

UNIVERSITÀ
DEGLI STUDI
DI BRESCIA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

DOTTORATO DI RICERCA IN TECHNOLOGY FOR HEALTH

ING-INF/05 - SISTEMI DI ELABORAZIONE DELLE INFORMAZIONI

Ciclo XXXVIII

**INTEGRATING ARTIFICIAL INTELLIGENCE
INTO DIAGNOSTIC IMAGING:
FROM AI-DRIVEN SIMULATION AND QUANTITATIVE IMAGING TO
PERSONALIZED CONTRAST ADMINISTRATION IN CT**

Giuseppe Roberto BONFITTO

Prof. Alberto SIGNORONI

Abstract - Italiano

Background

La somministrazione di mezzo di contrasto iodato nella tomografia computerizzata si basa ancora prevalentemente su protocolli fissi o aggiustati per il peso corporeo, che non tengono conto della marcata variabilità inter-individuale in termini di composizione corporea, volume vascolare e caratteristiche di attenuazione. Questa mancanza di personalizzazione determina un potenziamento non uniforme, un utilizzo eccessivo di iodio e frequenti scostamenti dagli obiettivi diagnostici. In parallelo, l'intelligenza artificiale si è affermata come strumento promettente in radiologia, con applicazioni che spaziano dal supporto alla comunicazione con il paziente all'ottimizzazione quantitativa dell'imaging.

Methods

Il lavoro di dottorato è stato strutturato su tre componenti metodologiche principali: la simulazione di scenari di claustrofobia in risonanza magnetica mediante large language models, per valutare il ruolo delle strategie comunicative sulla compliance del paziente; la creazione di un dataset multicentrico per l'analisi della composizione corporea in TC, basato su segmentazioni manuali ai livelli L1 e L3, quale standard di riferimento morfometrico; lo sviluppo di una pipeline di deep learning multimodale che combina l'encoding visivo delle immagini scout TC basato su DINOv2 con metadati strutturati di acquisizione e di dosaggio, finalizzata alla predizione del potenziamento vascolare e parenchimale. La deviazione del potenziamento è stata valutata quantitativamente in una coorte di 518 esami TC con mezzo di contrasto, ed è stato implementato un algoritmo di inversione della dose per generare raccomandazioni iodiche paziente-specifiche in un contesto simulativo.

Results

Oltre il 50% degli esami analizzati risultava al di fuori delle soglie diagnostiche accettabili di potenziamento per il parenchima epatico e renale, evidenziando i limiti dei protocolli standard. Il modello predittivo ha raggiunto errori assoluti medi inferiori a 15 HU negli organi addominali e ha consentito una valutazione simulativa (in silico) di strategie di dosaggio personalizzate. Tali strategie hanno ridotto la variabilità inter-paziente del potenziamento del 25–30% e hanno indicato un potenziale di riduzione della dose di iodio nei pazienti con profili di attenuazione più bassi. Un'analisi ambientale accessoria ha mostrato che il 31,7% del mezzo di contrasto caricato in sistemi multidose viene sprecato, supportando i potenziali benefici ecologici di approcci di dosaggio di precisione.

Conclusions

Questa tesi dimostra la fattibilità tecnica e traslazionale di un framework basato su intelligenza artificiale per la somministrazione personalizzata del mezzo di contrasto in TC. Sfruttando le informazioni derivanti da immagini scout a bassa dose e da dati strutturati di protocollo, l'approccio consente raccomandazioni di dose individualizzate, finalizzate a migliorare l'uniformità del potenziamento e a ridurre l'esposizione iodica non necessaria, contribuendo anche agli obiettivi di sostenibilità istituzionale. Sebbene il framework sia progettato per una integrazione fluida nei flussi di lavoro correnti, sono necessari ulteriori studi per validarne le prestazioni in contesti multicentrici e per incorporare stime di incertezza prima dell'impiego clinico. La valutazione è stata condotta retrospettivamente mediante analisi simulativa; studi prospettici in modalità silente e validazioni multicentriche sono richiesti per confermarne efficacia clinica e robustezza prima dell'adozione routinaria.

