Team Virginia-Mason
Project Proposal

Team Bios

Jennifer Kim
I majored in economics for undergrad, and have experience in legal editing and digital marketing, from which I gained solid communication, analytical and marketing skills. I am passionate about leveraging data to solve real world problems. I look forward to applying the knowledge and skills I gained in the MSDS program to real problems in the healthcare industry.

Vanja Glisic
My team role is the communicator between our team and outside stakeholders. I have been working for about six years in the insurance industry, with most of that time spent in a data science department, specifically on a demand and decision science team. My work focuses on understanding customer behavior and how pricing decisions impact it. In this work, I spend a lot of time coding in Python and occasionally writing SQL queries.

Hari haran Suresh kumar
I have a bachelor's in Computer Science and two years of experience at Cognizant Technology Solutions as an analyst. I worked on an internal tool that relied on postgreSQL and SQl queries, and frequent analysis. I am excited to use this capstone to apply my skills to real problems, and better position myself for data driven decision making.

Problem Statement

The goal of this project is to support surgeons in planning spinopelvic surgery by designing an algorithm that generates multiple viable surgical plans. The surgical plans will be ranked based on their ability to achieve radiographic alignment of the spine. The surgical plans consist of a series of spinopelvic parameters to be adjusted during surgery based on the recommendations from our algorithm. The target spinopelvic alignment will be informed by literature review, surgeon-defined benchmarks, and consultation with other surgeons. We are directly working with a team of surgeons at Virginia Mason Medical Center, who provide clinical insight for this project. Currently, surgeons heavily rely on manual calculations and experience to decide on an appropriate alignment plan, and the results of our project will provide a more systematic and data-driven approach for developing those plans.

Our approach is to build an algorithm that adjusts spinopelvic parameters and generates several feasible surgical plans. Each plan will be evaluated using a spinopelvic alignment fitness score, which we will define in collaboration with the surgeons. We expect to provide multiple options that are feasible for a given input set of spinopelvic parameters for a given patient. The surgeons will then be able to use their medical judgement and surgery simulator technology to select the option that is best for the patient. It is not the expectation that we will select one best option; rather, we aim to provide them with multiple possible plans for reaching radiographic alignment, with which they can use their medical judgement to pick the best one.

Some challenges we may encounter are selecting criteria for what is considered radiographic alignment of the spine. Different surgeons might have many ideas about what constitutes a successful surgery. We will need to conduct more research and collect expert advice to understand this better. We are working with the sponsor team for both of these. They have mentioned using alignment targets such as the GAP score, SRS-Schwab, Roussouly Laouissat, and the L4-L1-Hip Axis. We will need to develop a fitness function which will be used in the algorithm to assess the spinopelvic alignment, given a set of spinopelvic parameters. Another challenge we may encounter is that our data set only has 120 rows of patient records; this may limit our ability to create an ML model as the fitness function. We may be able to include a black-box model to query the outcome for proposed surgical plans, which may allow us to proceed with a small number of training vectors.

Data Pipeline

Our Virginia-Mason sponsor shared the dataset with us in Excel format post the capstone kickoff. It consists of 120 complex spine alignment surgery cases with 112 pre-operative and post-operative (collected after surgery) parameters, with each row corresponding to a single surgery case. It also includes the age and gender information of the patients. We will keep the dataset in our shared Google drive along with other capstone related documents.

Since the dataset contains only 120 spine cases, we are a bit concerned that this may be insufficient for a more rigid analysis. However, our team is hopeful that a proof of concept would also be valuable for future applications once more data becomes available.

Also, the postoperative parameters are missing values for some of the spine surgery cases. In particular, we are missing the alignment indication column, which is the measure of success we are aiming for. Our sponsor is aware of this and plans to provide the measurements at a later date.

Proposed Schedule

We created a tentative schedule board using the Monday management app. The app allows us to track tasks and deadlines and manage milestones. So far, it appears to be an efficient visualization and management tool for monitoring the entire project timeline, which will help us check the overall progress and classify and assign tasks by priority level for timely execution.

Before January, we plan to (deliverables)

- Agree on an initial definition of the alignment fitness score.
- Identify which parameters matter most
- Clarify how to distinguish between viable and non-viable plans

This will help implement an algorithm, run experiments on the dataset, and evaluate whether the generated plans look reasonable from a clinical point of view.

We will discuss the schedule and milestones with our sponsor and make adjustments as necessary.

Summary of Background Research

We did some background research on the following topics:

- Spinopelvic parameters (i.e. normal ranges and relationships between them)
- Using genetic algorithms in image-guided neurosurgery
- Automated spine measurement using deep learning

<u>Spinopelvic parameters</u>
One article pulled together information from multiple studies to provide a comprehensive overview on some key spinopelvic parameters and what it means for a spine to be radiographically aligned. The sources did not always specify sample sizes or other considerations so we should still confirm any quantitative relationships with our sponsors before relying on them. The paper notes that there is not just one way to have an aligned spine in the normal population, but generally keeping the center of gravity near the center of the hips helps. Some parameters, such as pelvic incidence and spinopelvic angle, have known geometric relationships, and others, like lumbar lordosis, have formulas suggesting optimal values. For example, pelvic incidence (PI) is equal to the sum of pelvic tilts (PT) and sacral slope (SS), and the optimal value of lumbar lordosis is the pelvic incidence plus 9 degrees. This information could help guide our objective function and constraints.

