

PhD Thesis Subject

Research Project

*Multimodal Modeling for the Diagnosis and Prognostic
Assessment of Shoulder Disorders*

Team and Supervisors

- Pr Marc-Olivier Gauci (CHU of Nice, University Institute of locomotion and Sport, Director of ICARE lab (IBV, INSERM U1091). marc-olivier.gauci@univ-cotedazur.fr

- Dr Jean-Baptiste Masson (DBC, INRIA, Institut Pasteur, Paris, Institut 3IA PRAIRIE). jbmasson@pasteur.fr

Scientific Context and Challenges

The Clinical Challenge

In clinical practice, particularly in shoulder orthopaedics, decision-making does not rely on a simple causal approach. Medical reasoning does not follow a linear logic such as “symptom → imaging → diagnosis → treatment.” Instead, it is an integrative process through which the surgeon progressively builds a body of evidence. In order to establish a diagnosis and propose an appropriate therapeutic solution, the surgeon must combine heterogeneous sources of information, including the clinical examination (pain, range of motion, functional scores), imaging data (X-rays, MRI, CT scans), and the patient’s individual context (age, comorbidities, occupational activity).

These elements are often incomplete and may sometimes even appear contradictory. In this context, imaging does not represent a direct consequence of the clinical presentation, but rather a complementary dimension. Although this integrative reasoning lies at the heart of medical practice, it remains an implicit cognitive process that is highly dependent on the practitioner’s experience and is therefore difficult to formalize.

The Scientific Challenge

The main scientific challenge of this project is to formalize this complex clinical reasoning so that it can be leveraged by algorithms. The aim is no longer to establish a simplified cause-and-effect relationship between the different modalities, but rather to model their complementarity in order to structure medical decision-making.

Recent advances in machine learning, computer vision, and natural language processing (NLP) now offer new opportunities. It is becoming possible to extract explicit geometric and morphometric descriptors from medical imaging, including in complex pathological settings. At the same time, NLP techniques make it possible to structure clinical textual data that have so far been underexploited, by identifying relevant variables related to the patient’s medical history, symptoms, or lifestyle and clinical context.

The challenge is to develop multimodal approaches capable of learning how to combine these different sources of information without arbitrarily prioritizing one over another. This modelling approach will make it possible to objectify clinical reasoning and the characterization of shoulder pathologies in order to support therapeutic decision-making. Within this framework, predicting

disease progression will naturally extend this modelling process once the underlying mechanisms have been properly captured.

Scientific Objectives

The main objective of this project is to develop a multimodal computational model integrating clinical and imaging data in order to structure and objectify clinical reasoning in shoulder pathologies, with the aim of improving diagnosis and therapeutic decision-making.

To achieve this objective, the project has been divided into three major scientific objectives:

1. Geometric Extraction and Modelling (Imaging)

The objective is to use the raw input images to enable the model to capture the pathology as a whole (e.g., detection of osteoarthritis) before analysing its geometry. The next step will consist of the automatic, selective, and progressive extraction of targeted morphometric descriptors (angles, curvature, wear volumes) and their spatial reference systems whenever required by the pathological context.

This progressive strategy ensures the feasibility of the project on large cohorts while extracting the mathematical variables required to objectively quantify the biomechanical degradation of the joint, even in the presence of diverse and complex pathologies.

2. Structuring of Textual Data (Text)

The second objective is to automatically extract the patient's functional dimension from unstructured data (radiology reports, medical records) using Natural Language Processing (NLP) techniques.

The challenge is to standardize and project these heterogeneous texts into a structured space in order to model key variables such as pain history, comorbidities, and patient-reported experience. In addition to these extracted data, the development of adaptive questionnaires (e.g., Computerized Adaptive Testing, CAT) specifically designed for shoulder conditions is also envisaged. These questionnaires would generate accurate functional scores that could directly enrich the model.

3. Multimodal Integration, Diagnostic Support, and Decision-Making

The final objective is to fuse the geometric parameters derived from imaging with the structured clinical variables extracted from text. The challenge will be to computationally reproduce the complex integration process underlying medical reasoning.

By combining both the physical structure of the shoulder and the patient's clinical profile, this multimodal model will bring together a body of evidence to first refine the diagnosis and then model the optimal therapeutic decision-making process.

Originality of the Approach

The originality of this project lies in its methodological and clinical positioning.

From a technological perspective, Vision-Language Models (VLMs) have demonstrated strong performance in many fields, but they remain limited and are often poorly suited to the reality and scale of local medical databases. Our approach proposes a robust alternative based on the extraction of structured parameters.

From a clinical perspective, while medical innovation has largely focused on surgical planning, which offers high commercial potential, this project addresses a relatively neglected yet essential area centred on the earlier stages of patient care (diagnosis and patient pathway), which are major contributors to diagnostic uncertainty and delays in appropriate care.

Implementation Milestones

The PhD project is structured around a three-phase methodological progression:

Phase 1 / Year 1 – Formalization of Representation Spaces (Extraction)

- Development of robust, fully automated 3D geometric extraction algorithms applied to computed tomography (CT) scans.
- Implementation of an NLP pipeline for structuring clinical variables and creating the patient state space.

Milestone: Establishment of a fully structured multimodal database.

Phase 2 / Year 2 – Coupling Strategies and Predictive Modelling

- Exploration of mathematical architectures for data fusion.
- The objective of this phase is to test different models capable of learning the complementarity between anatomy (CT) and the clinical profile (text).
- The aim will be to mathematically model this body of multimodal evidence in order to reproduce the mechanisms underlying disease characterization and medical decision-making.

Milestone: Proof of Concept (PoC) of a multimodal clinical characterization model.

Phase 3 / Year 3 – Robustness, Uncertainty Quantification, and Translation

- Rigorous evaluation of the model's limitations.
- Assessment of prediction sensitivity to outlier data (extreme bone deformities, incomplete texts).
- Quantification of the uncertainties associated with the model outputs.
- Evaluation of the clinical impact, particularly in terms of reducing diagnostic uncertainty and delays in patient management.

Milestone: Finalization of the clinical decision-support algorithm for diagnosis and prognosis.