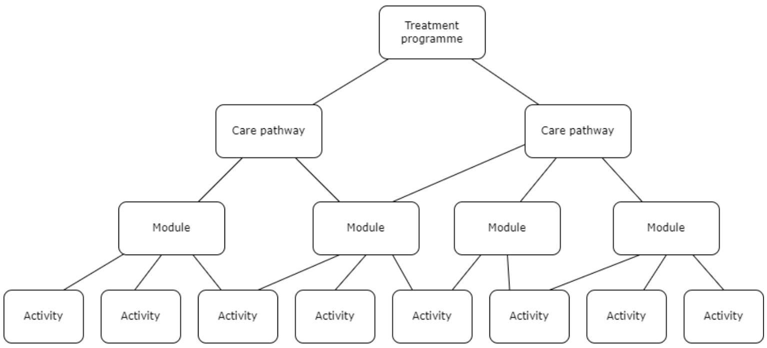

Figure 8: Diagram with relationships between all components of a treatment programme at Company X.

The care pathways are structured in decision trees by Company X, where the characteristics of a patient's diagnosis lead to a certain care pathway. Some of the modules are a standard part of the care pathway, while others are optional and only recommended to use if it offers additional value for a patient.

2.3 Total available therapy capacity and

The improvement programme team has developed a capacity overview for the care teams aligned with the improvement programme at Location Y, which includes all employees their corresponding number of working hours. This is also an overview the total available capacity in the ADHD/ASS and SAOS teams. For ADHD/ASS 87% of available capacity without research and education can be spent on patient-related activities. For SAOS this is 85% of the available capacity. The norm Company X set regarding the percentage of available time that has to be spent on patient-related activities is 86%, so the norm is currently met for ADHD/ASS but not for SAOS. Managers strictly monitor whether their employees meet this norm since there are small profit margins in youth psychiatry.

Company X is working towards their goal of a division of 50% indirect and 50% direct time. Currently, Location Y is not meeting this requirement for direct and indirect time yet, indirect time still exceeds direct time. The estimated ratio of indirect to direct time is currently 60-40, but this varies widely over the types of therapists. Some treatment teams structurally have a higher ratio than other teams. This is not problematic, as long as the average is met and therapists try to aim for the ratio, while maintaining a good quality of care.

The SAOS team currently consist of 25 therapists of 13 different functions of which 2 therapists are in training, and the ADHD/ASS teams consists of 12 therapists of which there are 8 different functions, and 3 therapists are in training. This wide range of different functions within teams shows how specialized therapists at Company X are. All different functions have differing capabilities, which is the set of activities that the therapist is educated to perform. Therapists with the same function sometimes have more capabilities than the 'standard' because of additional education. Figure 9 illustrates the proportions and types of functions per care team.

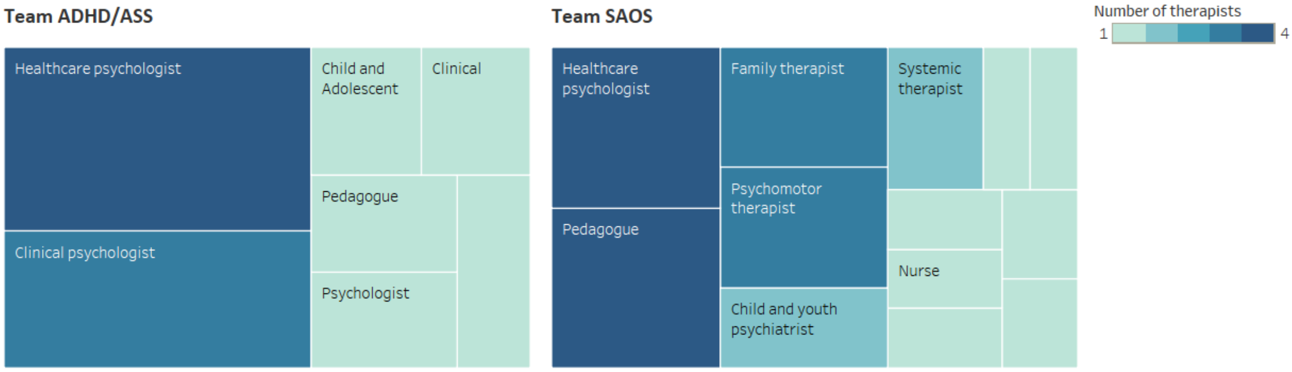

Figure 9: Visual overview of the types and amounts of therapists in the SAOS and ADHD/ASS teams in June 2023. The boxes represent the different types of therapists and the colour of the boxes the number of therapists of that type.

2.4 Planning procedure

Chapter 1 already shortly explained the transition of moving the planning responsibility from therapists to planners. This Section discusses the past, current and desired appointment planning procedure for Company X. Additionally, the required next steps and points for improvement are included for a smoother implementation of the new planning procedure.

Previously, therapists were responsible for scheduling the majority of appointments, with only a few exceptional cases where a dedicated appointment planner would handle the scheduling. As a result, therapists would sometimes spend a lot of time mailing or calling back and forth with patients, especially for appointments with multiple patients or therapists involved. When the improvement programme started, one of its goals was transferring the planning responsibility from therapists to planners entirely. Currently this is only the case for a small group of therapists, as more planners are needed to completely implement this change.

Hereby, several recommendations are given to Company X for the successful implementation of this change. The first recommendation involves a shift in the mindset of therapists. Therapists are still hesitant to transfer their planning responsibility to a planner, despite the fact that most therapists feel limited affinity towards planning tasks. To overcome this, it is advised to transition slowly, to show the benefits to therapists and patients, and to create familiarity between planners and therapists. Additionally, therapists should request planners to plan in appointment series weeks or months in advance, rather than on a weekly basis, as most therapist do now. By planning appointment series, less flexibility is needed from patients, since appointments are scheduled further in advance.

Another necessity to successfully implement the new planning procedure is a clear set of standard booking rules. This creates uniformity in the way of working amongst therapists and allows planners to structurally schedule appointments far ahead. Booking rules are sets of planning rules on how to schedule a type of module or activity. The appointment frequency, duration, present types of therapists, as well as other relevant information, are included in this set of rules. The booking rules need to be accepted by all relevant types of therapists, so this needs to be established during a meeting with the vast majority of all therapists present, just as their manager and a planner. Some distinctions can be added per the experience level of the therapists, especially for appointment duration, but all other characteristics can be equal for all relevant therapists.

In conclusion, the recommendations regarding the shift in mindset, planning in advance and establishing booking rules, are all necessary steps for implementing the new planning procedure. The second and third step are also required for the implementation of a workload forecasting model.

2.5 Data visualisation

In this subchapter we will perform a quantitative data analysis about the causes of the variability in the treatment processes at company X. This data visualisation consists of a data quality assessment in Section 2.5.1, after which the data visualisation method is explained in Section 2.5.2. The different data subjects are then separately discussed; appointment duration in Section 0, appointment frequency in Section 2.5.4 and care pathways in Section 2.5.5. Lastly, quantitative data for psychomotor therapists is analysed to be used in the forecasting model in Section 2.5.6.

2.5.1 Data assessment

The data used for this research was obtained in Excel format from the Business Intelligence department. The department already cleaned the data and extracted the data directly from the database. However, two separate data files were created, because of a missing data connection element between care pathways and appointments, so the data files cannot be linked. The care pathways can thus only be linked to patients, and not to appointments. Consequently, a good analysis of the relationship between care pathways and appointments was challenging. Therefore the data about pathways and appointments were analysed separately. Additionally, the data structure was adjusted to create more data visualisation possibilities.

Historic appointment data is used for Company X between 2019-2022, for the care teams SAOS and ADHD/ASS. To safeguard the confidentiality and privacy of patients and therapists, their names have been replaced with anonymized key values. This follows Company X’s policy regarding data usage.

As a result of using data between 2019 and 2022, the data during the COVID-19 pandemic is included, which might negatively influence the data analysis. We decided to use this dataset with a span of four complete years by selecting the most recent data available. Because of the many changes over the last years in therapy processes at Company X, we estimated that the use of recent data would be more beneficial for the data quality than excluding the COVID-19 pandemic by using less or older data.

2.5.2 Data visualisation method

A wide range of data visualisation programmes exist that make it easy to display data and communicate complex relationships. Data visualization is an important part of quantitative research, as is also confirmed by literature in Section 3.1.1. It improves data-driven decision-making and can be used as input for the forecasting model.

Two data visualisation programmes taught during the IEM bachelor programme are RapidMiner and Tableau. After conducting tests with both programmes to create graphs for visualising data, Tableau was selected as the preferred option. Tableau easily creates interactive dashboards to compare graphs, and it is possible to apply filters on graphs to explore relationships in the data. We selected Tableau over RapidMiner as there seemed to be more room for flexibility and creativity in creating dashboards.

The main purpose of the data visualisation is to find the sources of demand variability at Company X. For this purpose, the standard deviation is often used as a measure. The standard deviation is the square root of a value, found by comparing data points to the population mean. This commonly used measure for quantifying the dispersion of a dataset creates an understanding of variability by measuring the extent to which data points deviate from the average.

We will analyse two aspects of the variability in therapy processes at company X; appointment duration variability and appointment frequency. The data will be analysed, as well as a comparison with the appointment norms. We compare the results of the quantitative data analysis with relevant qualitative information that we gained during conversations and meetings with therapists, planners , managers and other members of the improvement programme team.

2.5.3 Appointment duration variability

Most therapist do not strictly adhere to the appointment duration norms and the norms are also often not used by the planners themselves. The appointment duration is not primarily based on the norms, but also often on therapists' preferences, resulting in unpredictable variability in appointment duration. This statement is supported by the data presented in Figure 10. The figure visualises the differences between appointment norms and average activity duration, with separate graphs for direct and indirect time. It is evident that the appointment duration norms per activity are indeed not strictly followed, as Figure 10 shows that differences of half an hour between norms and averages are not uncommon. The activities where norms are not tracked are excluded, so the graph cannot be used as if it provides a complete image, but at least for the 27 out of 79 activities that are displayed in Figure 10, we can conclude that the appointment duration norms are not strictly followed. This high variability in appointment duration variability makes it difficult to accurately forecast demand, and precisely predict bottlenecks and opportunities.

Figure 10: Comparison of appointment duration averages and norms, between 2019-2022 for 27 activity types. The first graph depicts the differences in direct time and the second graph for indirect time. Note that the vertical appointment duration scales are not the same for both graphs.

Figure 10 clearly displays differences between norms and averages, but the causes of the differences are difficult to directly derive from the quantitative research. Therefore, we have listed the causes of the duration variability according to the conversations and meetings with employees of company X. The primary cause for duration variability for indirect time is the experience level of the therapist, since more experienced therapists need less time for appointment preparation and documentation. Therapists' working methods also play a role, especially their efficiency and precision determine the required amount of time to prepare and document appointments. For direct time, variability often arises from the busy schedules of therapists and patients, which can lead to shorter appointments being planned, to prevent that patients need to wait longer for their appointments. Therapists' differing working methods can also play a role here, for example, some therapists let patients prepare more for their appointments, which can allow for shorter appointments. Additionally, the presence of a therapist in training typically leads to shorter appointment durations because of the additional help. A planner also mentioned that videocall appointments tend to be shorted than face-to-face appointments of the same activity type. Figure 11 compares the average appointment time per activity of face-to-face appointments and videocall appointments. While the planner's claim holds for most activities, there is no substantial difference, as the average face-to-face appointment is 2,3 minutes longer than the average videocall appointment. When zooming in on the differences per activity type, Figure 11 shows the differences between the average face-to-face and videocall appointment per activity type. It can be concluded that for the activities which both had videocall and face-to-face appointments between 2019 and 2022, the duration of videocall appointments was indeed shorter for 68,6% of all activity types.

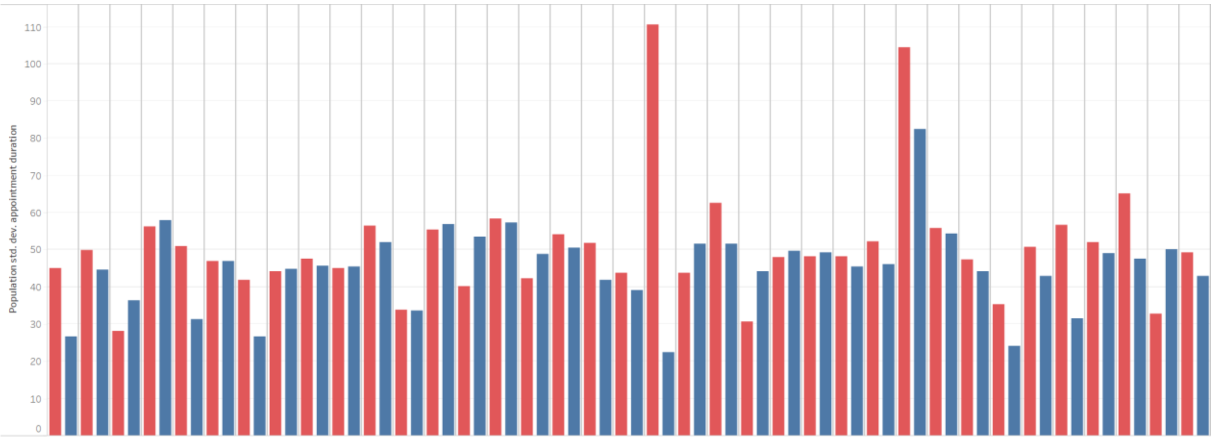

Figure 11: Comparison of the average appointment duration in minutes of face-to-face (red) and videocall (blue) appointments per activity type displayed in columns. The activity types that did not have both contact types in 2019-2022 were excluded, as this is invaluable data.

Figure 12 depicts a the duration variability among contact types, by showing the average duration and the standard deviation. From the standard deviation in Figure 12, we can conclude that the most variability originates from face-to-face appointments with patients, while there is less variability in contact via mail, telephone, WhatsApp, or videocalls. Face-to-face appointments also have the longest duration on average.