<u>Genetic algorithms in image-guided neurosurgery</u>
A key takeaway from this study is that the fitness function is based on a reciprocal of target registration error as the goal was to minimize target registration error. Their intended output was to find optimized configurations of coordinate markers on the patient's head for locating the lesion. The genetic algorithm involves passing down the genetic material of a set of individuals from a randomly determined population, each representing a possible fiducial configuration, to the next generation. The genetic material consists of a number of possible markers. Crossover is performed on each pair of individuals selected as performed. They strived to look for global minima, instead of being trapped with local optimal points, by introducing mutations and new individuals to each new population.

The algorithm's efficiency was evaluated in simulated settings by comparing the TRE values obtained from optimized configurations with unoptimized, randomly generated marker sets, for an internal target and an external surface target. Results showed a 15.9 +/- 0.7 % improvement for internal targets and a 51.4 +/- 8.9% improvement for surface targets.

<u>Automated spine measurement using deep learning</u>

One study we reviewed explored the use of deep learning and computer vision to automatically measure spinal alignment on whole-spine lateral X-rays. The authors trained a Mask R-CNN model to identify and segment each vertebra and the pelvis, then used computer vision methods to compute the spinopelvic angles that surgeons normally measure by hand. Their automated system produced measurements that were both accurate and more consistent than those of experienced clinicians, showing how machine learning can reduce variability and speed up routine analysis. The approach still has limitations—particularly in cases with severe deformities, hardware, or poorly visible anatomy—but it demonstrates that automated alignment measurement is feasible and could support future tools for surgical planning and research.

Proposed Solution

The overall goal is to design an algorithm which generates multiple feasible surgical plans for spinopelvic alignment. Each surgical plan is a sequence of surgical steps where each step adjusts one of the spinopelvic parameters/measurements by some determined amount. The goal of each surgical plan is to be a set of steps that leads to good alignment of the spine, which the surgeons can then evaluate with medical knowledge to select the best option to implement.

Deliverables

<u>Minimal Viable Product</u>
The minimal viable product for this project is the design and set-up of a solution that most effectively produces surgical plans for a given patient’s spinopelvic parameters. We will provide this to the team as a github repo containing the code.

This solution consists of two major components, the first being a fitness function for assessing the alignment of a spine, given a set of spinopelvic parameters. We will need to come up with the best option, given the information and data we can collect. Current options we are evaluating are (1) creating a mathematical equation based on physical properties, or (2) build an ML model with the data we have been provided. This can serve as a drop-in function that can be replaced in the future if a better option becomes available.

The second deliverable is an algorithm which uses the fitness function to create a set of best feasible solutions for reaching radiographic alignment of the spine, given a starting set of spinopelvic parameters. Currently we are evaluating optimization methods such as the Genetic Algorithm, Pareto, and Particle Swarm.

Additionally, we will deliver a written paper documenting the problem, selected methodology, and in depth explanations of why we decided on the methodology.

<u>Stretch Goals</u>
The stretch goals of this project are around the accuracy and complexity of the fitness function and the constraints which we choose to model. Once we have an initial minimal viable product set up to provide a solution, we can refine the fitness function to better predict alignment of the

spine given the parameters we have available. Also, we can put more time into researching how shifting one angle will affect the other angles, which will better represent the state of the spine. Refining these components will result in better surgical solutions given by our algorithm.

Contingency Plan

Our contingency plan is to create a very simple set up of the algorithm, with the bare minimum complexity to create feasible surgical plans. This set up will have the end-to-end setup, but the individual components will be simple versions of the models. For example, if we are unable to build a more complex or accurate physical equation or model to assess spinal alignment, we could instead build an ML model using our data (assuming we can assess the post-operative alignment in the existing data). However, this model may be overfit and have low accuracy given the limited data available initially.

Risks and Benefits

Risks

Potential bias in algorithms can lead to overrepresentation of particular ages or spinal deformity types might cause the algorithm to produce plans that work well for several patient types and poorly for others. The mean age in the patient cases is 66.5. The dataset only contains 120 cases, which can lead to overfitting and low accuracy if using an ML model for the fitness function.

Also, the algorithm might generate plans that are mathematically feasible but not clinically reasonable. Some parameter adjustments generated in plans may be impossible or unsafe. As we will be suggesting surgical plans that can be used for real-life patients, we should take extra care to test and validate the algorithm to ensure it is clinically sound and safe. We can do our best to add constraints to the algorithm to prevent this, but it may be difficult to capture all physical constraints without significant research beyond project scope.

As the fitness function and constraints rely heavily on surgeon guidance, if the surgeon availability is limited, the team may not have enough domain knowledge to tune the model properly.

Benefits

This approach provides an innovative and data-driven method for finding optimal solutions to the problem and sets precedent for future research in the field. It has the potential to significantly reduce time and effort that goes into developing surgical plans manually. With this time savings, surgeons can reallocate their efforts toward evaluating more surgical plans than would typically be feasible. By systematically exploring more options, the algorithm can uncover surgical plans that were not immediately apparent through the manual planning process. It also can help provide more consistency across surgical planning and possibly reduce variability from manual calculations. By providing surgeons with more information to guide their final surgical plan, this project can contribute to improved patient outcomes.