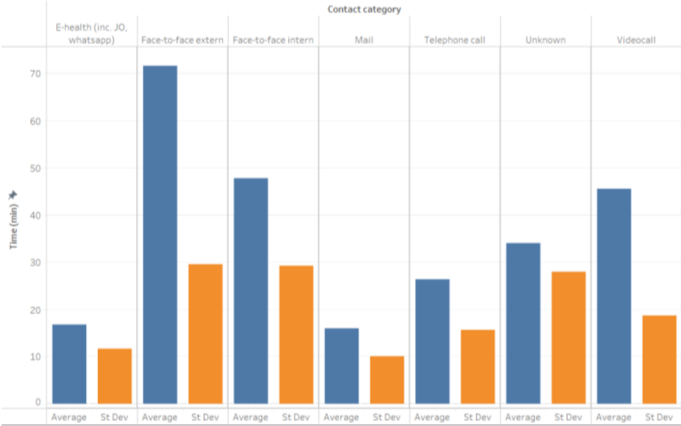

Figure 12: Comparison per contact category of average appointment duration (in blue) and the standard deviation (in orange) in 2019-2022.

The appointment duration also depends on the type of therapist linked to the appointment. Figure 13 shows the appointment duration average and standard deviation per type of therapist. A clear difference in averages is evident, as the average durations range from a minimum of 27,9 to a maximum of 54,3 minutes. The standard deviation also depends on the type of therapist, varying between 18 and 48 minutes. This shows that if historic data is used as input of a workload forecasting model, it is important to use appointment duration data for that type of therapist to obtain valid forecasting results.

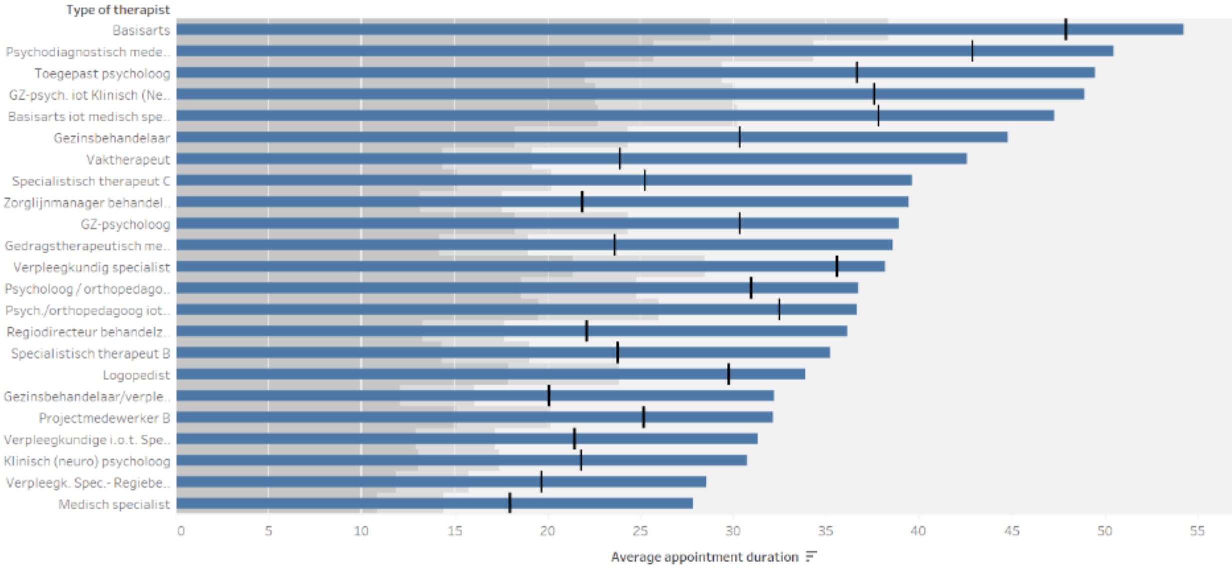

Figure 13: Comparison of the average duration (blue bars) and the standard deviation (vertical lines per bar) per type of therapist in 2019-2022. Therapists who have had fewer than 15 appointments or worked less than 5 hours during this period have been excluded, as their inclusion is not valuable, and the graphical representation is enhanced by their exclusion.

2.5.4 Appointment frequency variability

Therapists generally aim to see their patients with a standard frequency, for example once every one or two weeks. Descriptions of the prescribed appointment frequency per activity can be found on Company X's intranet for the majority of care pathways. Verifying whether this frequency is actually met for all 73 modules and 79 activities is a too time-consuming task to examine for this research. Moreover, conversations with Company X’s planner indicates that there is indeed variability in the appointment frequency, so not all appointments adhere to the frequency norms. There are various reasons for a disturbance of the frequency, therefore no strict frequency patterns can be detected in the appointment data.

Various causes of the variability in appointment frequency were found during conversations and meetings at Company X. For instance, a therapist's or patient's limited availability can cause appointments to be scheduled with shorter or longer intervals in between. Furthermore, already scheduled appointments can be cancelled by either the therapist or the patient, e.g. due to illness. For a therapist, there could also be an unexpected crisis with another patient causing the need to reschedule. Also, appointment frequencies differs depending on the type of therapy the patient receives.

If appointments are already scheduled, it is still not certain that the appointment will actually take place. Patients cancel appointments for various reasons, either within the acceptable time frame or at short notice. Additionally, there are patients who fail to show up for their scheduled appointments. The uncertainty caused by cancellations and no-shows causes frequency variability. By tracking this data, Company X gathered percentages for the cancellations and no-shows for Location Y between 2011 and June 2023, as presented in Table 2. This provides insight into the exact effect of cancellations and no-shows on appointment frequency. This is also useful input data to increase the accuracy of the forecasting model.

Appointment status	**Percentage of total appointments**
Present	93,7%
Unknown	0,1%
No show	1,4%
Cancelled in time	3,9%
Cancelled too late	2,0%
Total Appointments	100,0%

Table 2: The appointment status percentages at Location Y from 2011 to June 2023.

2.5.5 Variability in the use of care pathways

We analysed the variability in the use of care pathways, since Company X is interested in the extent to which therapists use care pathways. The first notable result is that with 48,0%, almost half of the assigned pathways to patients are generic care pathways. A common cause of the large share of generic pathways is the fact that not all pathways can be selected in the system once the patient is in treatment by a certain care team. Since Company X is specialized in its transdiagnostic work approach, this flaw in the system can be improved by the ICT department, in order to improve the quality of monitored data. Besides, it can be useful to remind therapists that the care pathways are constructed using scientific psychiatric research, and its effectiveness has thus been proven. This can encourage the use of the standard pathways for therapists.

Then we performed a data analysis to investigate the use of care pathways in Tableau. The generic care pathway is used significantly more than the standardized pathways, with a frequency of 2646, which is 48,0% of the total pathways between 2019-2022 in the SAOS and ADHD/ASS teams in Location Y, while the frequency of the standardized pathways ranges between 1 and 183 (3,3%). One of Company X's therapists expressed the efficiency of the standardized care pathways, because of their construction and validation by scientific research, and stated that increasing their utilization would improve the quality of care. The company's business manager agreed and suggested aiming for an 80-20 ratio between standardized and generic pathways. This confirms that the 52,0% percent of generic pathways is not a desired percentage. There is a total of 116 different care pathways used during the 3-year period, while there are only 64 care pathways explained on the care programme page on Company X's intranet. The frequencies of the use of the care pathways in 2019-2022 in teams SAOS and ADHD/ASS are also depicted in Figure 14.

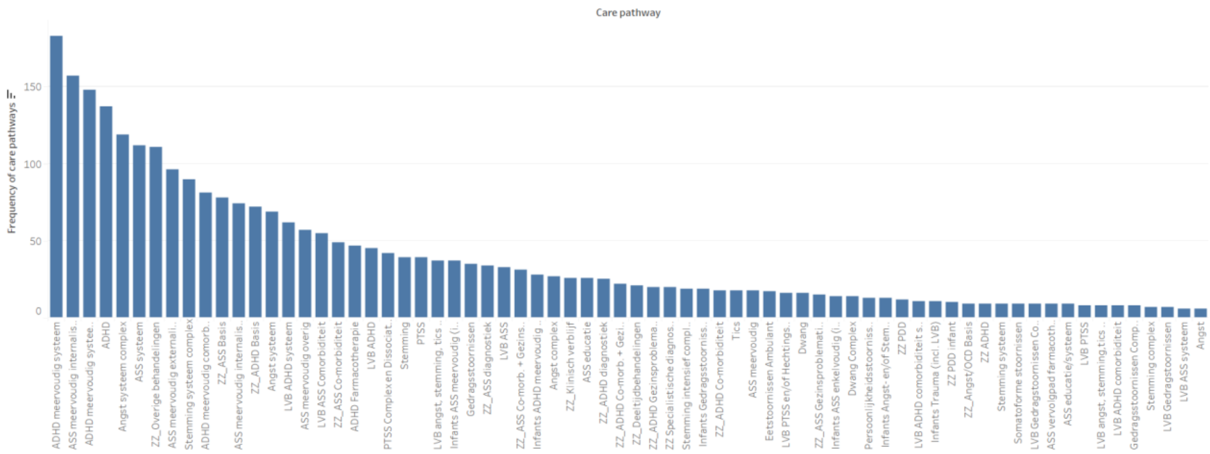

Figure 14 frequency of care pathways between 2019-2022 in teams SAOS and ADHD/ASS at Company X. The generic pathway and the 6 pathways where frequency is lower than or equal to 5 are excluded, to depict a clearer image of the distribution.

All SAOS and ADHD/ASS patients between 2019-2022 receive a treatment consisting of a number of care pathways ranging between 1 and 13, as is displayed in Figure 15. This data is useful for the construction of transition probabilities of the transitions between care pathways based on historic data since this is a useful input value for forecasting a patient´s treatment path from registration to deregistration. The histogram in Figure 15 also shows that with 53,7% the majority of all patients follow a total of two care pathways. Additionally, the numbers of patients drop exponentially as the number of pathways increases, with the exception of the patients that follow one pathway. Only 7,0% of all patients follow more than 4 pathways.

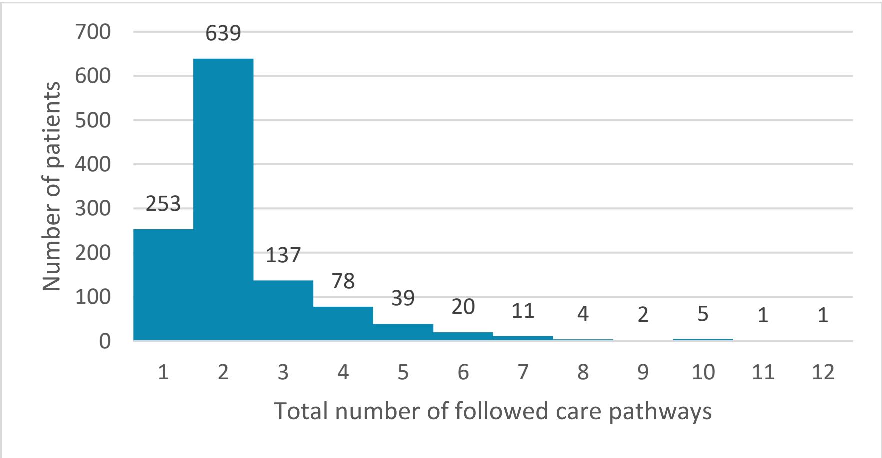

Figure 15: Histogram of the frequency of care pathways that are followed by a patient, between 2019-2022.

To conclude, 48,0% of all pathways are categorized as generic pathways. There are 115 other pathways tracked between 2019 and 2022, while there are only 64 standardized pathways on Company X´s intranet. In order to improve the data quality of patient demand, we recommend Company X to convey the importance of carefully and correctly tracking the pathways, including the termination of the pathway when it is finished to therapists. By implementing guidelines on documenting pathways, tracking patients' progress will be easier and it improves the accuracy of demand forecasting.

2.5.6 In-depth analysis of psychomotor therapy

This Section inspects all aspects of the general data visualisation for the specific case of patients who are treated by a psychomotor therapist. This will be used as input for the exact forecasting model in Chapter 3 and as input for the simulation forecasting model in Chapter 4.

2.5.6.1 Care pathways

A psychomotor therapist specified the lengths and types of appointments in their care pathways, to be used for the planners when taking over the planning responsibility from the therapists. By including Company X's total cancellation rate of 5,9% (too late and in time) and no-show rate of 1,4% this results in the rounded probability that an appointment takes place of 0,928. The lengths of the care pathways range from 7 to 12 weeks. Combining the care pathways lengths and the appointment probabilities results in the overview in Table 1Table 3. Table 3 also indicates whether a pathway is a termination pathway, so if the patient's treatment ends after following this care pathway. The type of activity that occurs every week is not particularly relevant to the output, because the number of appointments will be the same. By aggregating over all appointment types, the forecasting accuracy in Chapter 0 and Chapter 3 can be improved, since aggregated demand tends to exhibit less variability than demand observed at a more individualistic level (Dekker et al., 2004).

Care pathway	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12
Start PMT	0,928	0,928	0,928	0,928	0,928	0,928	0,928	0,928	0,928	0,928	0,928	
Start PMT (end)	0,928	0,928	0,928	0,928	0,928	0,928	0,928	0,928	0,928	0,928	0,928	0,928
Follow-up PMT	0,928	0,928	0,928	0,928	0,928	0,928	0,928					
Follow-up PMT (end)	0,928	0,928	0,928	0,928	0,928	0,928	0,928	0,928				

Table 3: Care pathways for a psychomotor therapist with probabilities that an appointment occurs for each week.

2.5.6.2 Appointment frequency

The no-show and cancellation rates from Section 2.5.6.2 are used as input for all types of therapists since these are tracked overall, instead of on an individual level and we assume these are similar for the patients of most types of therapists.

All care pathways with a standard or an optional module involving a psychomotor therapist are examined in terms of their appointment frequency. Out of all pathways, twenty of them offer an activity sequence that is executed by a psychomotor therapist. The module is an optional module in all pathways, so it is only offered when the therapist responsible for determining the patient's treatment programme deems it valuable. There is no recommended frequency given for the activity. However, in all other cases where the appointment frequency is specified, it is always every week for a sequence of ten weeks. However, even though the appointment variability is structurally set at once every week, practice can be very different, as explained in Section 2.5.4.

But since the exact frequency variability is unknown, a frequency of once every week for a series of activities performed by a psychomotor therapist seems to be a good input for a forecasting model. Especially since Company X is working on decreasing its demand variability, the actual values will keep approaching the given frequencies.

2.5.6.3 Appointment durations

Appointment duration varies widely among different types of therapists, as explained in Section 0. Therefore, using a general statistical distribution of appointment durations does not give a good image of the distribution for the duration of appointments of psychomotor therapists. As input values of the forecasting simulation model, a statistical analysis of the appointment durations based on historic data is needed. This resulted in the histogram in Figure 16.

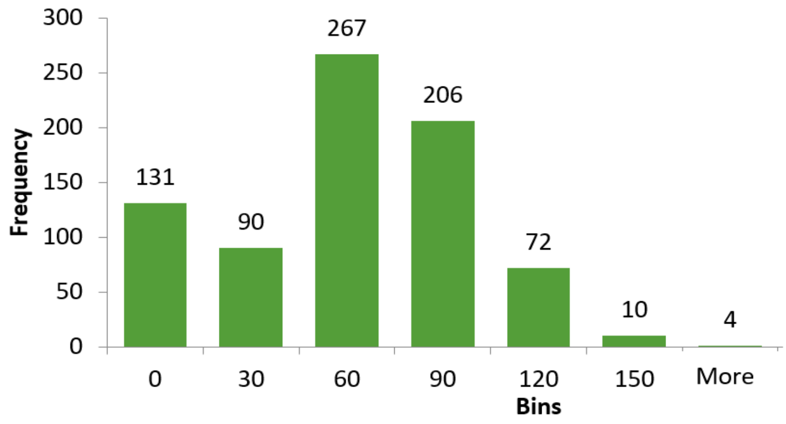

Figure 16: Histogram of the appointment durations of a psychomotor therapist of team SAOS between 2019-2022 at location Y.

The appointment duration data was extracted from the total dataset of appointments in the time range 2019-2022 for care teams SAOS and ADHD/ASS. The data was then filtered for the SAOS care team, module Psychomotor Therapy PMT and the appropriate type of therapist, such that only activities performed by psychomotor therapists who are parts of SAOS are considered when determining the distribution of appointment durations.

We performed the descriptive statistics data analysis in Excel, from which we derived that the average appointment duration is 54,0 minutes, the mode and median are both 60 minutes and the range in which all duration fall is 225 minutes. Subsequently, the statistical distribution of the data points was found by taking bins of 30 minutes and computing the frequencies and probabilities per bin. This results in the distribution visualized in the histogram in Figure 16. Bin sizes are often based upon rules of thumb, such as Sturges' rule (taking $1 + \log_2 n$ bins) or by choosing the bins such that there are at least k observations in each bin, where k varies per author. There are also more complex methods that determine the optimal bin width, making a tradeoff about the optimal amount of detail shown in the histogram (Wand, 1997). We chose the bin size not by aiming to visualize the actual distribution as closely as possible, but by choosing a distribution that is useful for the forecasting tool and makes sense for the therapists and tool users. Therefore bins of half an hour each are chosen, which shows the distribution and includes a level of detail, but not so much that it unnecessarily complicates the calculations.

The assumption is made that the values for which the appointment duration is zero, are cancelled appointments, so these values are excluded from the distribution. The statistical distribution and its probabilities that follow from the histogram in Figure 16 are then used to simulate the therapist's weekly workload in hours. This results in the probabilities in Table 4 for the psychomotor therapist. Only probabilities greater than 0,01 will be used in the VBA code, for the sake of simplicity and the insignificance of inclusion.

Probability	0,14	0,41	0,32	0,11	0,02	0,00
Range: start	0	30	60	90	120	150
Range: end	30	60	90	120	150	180

Table 4: Probabilities the appointment duration of a psychomotor therapist falls in the range between the range start and end value, based on data between 2019-2022.

2.6 Conclusion

This chapter answers the first research question: *"What are the characteristics of the demand and capacity at Company X?".*

The different sub-questions are answered in corresponding sections. The general patient journey is visualized in Figure 7. The therapists' working method consisting of standardized and generic pathways is then explained, as well as the different types of therapists in the SAOS and ADHD/ASS teams.

Next, it was analysed how the transition of moving the planning responsibility from therapists to planners can be successfully executed. The recommendations to Company X include creating a shift in the mindset of therapists, establishing clear booking rules and planning appointment series in advance instead of on a weekly basis. This will create insight into weeks with room for new intakes or an overflowing workload.

In this chapter, a data analysis and visualization were performed in the context of patient demand variability at Company X. We assessed the quality of the data and explored the data visualization method using Tableau. Our analysis focused on appointment duration variability, appointment frequency variability, and the variability in the use of care pathways.

Regarding appointment duration variability, it was found that therapists do not strictly adhere to the appointment duration norms, resulting in significant differences between the norms and the actual average appointment durations. The causes of this variability differ between direct and indirect time, with therapist experience and working methods playing a significant role. Appointment duration was also observed to vary depending on the type of therapist and the mode of appointment. For appointment frequency variability, various reasons for disturbances in the prescribed appointment frequency were identified, including therapist and patient availability, cancellations, and no-shows.

In terms of the use of care pathways, half of the tracked pathways are generic pathways. We recommend Company X to encourage therapists towards the increased use of standardized pathways. This will achieve a uniform way of working and could improve the quality of care. Additionally, the documentation was found to be suboptimal, indicating the need for Company X to create data-tracking guidelines and express their importance, in order to enable data-driven decision-making.

Variability in appointment frequency and duration were found for appointments that involve psychomotor therapist, as these can be used as input data for the forecasting model and simulation model.

3 Forecasting model

This chapter answers the second research question:

> *"How to create a forecasting model to predict a therapist's workload?"*

In order to answer this research question, we present a literature study in Section 3.1, consisting of a study of optimization modelling in healthcare and forecasting models. Section 3.2 describes the exact forecasting model and Section 3.3 shows the application of the model to a small sample.

3.1 Literature study of forecasting models

3.1.1 Optimization modelling and decision-making in healthcare

Logistical optimization has a long history of success stories in manufacturing. This differs considerably from the healthcare industry, because of the medical autonomy of therapists, differences between care processes and supply chains and the fact that patients cannot be treated in the same manner as physical products (Vanberkel, 2011). The context, priorities, predicament, and decisions are unique for every patient (Plsek & Greenhalgh, 2001). Often management in healthcare mainly focuses on the performance of individual departments, but not on the entire supply chain. This results in diminished access of patients without a significant cost reduction. The high level of complexity and uncertainty in healthcare makes it difficult for managers to use a systems approach in decision-making. Operations research model can be a guide towards improved and data-driven decision-making in the presence of uncertainty and complexity (VanBerkel, 2011).

Data-driven decision-making relies on data analysis rather than solely on intuition to enhance the operational performance of a company or organisation (Provost & Fawcett, 2013). In a study on the effects of data-driven decision-making, it was statistically shown that there is a positive effect between the level of data-driven practises in a company and its productivity (Brynjolfsson et al., 2011). This research was performed in a business context, in healthcare however, it is important to maintain a good balance between basing decision-making on guidelines, models, and protocols, while weighing the effects on employees' and patients´ experiences (Williams, 2002). This is confirmed by Ham & Alberti, who state that it is important to preserve the autonomy of therapists to prevent therapists from blindly adhering to established regulations (2002).

In order to handle the complexity and uncertainty in psychiatric organisations, it is essential to be aware of the high demand variability and to actively monitor emerging opportunities and patterns. By reacting flexibly and utilizing therapists' autonomy, complexity can be managed effectively (Williams, 2002). Demand forecasting in psychiatric organisations can mitigate the impact of complexity and uncertainty, by predicting emerging opportunities and bottlenecks.

3.1.2 Forecasting models

There are two types of forecasting methods; quantitative and qualitative forecasting methods. Qualitative forecasting is often used in situations with limited historic data, since the forecasts are based upon knowledge of experienced managers. Quantitative forecasting can only be applied when the situation complies with the following three conditions: historical data is available, the data can be quantified as numerical data and there is an assumption of continuity of historical patterns (Makridakis et al., 2008). Since the data in this research meets these three conditions and quantitative forecasting is generally more accurate (Armstrong, 2001), quantitative forecasting will be utilised in this research.

Data visualization plays a crucial role in identifying a suitable forecasting method, as it not only reveals historical data patterns that influence future data behaviour but can also suggest relationships between variables (Armstrong, 2001). Data visualization mining techniques are particularly valuable for exploratory data analysis, especially when dealing with large datasets. Interaction and distortion of the data is an important aspect of the visualization. Three steps are usually followed in the data visualization process: creating an overview, zooming and filtering, and providing details-on-demand (Keim, 2001). The details-on-demand technique is an interactive interface where a user can select data aspects to be visualized in more detail (Rauschenbach & Schumann, 1999). While Company X has created a data visualizations in PowerBI, this was primarily from a capacity perspective, whereas the focus of this research lies in analysing demand variability. Therefore, we perform a new data visualization from a demand perspective in Section 0.

After the data visualization phase, an appropriate quantitative forecasting method will be selected for the situation at Company X. Statistical quantitative forecasting can be classified into two methods: causal forecasting and time series analysis (Ivanov et al., 2018). In causal forecasting, the relationships between independent and dependent variables are measured. Causal forecasting is most accurate when there is a strong correlation between the variables (Davis, 2012). Time series analysis is based on historical data and assumes that factors influencing the past and present will continue to influence the future. For data with constant demand, types of time series forecasting such as Moving Average or Simple Exponential smoothing can be used. If there is a trend present in the data, time series forecasting techniques such as Double Exponential Smoothing or Regression Analysis are applicable (Ivanov et al., 2018).

Another quantitative forecasting method that is especially used in healthcare is a model created by Peter Vanberkel in his PhD thesis (2011). Vanberkel modelled workload in hospital departments because of recovering surgical patients. The analytical model is based on queueing theory, so it uses arrival rates for the number of surgeries in an operating room(OR) in a certain block in the Master Surgery Schedule (MSS) and service times for a patient's length of stay in the hospital ward. The distribution of the number of patients in the system on each day of the MSS is the main output of the model (Vanberkel, 2011).

The previously discussed models are all deterministic forecasting models; however, forecasting can also be performed using probabilistic forecasting models. The results of deterministic models are given as discrete values. Probabilistic models however, present output as a distribution, in e.g. quantiles or prediction intervals. This creates a more comprehensive understanding of the forecasted results compared to conventional deterministic models, offering an improved perspective for data-driven decision-making. Yet, there are also advantages of deterministic forecasting over probabilistic forecasting. Evaluation is an important aspect of creating a forecasting model. Deterministic forecasts can be evaluated using straightforward measures such as the Mean Squared Error and the Mean Absolute Error (Bazionis & Georgilakis, 2021). For probabilistic models, such measures are not sufficient, and a more complex evaluation method based on the statistical scoring rule theory is needed (Astfalck et al., 2023). For a long time, deterministic forecasting models were the main research focus, and these are still continuously being developed aiming to improve their accuracy. In recent years, the emergence of probabilistic models has allowed for a wider view of possible forecasting outcomes, by observing and studying the uncertainty of forecasts (Bazionis & Georgilakis, 2021). This transition is happening across different scientific disciplines (Gneiting & Katzfuss, 2014).

Forecasting can thus also be approached using a probabilistic model. If the demand can be characterized as a series of independent Bernoulli trails that are not necessarily identically distributed, the demand can be forecasted using the Poisson binomial distribution. The probability mass function of a Poisson Binomial distribution is a discrete probability distribution of a sum of independent Bernoulli trials, where the trials are not necessarily identically distributed, unlike the identically distributed Bernoulli trials of a binomial distribution (Wang, 2023).

An attempt was made to apply Vanberkel's model; however, due to the contextual differences, it was found to be an inadequate fit, primarily attributed to the disparity between the utilization of the MSS in Vanberkel's model and the unrestricted distribution of patients throughout the week at Company X. The approach was then adjusted from a deterministic to a probabilistic model, and the Poisson Binomial distribution was applied to create an exact probabilistic forecasting model.

3.2 Model description

Company X could significantly enhance its decision-making processes by transitioning to data-driven decision-making rather than relying solely on intuition. Insight into future demand is one aspect where substantial improvements can be made with the use of forecasting. Currently, Company X is using only qualitative forecasting based on managers´ and therapists´ experience, if any form of forecasting is used at all. Therefore, we constructed a quantitative probabilistic forecasting model to forecast a therapist´s weekly workload for its caseload of patients. The model is based on the probability mass function of the Poisson binomial distribution. This is applicable to the process at Company X, if we assume that a patient either has one appointment per week or none, with a certain probability. This is comparable to the success probabilities of the Bernoulli trials, which is a random experiment with exactly two outcomes (Papoulis, 1984). The probabilities are not identically distributed over all patients and weeks, unlike the classical Bernoulli trials of a binomial distribution. The Poisson Binomial distribution is therefore a good fit.

Equation 1 gives the formula for the exact calculation of the probability of having k appointments per week, based on the probability mass function of the Poisson binomial distribution (Wang, 2023).

$$\Pr_w(K = k) = \sum_{A \in F_k} \prod_{i \in A} p_{i,w} \cdot (1 - r_c) \cdot (1 - r_n) \prod_{j \in A^c} (1 - p_{j,w})$$

Equation 1: Probability function that a therapist has k appointments in week w.

The therapist's caseload is given as a set of n patients $N = \{1, \dots, n\}$. In Equation 1, F_k is the set of all subsets of N for which k patients have an appointment in week n. Every patient i can either have an appointment in week w with probability $p_{i,w}$ or zero appointments with probability $(1 - p_{i,w})$. The week w is included as an index in the equation because the probabilities differ per week, depending on which care pathway the patient follows and how far ahead in the pathway the patient is. The sets A are all individual subsets of F_k and A^c is the complement of A. For example, if you want to find the probability that there is one appointment is week 1, and there are 5 patients in total, then if set A is $\{1\}$, set A^c is $\{2,3,4,5\}$. The no-show rate and the cancellation rate of appointments are also included the model by the cancellation rate r_c and the no-show rate r_n, these are the probabilities that an appointment is not cancelled and that there is not a no-show.

There are $\frac{n!}{k!}$ elements in F_k, which leads to an infeasibly high run time for finding the subsets in F_k, except if the number of patients n is small. For a relatively common caseload of 30 patients, there would already be 10^{20} subsets to consider. Despite the effective procedure for finding the subsets of a set in Section 3.2.2, the computation time remains substantial. Especially since this is only the running time for finding the subsets and not for running the code of the complete model. There is a Python extension available that enumerates all subsets of a set in a relatively short amount of time for the amount of computations, but that falls outside the scope of this research.

The number of forecasted weeks is adjustable, however the model is intended to forecast workload for a timeframe of a quartile or half year, as fits the situations at Company X best. The three levels of management decisions are; strategic, tactical, and operational. This forecasting model can be categorized as an operational model, because the operational level ranges from daily decisions until half a year from now (Schmidt & Wilhelm, 2000). The forecasting model can be used to make weekly decisions about overloading or underutilised workload. The model is also useful for quarterly decisions about intake planning, which are also operational decisions.

3.2.1 Model assumptions

To model a therapist's weekly workload several assumptions and simplifications are required.

- All patients have either zero appointments or one appointment per week. This is a small simplification of the actual situation. For patients of psychomotor therapists, this is an accurate simplification, but that may not be the case for all types of patients.
- If an appointment is cancelled by the therapist or patient, or if there is a no-show, the appointment is cancelled and not planned again. This is a simplification of the actual situation.
- The model does not take new patients that arrive during the forecasting horizon into account. This simplification is made such that the tool can be used for the intake planning.
- The probability that a patient has an appointment depends only on the care pathway that the patient is following and how far advanced the patient is in the care pathway. This is an assumption
- The no-show rate and the cancellation rate are the average rates based on Company X's data for Location Y in Section 2.5.6.2 and are thus assumed to be the same for all patients and care teams. This is an assumption of which the employees of Company X estimated that the effects will indeed be neglectable.
- The no-show rate is also included for appointments in indirect time since no distinction is made between appointments in direct and indirect time. This is a simplification that is made because the indirect and direct time data is often tracked in a different way.
- Seasonality is not included in the model. This is a simplification that does affect the accuracy of the forecasting results, since there are often less appointments during holiday weeks.
- This assumption is specific for psychomotor therapists. A patient in a follow-up block always has the same probability of continuing treatment with another follow-up block, regardless of the number of completed follow-up blocks. In reality, if the patient has already completed multiple follow-up blocks, the probability of terminating the treatment increases. This is thus a simplification.

3.2.2 Enumerating all subsets of a set of patients

In order to find all probabilities $\Pr(K = k)$ that the therapist's workload consists of k appointments in week w, we first need all subsets of patients that have an appointment if set A is the set of all patients in treatment with the therapists in week w.

The time it takes to find all subsets of a set increases exponentially as the number of elements in the set increases. Commonly used methods are Gray codes and lexicographic ordering, but for large sets these are sub-optimal. Loughry et al. (2000) developed an algorithm that systematically generates a sequence of subsets of a set by progressively increasing the number of elements. The article also includes an algorithm programmed in C++, this programming language was adjusted and converted to VBA to be used for the calculations in Section 3.3. The original code can be found in the article by Loughry et al. (2000).

3.3 Applying the model

The model described in Section 3.2 will in practice not be applied at Company X, due to its impractically high run time of finding all subsets of a set consisting of a complete caseload of patients. However, the model is able to exactly calculate the expected workload for small samples. So, it could be a valuable model in another context, e.g. for other psychiatric organisation where therapists have smaller caseloads. Therefore, this section will apply the model to a small sample to show its application.

The model will be applied to a psychomotor therapist with a caseload of 5 patients, forecasting the therapist's workload for 12 weeks. The input values for calculations for a psychomotor therapist are given in Section 2.5.6. The characteristics of the four care pathways will be used, just as the rates for no-shows and cancellations from Section 2.5.6.2. In addition to that, transition probabilities of the transitions between care pathways are required, this data is given in Table 5. 'Start PMT (end)' and 'Follow-up PMT (end)' are termination pathways, so this is always the last pathway before a patient completes its treatment. All its transition probabilities from the termination pathways are thus always zero. The transition probabilities are not based on historic data because of data availability issues, but it is a rough estimation. We advise Company X to determine the probabilities by analysing historic data and inquiring therapists.

	Start PMT	**Start PMT (end)**	**Follow-up PMT**	**Follow-up PMT (end)**
Start PMT	0	0	0,3	0,7
Start PMT (end)	0	0	0	0
Follow-up PMT	0	0	0,1	0,9
Follow-up PMT (end)	0	0	0	0

Table 5: Transition probabilities between care pathways for a patient treated by a psychomotor therapist. The probabilities are given as: the probability that the pathway given in the column of a cell is next when the patient is currently in pathway given in the row of the cell.

The five patients are randomly assigned to a care pathway and a position in the care pathway as if the therapist would have given the input. The randomly generated start situation has one patient starting in each of the pathways, expect 'Start PMT (end)', where 2 patients start. The pathways indicated with 'end' are termination pathways, as these are the final pathway before a patient completed the treatment. Therefore three of the five patients are finishing their treatment after the current pathway and two patients are assigned to one of the other pathways with given transition probabilities.

Now we will apply the exact model to a random sample of a psychomotor therapist with a caseload of five patients, forecasting the therapists' workload for 12 weeks. All patients i have a probability $p_{i,w}$ that they have an appointment in week w. All these probabilities for all i and w are manually calculated in Excel, using the input data. The subsets of the set of five patients are determined according to the method in Section 3.3. The subsets and probabilities $p_{i,w}$ were then used to calculate the probabilities $\Pr(K = k)$ of a therapist having a workload of k appointments per week for $w \in \{1, \dots, 12\}$ and $k \in \{0, \dots, 5\}$. The calculations for the therapists' caseload of 5 patients and 12 weeks were programmed in VBA according to Equation 1. The code is explained in pseudocode in Figure 17. Only the calculation for this small sample was coded in VBA, we did not program a generally applicable method for the exact forecasting model.

```
For all number of appointments per week
    For all weeks
        For all subsets of patients
            Determine the probability the number of patients in the subset has an appointment and
            the patients outside the subset do not
```

Figure 17: Pseudocode of finding a therapist's weekly workload using the exact workload forecasting model.

The results of the model is displayed in Table 6, with rounded probabilities of having k appointments in week w.

	Number of weeks											
Number of weekly appointments	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
0	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,005	0,005	0,005
1	0,000	0,000	0,000	0,000	0,001	0,139	0,139	0,139	0,139	0,986	0,986	0,986
2	0,006	0,006	0,006	0,006	0,053	1,723	1,723	1,723	1,723	0,017	0,017	0,017
3	0,099	0,103	0,103	0,103	0,579	0,000	0,000	0,000	0,000	0,000	0,000	0,000
4	0,267	0,267	0,267	0,267	0,742	0,000	0,000	0,000	0,000	0,000	0,000	0,000
5	0,688	0,688	0,688	0,688	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

Table 6: The probabilities (rounded to 3 decimal places) that a number of appointments take place, which is the outcome of the exact workload forecasting model.

Now that we have shows the forecasting results of the exact model, we will illustrate the calculation method more clearly. First we calculate all weekly probabilities for all patients in the therapist's caseload. We show an example calculation of a weekly (w) probability of a patient (i) in Equation 2. Equation 2 shows the calculations of the probability that a patient has an appointment in week 4, who was in position 10 of the 'Start PMT' pathway in week 1. After this patient finished their current pathway, the patient has a 0,3 probability of going to the 'Follow-up PMT' and 0,7 probability of going to the 'Follow-up PMT (end)' pathway.

$$\begin{aligned} &\mathrm{P}(\text{"}patient\ i\ has\ an\ appointment\ in\ week\ 4\text{"}) = p_{i,w} \cdot (1 - r_c) \cdot (1 - r_n) \\ &= p_{1,4} \cdot (1 - 0{,}059) \cdot (1 - 0{,}014) \\ &= 0{,}3 \cdot 0{,}928 + 0{,}7 \cdot 0{,}928 \cdot (1 - 0{,}059) \cdot (1 - 0{,}014) = 0{,}881 \end{aligned}$$

Equation 2: Calculation of the probability that a patient, who is in position 10 of the 'Start PMT' pathway in week 1, has an appointment in week 4. The numbers are rounded to three decimals.

Then we calculate the probabilities that there are k appointments, with $k \in (0, 1, \ldots, n)$. Using the probabilities for all weeks (w) and all patients (i), we can calculate the probabilities that there are k weekly appointments. This is calculated in Excel using Equation 1 and the VBA code in Figure 17. Equation 3 shows to calculate the probability that there are 0 appointments in week 1. The other calculations of k are similar. If the values of $P(K = k)$ are found for all k, the exact model calculations are completed.

$$\begin{aligned} P(0\ appointments\ in\ week\ 1) = &\sum_{A \in F_k} \prod_{i \in A} p_{i,1} \cdot (1 - r_c) \cdot (1 - r_n) \prod_{j \in A^c} (1 - p_{j,1}) \\ &= (1 - p_{1,1}) * (1 - p_{2,1}) * (1 - p_{3,1}) * (1 - p_{4,1}) * (1 - p_{5,1}) \end{aligned}$$

Equation 3: Formula for finding the probability that there are zero appointments in week 1, if the therapist has a caseload of five patients.

3.4 Conclusion

This chapter researched the question: *"How to create a forecasting model to predict a therapist's workload?"*.

The literature study highlights the challenges of uncertainty and complexity of demand and the importance of data-driven decision-making in healthcare. It then discusses the different types of forecasting models, focussing on quantitative forecasting methods. Different forecasting methods are proposed, including causal and time series forecasting, and deterministic and probabilistic forecasting.

Then a model description is given of the quantitative probabilistic exact forecasting model, based on the probability mass function of the Poisson binomial distribution. The model calculations require all subsets of the set of all patients, which leads to an exponentially increasing run time as the number of patients in the therapists' caseload increases. The determination of all subsets of a set is also given in this chapter. The exact forecasting model is in practise only applicable to a therapist with a small caseload, especially since the improvement programme eventually aims to model workload for groups of therapists. The chapter concludes with the application of the previously described model to a small sample to show its theoretical application.

Based on this chapter we have recommendations to Company X, starting with finetuning the model's paraments based on historical data and/or inquiring therapists, especially the transition probabilities. The other variables are already based on historic data, but the transition probabilities are a rough estimation.

4 Simulation-based Forecasting Approximation model

Since the exact workload forecasting model as constructed in Chapter 3 is infeasible for practical application due to its high run time, we will construct a Monte Carlo simulation model in this chapter. This answers the third research question:

> *"How to construct a forecasting simulation model to predict a therapist´s weekly workload?"*

Section 4.1 is a literature study of Monte Simulation. Section 4.2 describes the simulation-based forecasting approximation model. Section 4.3 shows the VBA implementation of the model and Section 4.4 evaluates the impact of demand variability on the workload based on the model. Section 4.5 lastly recommends how the simulation model can be improved.

4.1 Literature study of Monte Carlo simulation modelling

Monte Carlo simulation is defined as a type of discrete-event simulation where stochastic models are used to randomly generate configurations of a system (Earl & Deem, 2008). Large amounts of random data are generated in order to accurately describe a system and gain insights into the probability of an event (Sousa et al., 2020). Monte Carlo simulation is a commonly used technique for the probabilistic analysis of a system. Statistics of the input variables of a system are given as input and the statistics of output variables are the output of the simulation. A Monte Carlo simulation consists of a high number of simulation runs. The simulation results will become more accurate as the number of runs approaches infinity. During each iteration of the simulation, random values are assigned to the variables, based on the given statistics of the input variables by the user. The high number of runs creates an accurate image of the statistical distribution of the output variables (Mahadevan, 1997). Another advantage of Monte Carlo simulation is that is applicable for complex systems but is still relatively simple to implement (Papadopoulos & Yeung, 2001).

Monte Carlo simulation is based on the principle of the 'law of large numbers', which describes the convergence of a distribution of large samples to the distribution of the underlying population (Gilli et al., 2019). Consequently, a large number of runs is needed in order to construct a reliable and accurate Monte Carlo simulation model. The law of large numbers is formally stated in Equation 4.

$$\lim_{n \to \infty} \sum_{i=1}^{n} \frac{X_i}{n} = \bar{X}$$

Equation 4: Law of large numbers.

According to Horcas Aguilera et al., there are three typical uses of Monte Carlo simulation (2021). First, it involves sampling to obtain information about a random object through the observation of many realisations, such as simulating the behaviour of a system. Second, it allows for the estimation of specific numerical quantities, for example, the expected throughput in a Software Product Line. Third, Monte Carlo can be used for optimizing a complex objective function, for instance for a search-based technique. The aim of this research is in line with the second use of Monte Carlo simulation since the numerical quantity workload is estimated, consisting of a number of appointments and a number of working hours.

4.2 Model description

In this chapter, we will construct a forecasting simulation model that predicts a therapist's weekly workload based on Monte Carlo simulation. The decision to apply Monte Carlo simulation was two-fold. First, from the researcher's perspective, the model is applicable to complex situations, including the context of this research, where complexity is caused by the high level of variability in treatment between patients and therapists. The outcome of the model can be presented either discretely or as a probabilistic distribution of all outcomes. Second, from Company X's perspective, the decision was made to implement a relatively straightforward model as Monte Carlo, such that the model could be adjusted at a later stage if necessary. This will allow for a better understanding of the model by Company X's employees.

The simulation model is based on the same input values as the exact model in Chapter 3. All assumptions described for the exact forecasting model in 3.2.1 also hold for the simulation forecasting model. The simulation model is implemented by developing a dashboard in Excel and programming the Monte Carlo simulation in VBA. The scope of the simulation model is the ability to forecast a psychomotor therapist´s workload for its current caseload. The demand is forecasted for a psychomotor therapist because the appointment series are clearly characterized for this type of therapist (in Section 0), enabling us to forecast the workload based on Company X´s data.

Figure 18 shows the dashboard of the simulation forecasting model in the first Excel sheet called "Dashboard". The input data on the left side of the dashboard is divided into therapist-specific data and general data. The therapist-specific input includes the care team, the type of therapist, and the number of patients in the therapist's caseload. The description of the caseload can be entered in the worksheet "Input Caseload". This input consists of the care pathway and position in the care pathway during week 1 of the forecast. The general input data consists of probabilities of no-shows and cancellations (either in time or too late), and both originate from the probabilities in Section 2.5.6.2. The complete monthly data of no-shows, cancellations etcetera used to determine these probabilities are given in the worksheet "Cancellations and No-Shows". The percentages on the dashboard automatically change if the data in this worksheet about cancellations is updated. The number of simulation runs and the number of forecasted weeks are also part of the general input.

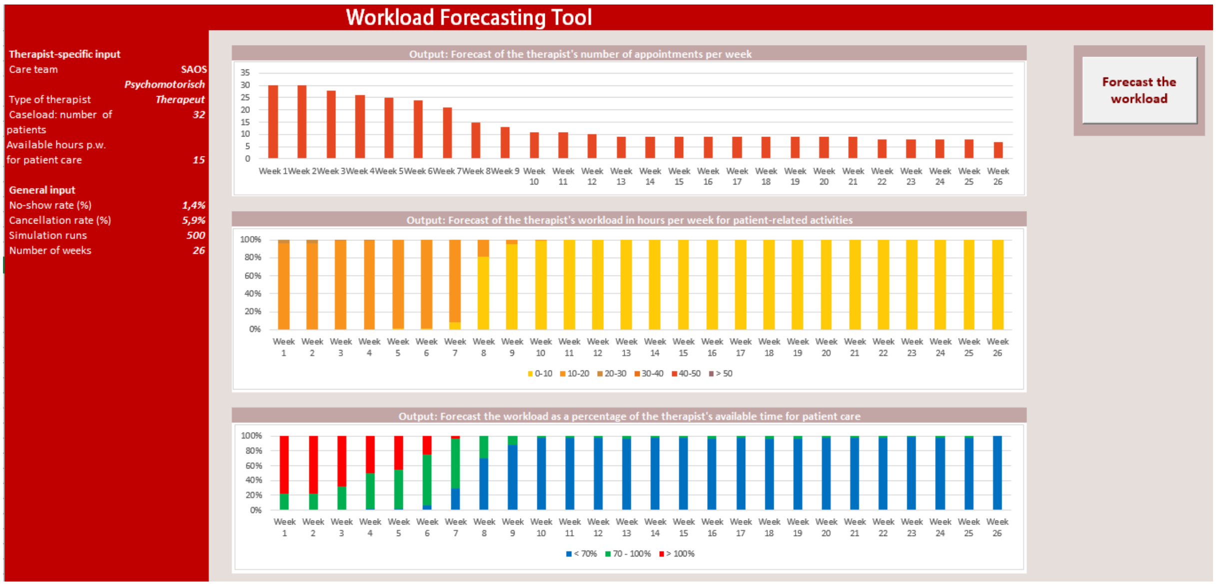

Figure 18: the layout of the dashboard of the workload forecasting simulation model created for Company X in Excel VBA, here for the input values of a psychomotor therapist for 32 patients and 26 weeks.

Additional input values have to be specified per type of therapist in the worksheet "Input data". These input values are the care pathway characteristics and weekly probabilities per care pathway, as was done in Table 3. Also, the probabilities that the appointment duration is in a certain range have to be specified on this worksheet. The ranges are given and are half an hour long each. The probabilities are derived from the historical data of all appointments at company X in Location Y between 2019 and 2022. For psychomotor therapists these probabilities are provided in Table 4 in Section 2.5.6.3. The transition probabilities are given in Table 5 and are used in the same method as for the exact model in Chapter 3.

As explained in Section 4.1, the input variables are randomly assigned during each run based on the input distribution. The starting care pathways and positions are given by the user of the simulation model. Afterwards, the remaining sequence of the pathways is followed. If the current pathway is not a termination pathway, another pathway is randomly assigned using the transition probabilities. A large number of possible scenarios is hereby simulated, and either a probability distribution of the possible outcomes or the most probably outcome is the result of the simulation model. The principle of Monte Carlo simulation is not only applied for randomly generating sequences of pathways and appointments per patient, but also the appointment duration is simulated using Monte Carlo based on the statistical distribution that followed from Table 4.

The Monte Carlo simulation results are presented as output on the dashboard in three graphs. The first output graph shows the most probable forecasted number of appointments per week. The second graph visualises the distribution of the number of hours per week the workload consists of. The third graph presents a distribution of the number of hours worked in comparison with the available number of hours for patient-related tasks. The output is given in the worksheet "Output simulation" and is then used to create the tables on the dashboard.

The user of the simulation model can use the results of the simulation for example for seeing when there is enough room in the therapist's weekly schedule to plan new intakes. Additionally, the model is useful for examining which therapists have an overloaded weekly workload.

4.3 Implementation of the forecasting simulation model in VBA

4.3.1 Programming the forecasting simulation model

The workload forecasting simulation model is developed by programming the Monte Carlo simulation in VBA Excel according to the pseudocode in Figure 19. The complete VBA code is given in Appendix A.

```
For all runs
    For all patients
        Randomly appoint an appointment series
        Randomly appoint a location in the appointment series
        For all weeks
            Assign whether the patient has an appointment or not from the distribution
For all weeks
    For all runs
        Calculate the average weekly appointments of all runs
```

Figure 19: Pseudocode for forecasting a therapist's workload.

The number of runs is an input variable of the workload forecasting simulation model. However, this is by default set at 10.000 runs in order to make an accurate forecast in accordance with the law of large numbers as explained in Section 4.1. The simulation model has a short runtime considering the number of runs it performs, as one million simulation runs usually take less than one minute.

4.3.2 Verifying the performance of the simulation model

The expectation was that the number of appointments would decrease as the weeks progresses. In addition to that, the number of hours that the weekly workload consists of should also decrease as the weeks pass by. After running the simulation model this indeed proved to be the case. When the simulation was run for a psychomotor therapist in Location Y with a caseload of 33 patients for 26 weeks, it can be concluded that the number of appointments was halved in week 9 and almost zero in week 14. This provides useful input for the decision-making for the intake planning.

The evaluation of the simulation model's accuracy is performed using the data of the exact forecasting model about the sample of 5 patients over 12 weeks in Section 3.3. The same input data as for the model is also entered into the simulation model. The output of the model and the simulation model is compared. For the simulation model, the most probable number of appointments is given, so the results are integers. The forecasting model is probabilistic and thus gives its results in decimal numbers. Table 7 gives a comparison of the results of the model and the simulation model per week, which is also visually displayed in the graph in Figure 20. Given that the same input data is used for both the simulation model and model, and that many of the same assumptions are made, the output should be similar, regardless of the difference in forecasting methods.

Weeks	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Simulation model	5	5	5	5	5	5	4	2	2	2	2	2
Exact model	4,6	4,6	4,6	4,6	4,6	4,6	3,7	1,9	1,9	1,9	1,9	1,8

Table 7: Comparison of the results of the exact model and simulation model, for the workload forecast of 5 patients over 12 weeks. The model's values are rounded on one decimal place.

The results of the simulation model and the forecasting model can be compared using Table 7 or Figure 20. It can be concluded that the results of the simulation model and model are equal if the model's results are rounded up to integers. Given that the same input values are used, but with a different forecasting method, this can be interpreted as a confirmation of the correctness of the simulation model's results. This is because the results of the model are exactly calculated so we can assume these are true. However, this is of course only evaluated based on one calculation, so for a more extensive comparison, it would be useful to compare the results for more randomly picked samples. Additionally, the results may be limited by incorrect assumptions or the accuracy of the input data. However, if these assumptions approximate the real world and the input data is accurate, then the simulation model's forecasts can be considered accurate based on this first example at least.

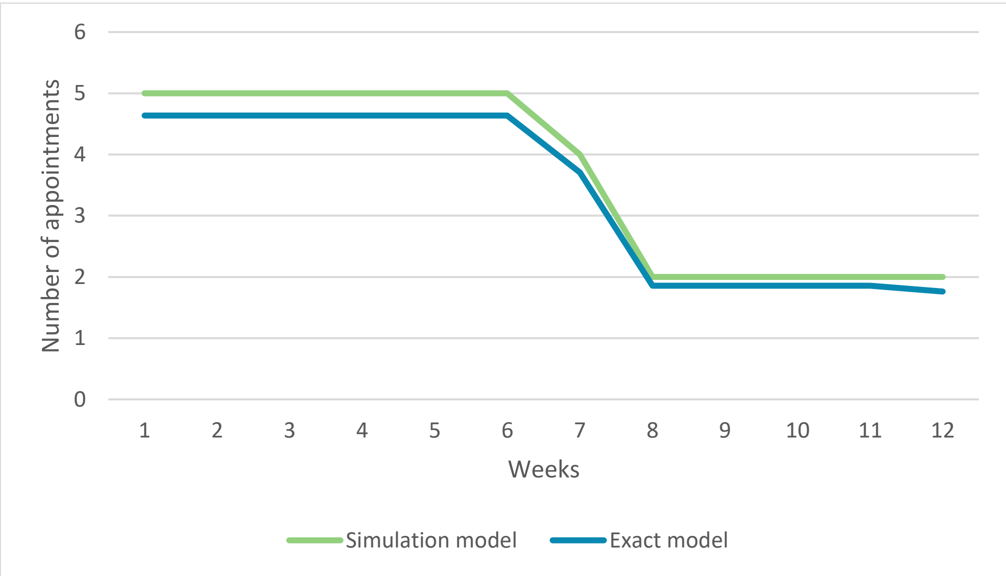

Figure 20: Comparison of the results of the exact model and simulation model, for the workload forecast of 5 patients over 12 weeks. Based on the example in Section 3.3.

Verifying the model by only testing with one sample is not strong method. Therefore, we include another sample to test the model on. Again, we forecast for a caseload of 5 patients and 12 weeks, however we take different starting conditions for the patients. Table 8 shows the starting conditions of the therapist's caseload of 5 patients, consisting of the patient's care pathways and positions in week 1.

Patient	Start care pathway	Start position
1	1	3
2	3	4
3	4	2
4	2	5
5	4	7

Table 8: Starting conditions of the caseload in the second sample on which we tested the forecasting simulation model.

We forecasted the number of weekly appointments using the exact forecasting model and the simulation forecasting model. Figure 21 compares the results of the exact and simulation model for the second sample of 5 patients. We can conclude from Figure 21 that the rounded results from both forecasting models are equal for this sample, so the simulation model correctly forecasts the number of weekly appointments for this sample, assuming that the calculations from the exact forecasting model are correct.

Two samples still provides a limited proof that the simulation model works, therefore we advise Company X to evaluate the model's results by comparing with actual data. For this research, there was no data available yet to test this model on, but we do highly recommended company X to take this as a next step in improving the model. By finding the gaps in between the forecasted data and the actual data, the tool can be further improved.

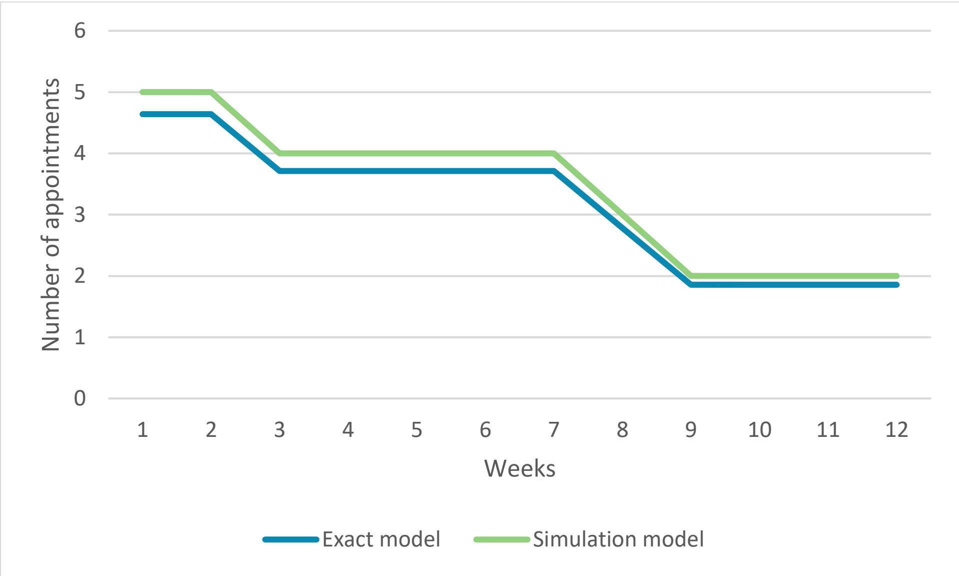

Figure 21: Comparison of the results of the exact model and simulation model, for the workload forecast of 5 patients over 12 weeks. Based on the second sample to further verify the simulation model.

4.4 Impact of demand variability on workload

There are different aspects of demand variability, as we explained in Chapter 0. These all have a different impact on the results of the forecasting simulation model. The different aspects are the variability of appointment duration, appointment frequency and use of care pathways.

Appointment duration is measured in the second graph on the dashboard by a measure of variability since the result is given in total duration ranges with its probability. If the appointment duration for appointments would always be the same or fall in a small range, the model would forecast the weekly workload with almost certainty. However, when there is a large spread of durations because of high appointment duration variability, then there is a high level of uncertainty in the forecast. The number of patients can also influence the uncertainty in the appointment duration results of the simulation model.

Two examples of outcomes of the forecasting simulation model are given in Figure 22 and Figure 23. Here a clear difference in the certainty of the interpretation of results is visible, as the level of certainty when interpreting the results is clearly higher in Figure 22 as opposed to Figure 23. These figures are both for the appointment duration range of a psychomotor therapist, but if the output would be compared for different types of therapists with varying appointment duration distributions, there would also be a difference regarding to what extent the variability impacts the workload.

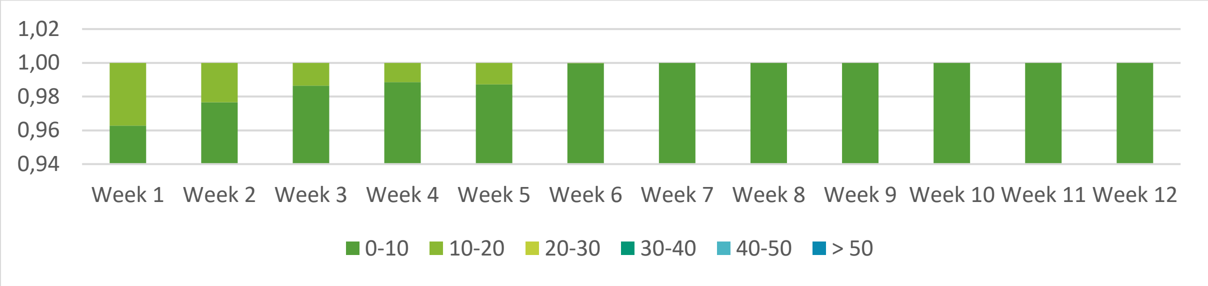

Figure 22: Forecasted of the therapist's workload in ranges of hours and probabilities, for 12 weeks and 14 patients. This figure is output of the forecasting simulation model.

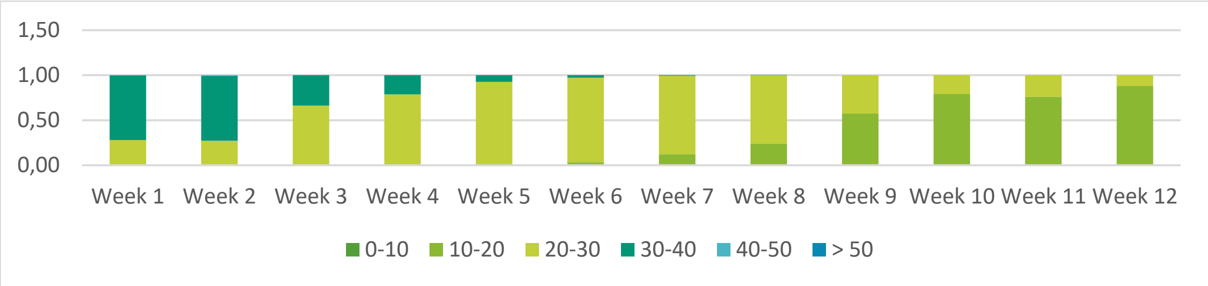

Figure 23: Forecasted of the therapist's workload in ranges of hours and probabilities, for 12 weeks and 63 patients. This figure is output of the forecasting simulation model.

Secondly, variability in appointment frequency affects the workload. In principle, the patients would follow a series of appointments with psychomotor therapists on a weekly basis. However, because of the impact of cancellations and no-shows, this weekly frequency is rarely completely reached. This affects the forecasted workload, as can be seen in Figure 24. The therapist starts with a caseload of 33 patients, which is then slowly decreased as more patients finish their treatment. However, even in the first weeks when the complete caseload is still in treatment with the therapist, there are never 33 appointments per week, because of cancellations and no-shows.

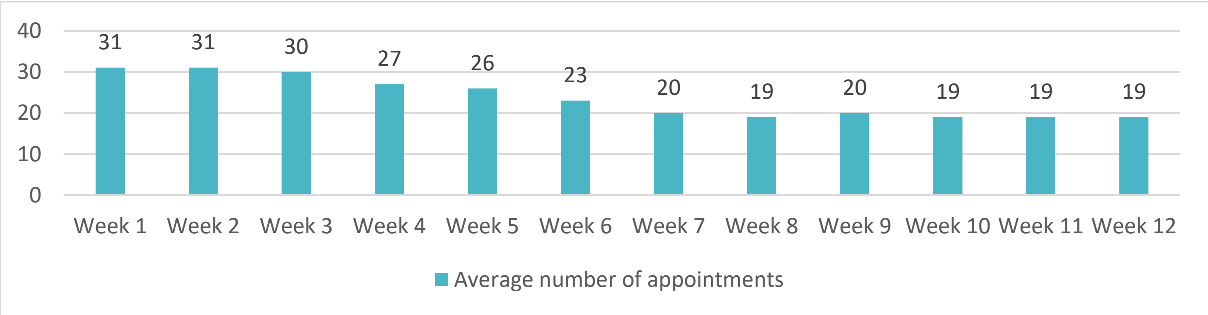

Figure 24: Forecast of the number of weekly appointments for 12 weeks and a caseload of 33 patients. This is output of the forecasting simulation model.

Lastly, the variability of care pathways plays a role. There are different care pathways with differing lengths and characteristics, some are termination pathways, and the lengths range from 7 to 12 weeks. If the patients all would follow the same pathway or if the number of pathways would be doubled, the forecasting results from the simulation model would be significantly different. This becomes evident when the current input values from the simulation model are compared with a case where all patients first do a Start PMT block, continue with a follow-up block, and then complete their treatment. This would result in a steeper decrease in the forecasted number of weekly appointments as can be seen in Figure 25 of the current situation and Figure 26 of the more standardized case.

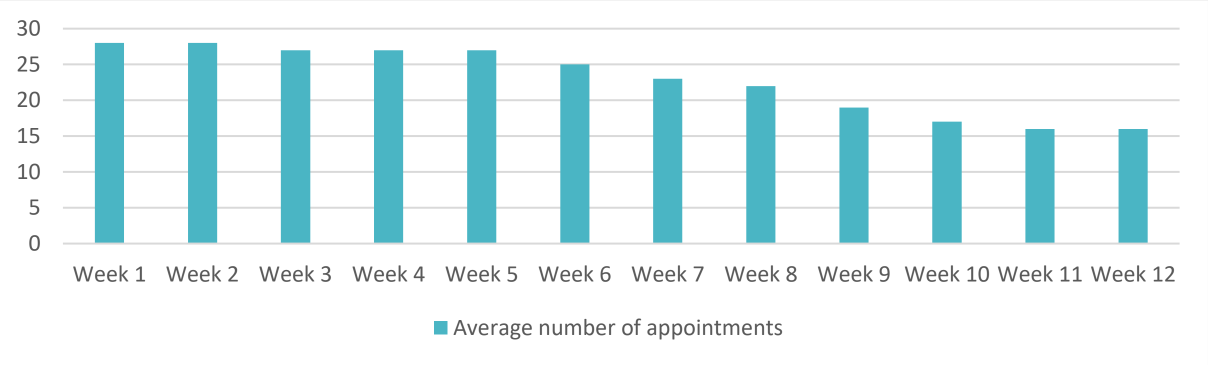

Figure 25: Forecasted number of weekly appointments for 30 patients and 12 weeks, using the standard input data of the forecasting simulation model.

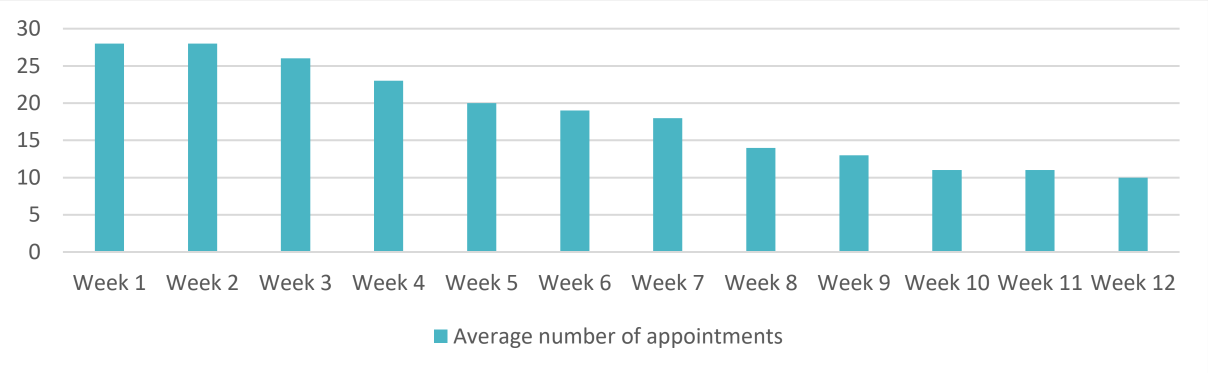

Figure 26: Forecasted number of weekly appointments for 30 patients and 12 weeks, using a more standardized treatment path where patients always do the start PMT block and the follow-up block.

4.5 Improving the forecasting simulation model

First, we will explain the improvements that were already made to the simulation model in discussion with Company X. Afterwards, we present our recommendations to Company X on how to further improve the simulation model.

Initially, we planned to create a simulation model that forecasts the demand of the next quartile, after revision and discussion with Company X we decided that it is more useful to let the number of weeks be adjustable by the user. It remains an operational model with a medium-time frame between a few months a year, but it does grant more freedom to the user. Additionally, when presenting an earlier version of the forecasting simulation model to an employee of the improvement programme, it was indicated to be more useful if a description of the caseload is given as user input, instead of generating this randomly. Therefore, a patient’s current care pathway and position in the pathway are used as input values of the simulation model.

We will list the key improvement areas that we strongly recommend Company X to prioritise to enhance the simulation model's performance. The obvious recommendation is to implement the model into the organisation by expanding to all types of therapists. This includes determining the input data for all types of therapists, including the characteristics of the care pathways, transition probabilities between care pathways and the probability distribution of the appointment durations.

Furthermore, the possibility for a patient to have multiple appointments per week can be implemented, either by adjusting the series characteristics or by changing the appointment duration distribution into the total duration of appointments per week. Both are approximately equally accurate, but the second option is probably the easiest to implement in the current model. Patients in treatment with psychomotor therapists in most cases have a maximum of one appointment per week, but this varies widely among the different types of therapists.

The model can be expanded to be functional for all types of therapists, by using the same procedure as was described to create the model for a psychomotor therapist. Eventually, the improvement programme team is working towards creating a model to forecast the workload of groups of therapists, for example of the same type of therapist or in the same care team. A starting point for that is given in this research by following the principles of the forecasting model created in this research.

4.6 Conclusion

This chapter answered the final research question: *"How to construct a forecasting simulation model to predict a therapist´s weekly workload?"*.

We created a simulation model using Monte Carlo simulation. We developed the dashboard in Excel by programming a Monte Carlo simulation in VBA. The model forecasts a therapist's weekly workload based on various input variables, such as a description of the caseload and the transition probabilities and care pathways. Many scenarios are simulated to determine a distribution of the possible outcomes with their probability. The number of weekly appointments and workload in hours are forecasted. Certain assumptions and limitations are considered in the design of the simulation model, such as having a maximum of one appointment per week, fixed appointment frequencies and using an average cancellation rate.

The performance of the simulation model is evaluated by comparing the results of the exact calculation based on the forecasting model for a small sample in Section 0 and the results of the forecasting model for the same input data. Since the results exactly match, we assume the validity of the results simulation model. The validation is limited by the fact that only one comparison was made for a sample of 12 weeks and 5 patients. Additionally, this also assumed that the input values and the simulation model´s assumptions are correct. Ideally, we would have compared the model's output to historical data, but because of data availability this was not yet possible in this research.

The simulation model can be used by the improvement programme team of Company X to create more insight into the expected weekly demand for example for the next quartile. The simulation model can be extended to all types of therapists, by following the method as in this Chapter.

5 Conclusion

This research was undertaken to evaluate the impact of the demand variability in youth mental healthcare on the required therapy capacity. The core problem of this research is the high level of demand variability in the treatment processes of psychiatric patients, which causes fluctuating workloads, and long waiting lists. This is a problem that mental healthcare organisations all over the Netherlands are currently facing. That makes this research not only relevant to Company X, but also to other psychiatric organisation in the Netherlands. Furthermore, the research makes a practical contribution by creating a method to create a simulation model for Company X to implement in order to foresee when there is room in the workload for new intakes and when solutions are needed for therapists with an excessive weekly workload. Additionally, the research makes a scientific contribution since we are the first to apply Monte Carlo simulation to forecast therapists' workload in a psychiatric context. This simulation model can thus help psychiatric patients and therapists throughout the Netherlands with shorter access times and stabilized workloads.

We first created the exact forecasting model as an early attempt to forecast workload using an exact probabilistic model. This model is based on the probability mass function of a Poisson binomial distribution and can precisely calculate the probabilistic distribution of how many appointments will occur weekly. A disadvantage of the model is the high run time due to the fact that the model requires all subsets of the number of patients in the caseload as its input. Even if the determination of subsets is effectively calculated in a coded algorithm, the time-consuming runtime of the model restricts the model from a practical implication in this context. However, the model might be useful in another organisation where therapists always have a small caseload of patients.

Due to the impractical runtime of the exact forecasting model, we developed a Monte Carlo simulation model. Random scenarios are generated for a patient's treatment during the forecasted timeframe, as well as the duration of appointments. This is generated for a high number of simulation runs, since the results approach the exact solution as the number of runs reaches infinity. We developed a simulation model that forecasts the weekly workload of a psychomotor therapist, to illustrate to Company X how the simulation model can be expanded for other types of therapists.

The results of the simulation model can be used to make tactical and operational data-driven decisions, about workload division, intakes and waiting lists. The bottlenecks regarding workload can be solved in advance with the gained insight about which weeks have an excessive workload. Weeks with underutilized workload are also observable from the simulation model's results which is valuable quantitative input to be used for the planning of intakes. Improving these aspects of decision-making can lead to shortened waiting lists and a stabilized workload.

The performance of the forecasting simulation model was evaluated by comparing calculations made by an exact forecasting model. The approximated forecasting results of the simulation model and the calculations based on the exact model appeared to be equal for the sample on which it was tested. So, assuming the input values are correct, and the assumptions hold, we conclude that the results of the forecasting simulation model are valid. As soon as the company has testing data available, the forecasting results should also be compared with historic data, to improve the model's accuracy.

In addition to the implementation of the simulation model, we have further recommendations to Company X, starting with a more uniform way of working amongst different therapists of the same type. Creating a set of booking rules for therapist groups can be a good step in the right direction, in addition to encouraging therapists to follow the norms for appointment duration and frequency as long as there is no healthcare-related reason to deviate. This more uniform way of working also includes using standardized care pathways if possible, since these are created and validated by psychiatric scientific research and can decrease inter-therapist variability. By explaining the positive effects on the stability of the workload and the lengths of the waiting lists, the therapists can be convinced of the importance of a uniform way of working and its increased predictability of patient demand. Moreover, we advise improving the quality of data by setting strict data tracking standards for planners and therapists. Improving the data quality will enhance Company X's data-driven decision-making abilities.

6 Discussion

The research presented in this thesis is subject to several limitations, most of which are centred around the limited timeframe of this thesis. Starting with the missed opportunity to further validate the forecasting simulation model based on more comparisons between the exact forecasting model and the forecasting simulation model. Additionally, it would be beneficial to compare forecasted results with actual data, instead of only comparing it with the exact theoretical model. It was attempted to validate the simulation model with actual data, but this was constrained by data availability and data quality, because it was not possible to derive the data from the database in the right format to compare it with the forecasting results. There are probably options to do so, if Company X invests more time here.

The limited data availability also limited the possibility to develop a fully functional forecasting simulation model for all types of therapists, as the required booking rules were only available for psychomotor therapists. While appointment duration and frequency data could be obtained for all therapists, the absence of structured care pathways as provided in the booking rules prevented the creation of the simulation model for all types of therapists.

There is a high level of variability between the ways of working of different therapists at Company X, which is amplified by the frequency variability, duration variability and variability caused by cancellations and no-shows. The theoretical norms and guidelines are therefore not close to the actual therapy processes at Company X. By implementing measures for variability regarding appointment duration, cancellations, and no-shows an attempt is made at including the variability in the forecasting model. However, the forecasted results will still differ from historic data, which is why it would be a good point for future research to improve the forecasting simulation model based on testing with historic data.

The assumptions that had to be made in order to complete this thesis within the available time, did limit the accuracy of the forecasting results. The forecasting simulation model is limited by its assumption to have a maximum of one weekly appointment, which was a suitable assumption for psychomotor therapists, but not for all therapist types. This can be done either by adjusting the series characteristics and the coded procedure or by changing the appointment duration distribution. While the second option is easier to implement, the first option is more accurate. Another assumption that decreases the accuracy of the forecasted results of the simulation model is the fact that appointments that are cancelled because of a cancellation or no-show by a patient do not take place at a later stage. This could be improved by adding this as an extra rescheduling feature in the code, but because of time limitations, this was not possible in this research. Distinguishing between appointments in direct and indirect time can further improve the simulation model. The no-show and cancellation rate can then be applied only to appointments in direct time and separate more accurate appointment duration distributions can be used for indirect and direct time. The forecasting simulation model could be improved by including external factors such as seasonality. This requires adding the date of the starting week of the forecast and an analysis of seasonality throughout the year, such as the decrease in appointments during the summer holidays. By integrating seasonality into the model, the forecasting accuracy would improve. Especially because of the challenges of planning during the summer holidays.

Appendix

A. VBA code for the Monte Carlo Simulation of forecasting a therapist´s workload

```
Sub Simulation()

'Declare variables

Dim NrPatients As Long                  'number of patients in the therapist's caseload

Dim NrWeeks As Long                     'number of forecasted weeks

Dim MaxLengthSeries As Long             'maximum length of the longest appointment series

Dim NrSeries As Long                    'number of different standardized appointment series

Dim CancelRate As Double                'probability that a patient cancels an appointment

Dim NoShowRate As Double                'probability the patients does not show up at an appointment

Dim NrRuns As Long                      'number of runs for the Monte Carlo simulation

Dim CurrentSeries As Long               'which type of appointment series is the patient currently in

Dim CurrentPositionSeries As Long       'which position is the patient in when a (new) series is assigned

Dim AvWeeklyAppointments As Long        'average number of appointments per week

Dim RandomNumber As Double

Dim PastWeeks As Long

Dim Ranges As Long                      'ranges of single appointment duration of the probability distribution

Dim NrRanges As Long                    'ranges of weekly workload duration

Dim Total As Long                       'used for calculating averages

'declare arrays

Dim LengthSeries() As Long              'length of each series

Dim AppointmentsPerPatientPerWeek() As Long 'whether the patient has an appointment (1) or not (0) in a
week

Dim AppointmentSeries() As Double       'the standardized appointment series

Dim WeeklyAppointmentsRuns() As Long    'the number of appointments per week per run

Dim ForecastWeeklyAppointments() As Long 'the forecasted number of weekly appointments

Dim TransitionMatrix() As Double        'transition probabilities between appointment series, e.g.
TransitionMatrix(a,b) gives the prob of going from a to b

Dim CumulativeTransitionMatrix() As Double 'cumulative probabilties of transitions between appointment
series based on the TransitionMatrix()

Dim PatientSeries() As Variant          'the randomly generated appointments per patient for the NrWeeks

Dim TreatmentEnded() As Boolean         'whether a patients has completed its treatment (True) or (Not)

Dim Duration() As Variant               'the duration of a patient's appointment in week w

Dim DurationProb() As Variant           'cumulative probabilities and the start values of the ranges of 0.5 hours
each

Dim WeeklyDurations() As Double         'weekly total duration per run

Dim DurFreqArray() As Long              'duration ranges and the frequency the weekly workload is in this range
```

```
Dim DurProbArray() As Variant            'duration ranges and the probability the weekly workload is in this
range

'declare counter variables

Dim i As Long                    'general counter

Dim j As Long                    'general counter

Dim r As Long                    'simulation runs counters

Dim p As Long                    'patient counter

Dim w As Long                    'week counter

'declare worksheets

Dim DashboardSheet As Worksheet

Dim InputSheet As Worksheet

Dim SimulationSheet As Worksheet

Dim TransitionProbSheet As Worksheet

Dim CaseloadSheet As Worksheet

'set sheets

Set DashboardSheet = ThisWorkbook.Sheets("Dashboard")

Set InputSheet = ThisWorkbook.Sheets("Input data")

Set SimulationSheet = ThisWorkbook.Sheets("Output Simulation")

Set TransitionProbSheet = ThisWorkbook.Sheets("Input Transition Probabilities")

Set CaseloadSheet = ThisWorkbook.Sheets("Input Caseload")

'assign values to variables from the dashboard´s input values

NrWeeks = DashboardSheet.Cells(14, 2)

CancelRate = DashboardSheet.Cells(12, 2).Value

CancelRate = 1 - CancelRate

NoShowRate = DashboardSheet.Cells(11, 2).Value

NoShowRate = 1 - NoShowRate

NrRuns = DashboardSheet.Cells(13, 2).Value

NrPatients = DashboardSheet.Cells(6, 2).Value

NrRanges = 6

'set arrays

ReDim TreatmentEnded(1 To NrPatients)

ReDim AppointmentsPerPatientPerWeek(1 To NrPatients, 1 To NrWeeks)

ReDim WeeklyAppointmentsRuns(1 To NrRuns, 1 To NrWeeks)

ReDim ForecastWeeklyAppointments(1 To NrWeeks)

ReDim PatientSeries(1 To NrPatients, 1 To NrWeeks)

ReDim Duration(1 To NrPatients, 1 To NrWeeks)
```

```
ReDim AvWeeklyDuration(1 To NrWeeks)
ReDim WeeklyDurations(1 To NrWeeks)
ReDim DurProbArray(1 To NrRanges, 1 To NrWeeks)
ReDim DurFreqArray(1 To NrRanges, 1 To NrWeeks)
'find the amount of appointment duration ranges in the probability distribution
i = 1
Do Until InputSheet.Cells(22 + i, 1) = ""
   Ranges = i
   i = i + 1
Loop
'assign duration values to DurationProb() and probabilities
ReDim DurationProb(1 To Ranges, 1 To 2)
DurationProb(1, 1) = InputSheet.Cells(23, 1)
DurationProb(1, 2) = InputSheet.Cells(23, 2)
For i = 2 To Ranges
   DurationProb(i, 1) = DurationProb(i - 1, 1) + InputSheet.Cells(22 + i, 1)
   DurationProb(i, 2) = DurationProb(i, 2) + InputSheet.Cells(22 + i, 2)
Next i

'derive NrSeries from the worksheets
i = 1
Do Until InputSheet.Cells(3 + i, 5) = ""
   NrSeries = i
   i = i + 1
Loop
'assign values to TransitionMatrix() from the worksheet
ReDim TransitionMatrix(1 To NrSeries, 1 To NrSeries)
For i = 1 To NrSeries
   For j = 1 To NrSeries
      TransitionMatrix(i, j) = TransitionProbSheet.Cells(i + 3, 1 + j)
   Next j
Next i
'assign values to the cumulative transition matrix
ReDim CumulativeTransitionMatrix(1 To NrSeries, 1 To NrSeries)
For i = 1 To NrSeries
   For j = 1 To NrSeries
```

```
        If j = 1 Then
            CumulativeTransitionMatrix(i, j) = TransitionProbSheet.Cells(i + 3, 1 + j)
        Else
            CumulativeTransitionMatrix(i, j) = CumulativeTransitionMatrix(i, j - 1) + TransitionProbSheet.Cells(i +
3, 1 + j)
        End If
    Next j
Next i
'Read max MaxLengthSeries from the sheet and redim the AppointmentSeries array
i = 1
Do Until InputSheet.Cells(3, 4 + i) = ""
    MaxLengthSeries = i
    i = i + 1
Loop
ReDim AppointmentSeries(1 To NrSeries, 1 To MaxLengthSeries)
'redim and assign values to Lengthseries array with the input therapy blocks
ReDim LengthSeries(1 To NrSeries)
For j = 1 To NrSeries
    i = 1
    Do While InputSheet.Cells(3 + j, 4 + i) <> ""
        LengthSeries(j) = i
        i = i + 1
    Loop
Next j
'read the series specifications from the worksheets
For i = 1 To NrSeries
    For j = 1 To MaxLengthSeries
        'correct the probabilities for the cancellation rate and the no show rate
        AppointmentSeries(i, j) = InputSheet.Cells(3 + i, 4 + j) * CancelRate * NoShowRate
    Next j
Next i
For r = 1 To NrRuns
    'loop through patients
    For p = 1 To NrPatients
    TreatmentEnded(p) = False
        'assign patient to their current series and position based on the user's input
```

```
        CurrentSeries = CaseloadSheet.Cells(2 + p, 2).Value
        CurrentPositionSeries = CaseloadSheet.Cells(2 + p, 3).Value
        'find the probability that patient p has an appointment in week = 1 based on the CurrentSeries and the CurrentPosition
        PatientSeries(p, 1) = AppointmentSeries(CurrentSeries, CurrentPositionSeries)
        PastWeeks = 1
        'if the patient is currently in the last treatment week, assign to a new pathway
        If CurrentPositionSeries = LengthSeries(CurrentSeries) Then
          'randomly assign a next series based on the transition probabilities
          Randomize
          RandomNumber = Rnd()
          For i = 1 To NrSeries - 1
              If RandomNumber < CumulativeTransitionMatrix(CurrentSeries, 1) Then
                CurrentSeries = 1
                CurrentPositionSeries = 1
                PatientSeries(p, 1) = AppointmentSeries(CurrentSeries, CurrentPositionSeries)
                PastWeeks = 0
              ElseIf RandomNumber > CumulativeTransitionMatrix(CurrentSeries, i) And RandomNumber <= CumulativeTransitionMatrix(CurrentSeries, i + 1) Then
                CurrentSeries = i
                CurrentPositionSeries = 1
                PatientSeries(p, 1) = AppointmentSeries(CurrentSeries, CurrentPositionSeries)
                PastWeeks = 0
            ElseIf RandomNumber > CumulativeTransitionMatrix(CurrentSeries, NrSeries) Then
                CurrentSeries = NrSeries
                CurrentPositionSeries = 1
                PatientSeries(p, 1) = AppointmentSeries(CurrentSeries, CurrentPositionSeries)
                PastWeeks = 0
            End If
          Next i
        End If
        For w = 2 To NrWeeks
          If TreatmentEnded(p) = False Then
            PastWeeks = PastWeeks + 1
            'check if the patient finished the current appointment series, if so assign a new series
            If CurrentPositionSeries + (PastWeeks - 2) <> LengthSeries(CurrentSeries) Then
```

```
                PatientSeries(p, w) = AppointmentSeries(CurrentSeries, CurrentPositionSeries + PastWeeks - 1)
'base this upon an array with the standard blocks

            Else    'if the current therapy block is finished, then assign a new path or not

                Randomize

                RandomNumber = Rnd()

                'assign the next series based on the transition probabilities from the current series

                If RandomNumber < CumulativeTransitionMatrix(CurrentSeries, 1) Then

                    CurrentSeries = 1

                    CurrentPositionSeries = 1

                    PatientSeries(p, w) = AppointmentSeries(CurrentSeries, CurrentPositionSeries)

                    PastWeeks = 0

                ElseIf RandomNumber > CumulativeTransitionMatrix(CurrentSeries, NrSeries - 1) And
CumulativeTransitionMatrix(CurrentSeries, NrSeries) <> 0 Then

                    CurrentSeries = NrSeries

                    CurrentPositionSeries = 1

                    PatientSeries(p, w) = AppointmentSeries(CurrentSeries, CurrentPositionSeries)

                    PastWeeks = 0

                ElseIf CumulativeTransitionMatrix(CurrentSeries, NrSeries) = 0 Then

                    TreatmentEnded(p) = True

                    Exit For

                Else

                   For j = 1 To NrSeries - 1

                    If RandomNumber > CumulativeTransitionMatrix(CurrentSeries, j) And RandomNumber <=
CumulativeTransitionMatrix(CurrentSeries, j + 1) Then

                    CurrentSeries = j + 1

                    CurrentPositionSeries = j + 1

                    PatientSeries(p, w) = AppointmentSeries(CurrentSeries, CurrentPositionSeries)

                    PastWeeks = 0

                    End If

                  Next j

                End If

                End If

            End If

            Next w

            PastWeeks = 1

        For w = 1 To NrWeeks
```

```
        'if rnd is smaller than probability an appointment will occur then there is an appointment (1), otherwise
not (0)
        If Rnd <= PatientSeries(p, w) Then
          AppointmentsPerPatientPerWeek(p, w) = 1
        End If
      Next w
    Next p
    'assigning durations to all appointments from the duration distribution
      For p = 1 To NrPatients
        For w = 1 To NrWeeks
        Randomize
        RandomNumber = Rnd()
          If RandomNumber <= DurationProb(1, 1) Then
            Duration(p, w) = DurationProb(1, 2) + 0.25
            'if there is an appointment in week w the duration will be multiplied with 1 otherwise with 0
            Duration(p, w) = Duration(p, w) * AppointmentsPerPatientPerWeek(p, w)
          ElseIf RandomNumber > DurationProb(Ranges, 1) Then
             Duration(p, w) = DurationProb(Ranges, 2) + 0.25
             'if there is an appointment in week w the duration will be multiplied with 1 otherwise with 0
             Duration(p, w) = Duration(p, w) * AppointmentsPerPatientPerWeek(p, w)
          End If
          For i = 1 To Ranges - 1
            If RandomNumber > DurationProb(i, 1) And RandomNumber < DurationProb(i + 1, 1) Then
              Duration(p, w) = DurationProb(i, 2) + 0.25
              'if there is an appointment in week w the duration will be multiplied with 1 otherwise with 0
              Duration(p, w) = Duration(p, w) * AppointmentsPerPatientPerWeek(p, w)
            End If
          Next i
          'determine the total duration per week over all runs
          WeeklyDurations(w) = WeeklyDurations(w) + Duration(p, w)
        Next w
      Next p
      For w = 1 To NrWeeks
        'check in which range the weeklyduration(w) is
        If WeeklyDurations(w) < 10 Then
          DurFreqArray(1, w) = DurFreqArray(1, w) + 1
```

```
            ElseIf WeeklyDurations(w) > 10 And WeeklyDurations(w) <= 20 Then
                DurFreqArray(2, w) = DurFreqArray(2, w) + 2
            ElseIf WeeklyDurations(w) > 20 And WeeklyDurations(w) <= 30 Then
                DurFreqArray(3, w) = DurFreqArray(3, w) + 2
            ElseIf WeeklyDurations(w) > 30 And WeeklyDurations(w) <= 40 Then
                DurFreqArray(4, w) = DurFreqArray(4, w) + 2
            ElseIf WeeklyDurations(w) > 40 And WeeklyDurations(w) <= 50 Then
                DurFreqArray(5, w) = DurFreqArray(5, w) + 2
            ElseIf WeeklyDurations(w) > 50 Then
                DurFreqArray(6, w) = DurFreqArray(6, w) + 2
            End If
        Next w
        'erase values but keep the dimensions
        ReDim WeeklyDurations(1 To NrWeeks)
        'determine all weekly appointment totals per run
        For p = 1 To NrPatients
            For w = 1 To NrWeeks
                WeeklyAppointmentsRuns(r, w) = WeeklyAppointmentsRuns(r, w) +
AppointmentsPerPatientPerWeek(p, w)
                AppointmentsPerPatientPerWeek(p, w) = 0
            Next w
        Next p
Next r
'convert duration frequencies to probabilities
For w = 1 To NrWeeks
    'calculate total frequency
    For i = 1 To NrRanges
        Total = Total + DurFreqArray(i, w)
    Next i
    'calculate the probabilities
    For i = 1 To NrRanges
        DurProbArray(i, w) = DurFreqArray(i, w) / Total
    Next i
    Total = 0
Next w
'find the distribution for the forecasted workload in hours per week
```

```
For w = 1 To NrWeeks
    For r = 1 To NrRuns
        AvWeeklyAppointments = AvWeeklyAppointments + WeeklyAppointmentsRuns(r, w)
    Next r
    ForecastWeeklyAppointments(w) = AvWeeklyAppointments / NrRuns
    AvWeeklyAppointments = 0 'initialize
Next w
'write output to the simulation sheet for table 1
SimulationSheet.Range("B5:ZZ5").ClearContents
For w = 1 To NrWeeks
    SimulationSheet.Cells(5, w + 1) = ForecastWeeklyAppointments(w)
Next w
    'write output to the simulation sheet for table 2
SimulationSheet.Range("B10:ZZ15").ClearContents
For i = 1 To NrRanges
    For w = 1 To NrWeeks
        SimulationSheet.Cells(9 + i, w + 1) = DurProbArray(i, w)
    Next w
Next i
End Sub
```

References

Armstrong, J. (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Springer Science & Business Media.

Astfalck, L., Bertolacci, M., & Cripps, E. (2023). Evaluating probabilistic forecasts for maritime engineering operations. *Data-centric Engineering*, 4. https://doi.org/10.1017/dce.2023.11

Bazionis, I. K., & Georgilakis, P. S. (2021). Review of Deterministic and Probabilistic Wind Power Forecasting: Models, Methods, and Future Research. *Electricity*, 2(1), 13–47. https://doi.org/10.3390/electricity2010002

Brynjolfsson, E., Hitt, L. M., & Kim, H. (2011). Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? Social Science Research Network. https://doi.org/10.2139/ssrn.1819486

CHOIR. (n.d.). *Integral Capacity Planning and Management*. Universiteit Twente. https://www.utwente.nl/en/choir/research/Projects/ICM/

Davis, J. A. (2012). *Measuring Marketing* (2nd ed.). Wileys. https://doi.org/10.1002/9781119199205

Dekker, M., Van Donselaar, K. K., & Ouwehand, P. P. (2004). How to use aggregation and combined forecasting to improve seasonal demand forecasts. *International Journal of Production Economics*, 90(2), 151–167. https://doi.org/10.1016/j.ijpe.2004.02.004

Earl, D. J., & Deem, M. W. (2008). Monte Carlo Simulations. *In Methods in molecular biology* (pp. 25–36). Springer Science+Business Media. https://doi.org/10.1007/978-1-59745-177-2_2

Gabbay, U., & Bukchin, M. (2009). Does daily nurse staffing match ward workload variability? Three hospitals' experiences. International journal of health care quality assurance.

Gilli, M., Maringer, D., & Schumann, E. (2019). Generating random numbers. *In Elsevier eBooks* (pp. 103–132). https://doi.org/10.1016/b978-0-12-815065-8.00017-0

Gneiting, T., & Katzfuss, M. (2014). Probabilistic Forecasting. Annual Review of Statistics and Its Application, 1(1), 125–151. https://doi.org/10.1146/annurev-statistics-062713-085831

Ham, C., & Alberti. (2002). The medical profession, the public, and the government. *BMJ, 324*(7341), 838–842. https://doi.org/10.1136/bmj.324.7341.838

Heerkens, H., & Van Winden, A. (2021). *Solving Managerial Problems Systematically*. Routledge.

Horcas Aguilera, J. M., Márquez, A., Galindo Duarte, J. Á., & Benavides Cuevas, D. F. (2021). Monte Carlo Simulations for Variability Analyses in Highly Configurable Systems. *In CWS/ConfWS 2021: 23rd International Configuration Workshop (2021)*, pp. 37-44.. CEUR Workshop Proceedings (CEUR-WS. org).

Ivanov, D., Tsipoulanidis, A., & Schönberger, J. (2018). Demand Forecasting. *In Springer texts in business and economics* (pp. 319–333). Springer International Publishing. https://doi.org/10.1007/978-3-319-94313-8_11

Keim, D. A. (2001). *Visual exploration of large data sets. Communications of the ACM, 44*(8), 38–44. https://doi.org/10.1145/381641.381656

Loughry, J., van Hemert, J., & Schoofs, L. (2000). Efficiently enumerating the subsets of a set. *applied-math. org/subset. pdf*.

Lund, T. (2012). Combining Qualitative and Quantitative Approaches: Some Arguments for Mixed Methods Research. *Scandinavian Journal of Educational Research*, 56(2), 155–165. https://doi.org/10.1080/00313831.2011.568674

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (2008). *Forecasting Methods and applications* (3rd ed.). John Wiley & Sons.

Ministerie van Volksgezondheid, Welzijn en Sport. (2023, March 27). *Toezicht op aanpak wachttijden GGZ*. Geestelijke Gezondheidszorg | Inspectie Gezondheidszorg En Jeugd. Retrieved April 24, 2023, from https://www.igj.nl/zorgsectoren/geestelijke-gezondheidszorg/wachtijden-in-de-ggz

Ordu, M., Demir, E., & Tofallis, C. (2019). A comprehensive modelling framework to forecast the demand for all hospital services. *International Journal of Health Planning and Management*, *34*(2). https://doi.org/10.1002/hpm.2771

Ordu, M., Demir, E., & Tofallis, C. (2020). A decision support system for demand and capacity modelling of an accident and emergency department. *Health Systems*, 9(1), 31–56. https://doi.org/10.1080/20476965.2018.1561161

ORTEC. (n.d.). *Optimize healthcare with ORTEC's capacity management software*. Retrieved July 9, 2023, from https://ortec.com/en/solutions/integral-capacity-management

Papadopoulos, C., & Yeung, H. (2001). Uncertainty estimation and Monte Carlo simulation method. *Flow Measurement and Instrumentation*, 12(4), 291–298. https://doi.org/10.1016/s0955-5986(01)00015-2

Papoulis, A. (1984). "Bernoulli Trials". Probability, Random Variables, and Stochastic Processes (2nd ed.). New York: McGraw-Hill. pp. 57–63.

Plsek, P. E., & Greenhalgh, T. (2001). Complexity science: The challenge of complexity in health care. *BMJ*, 323(7313), 625–628. https://doi.org/10.1136/bmj.323.7313.625

Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, *1*(1), 51–59. https://doi.org/10.1089/big.2013.1508

Rauschenbach, U., & Schumann, H. (1999). Demand-driven image transmission with levels of detail and regions of interest. *Computers & Graphics*, 23(6), 857–866. https://doi.org/10.1016/s0097-8493(99)00116-8

Schmidt, G., & Wilhelm, W. E. (2000). Strategic, tactical and operational decisions in multi-national logistics networks: A review and discussion of modelling issues. *International Journal of Production Research*, 38(7), 1501–1523. https://doi.org/10.1080/002075400188690

Sousa, A. E., Pereira, M. O., & Matos, H. A. (2020). Stochastic Analyses on Relative Viscosity of Water-in-oil Emulsions. *In Elsevier eBooks* (pp. 121–126). https://doi.org/10.1016/b978-0-12-823377-1.50021-5

Tello, M. J., Reich, E., Puckey, J., Maff, R., Garcia-Arce, A., Bhattacharya, B. S., & Iyer, G. (2022). Machine learning based forecast for the prediction of inpatient bed demand. *BMC Medical Informatics and Decision Making*, 22(1). https://doi.org/10.1186/s12911-022-01787-9

Vanberkel, P. T. (2011). *Interacting Hospital Departments and Uncertain Patient Flows: Theoretical Models and Applications.*

Williams, R. (2002). Complexity, uncertainty and decision-making in an evidence-based world. *Current Opinion in Psychiatry*, *15*(4), 343-347.

White, S. A. (2004). Introduction to BPMN. *Ibm Cooperation*, *2*(0), 0.

Wood, S. J., Michaelides, G., & Totterdell, P. (2013). The impact of fluctuating workloads on well-being and the mediating role of work–nonwork interference in this relationship. *Journal of Occupational Health Psychology*, 18(1), 106–119. https://doi.org/10.1037/a0031067

Wand, M. P. (1997). Data-Based choice of histogram BIN width. *The American Statistician*, 51(1), 59–64. https://doi.org/10.1080/00031305.1997.10473591

Wang, Y. H. (1993). On the number of successes in independent trials. *Statistica Sinica*, 3(2), 295–312. http://www.jstor.org/stable/24304959