Abstract - English

Background

Iodinated contrast administration in computed tomography remains predominantly reliant on fixed or weight-adjusted protocols that overlook substantial inter-individual variability in body habitus, vascular volume, and attenuation characteristics. This lack of personalization leads to inconsistent enhancement, excessive iodine usage, and frequent deviation from diagnostic targets. Concurrently, artificial intelligence has emerged as a promising tool in radiology, with applications extending from patient communication support to quantitative imaging optimisation.

Methods

This doctoral work was structured around three core methodological components: simulation of MRI claustrophobia scenarios using large language models to evaluate the role of communicative strategies in patient compliance; creation of a multicentre dataset for CT-based body composition analysis via manual segmentation at L1 and L3 levels, providing a morphometric reference standard; and development of a multimodal deep learning pipeline combining DINOv2-based vision encoding of CT scout images with structured acquisition and dosing metadata to predict vascular and parenchymal enhancement. Enhancement deviation was quantitatively assessed in a cohort of 518 contrast-enhanced CT examinations, and a dose-inversion algorithm was implemented to generate patient-specific iodine recommendations under a simulation framework.

Results

Over 50% of analysed examinations fell outside acceptable diagnostic enhancement thresholds for liver and renal parenchyma, highlighting the limitations of standard protocols. The predictive model achieved mean absolute errors below 15 HU in abdominal organs and enabled simulation-based (in silico) evaluation of individualized dosing strategies. These simulation-based strategies reduced inter-patient enhancement variability by 25–30% and indicated the potential for iodine dose reduction in patients with lower attenuation profiles. An ancillary environmental analysis showed that 31.7% of contrast material loaded into multi-dose systems was wasted, supporting the potential ecological benefits of precision dosing approaches.

Conclusions

This thesis demonstrates the technical and translational feasibility of an AI-powered framework for personalized contrast media administration in CT imaging. By leveraging information from low-dose scout images and structured protocol data, the approach enables individualized dose recommendations aimed at improving enhancement uniformity and reducing unnecessary iodine exposure, and contribute to institutional sustainability goals. While the proposed framework is designed for seamless integration into current workflows, further work is needed to validate its performance across multiple centres and to incorporate uncertainty estimation prior to clinical deployment.

The framework was evaluated retrospectively using simulation-based analyses; prospective silent-mode validation and multicentre studies are required to confirm clinical effectiveness and robustness prior to routine clinical deployment.

Abbreviations and Acronyms

AI – Artificial Intelligence

CT – Computed Tomography

MRI – Magnetic Resonance Imaging

PET – Positron Emission Tomography

FDG – Fluorodeoxyglucose

HU – Hounsfield Unit

ICM – Iodinated Contrast Media

ROI – Region of Interest

LLM – Large Language Model

CNN – Convolutional Neural Network

GAN – Generative Adversarial Network

SSIM – Structural Similarity Index Measure

MAE – Mean Absolute Error

MAPE – Mean Absolute Percentage Error

SUVR – Standardized Uptake Value Ratio

BSA – Body Surface Area

LBW – Lean Body Weight

PACS – Picture Archiving and Communication System

RIS – Radiology Information System

FWHM – Full Width at Half Maximum

kVp – Kilovolt Peak

DICOM – Digital Imaging and Communications in Medicine

mgI – Milligrams of Iodine

mgI/kg – Milligrams of Iodine per Kilogram

ViT – Vision Transformer

DINOv2 – Self-supervised Vision Transformer model (Facebook/Meta AI)

Contents

Abstract..... 2

List of Figures..... 10

List of Tables..... 10

Chapter 1..... 11

Introduction and Rationale..... 11

1.1 Background..... 11

1.2 Rationale of the PhD: Building a Coherent Multistep Research Program..... 12

1.3 Research Pillars: Competence Building Across Domains..... 14

1.4 The Target Clinical Need: Optimising Iodinated Contrast Use in CT 16

1.5 Research Objectives 18

 Primary Objective 18

 Secondary Objectives 19

1.6 Structure of the Thesis 20

1.7 Project Timeline..... 22

Chapter 2..... 24

Large Language Models and AI-assisted Simulation in Radiography Education 24

2.1 Introduction..... 24

2.2 Theoretical Background on LLMs 25

2.3 Rationale and Objectives 26

2.4 Materials and Methods 27

2.5 Results 29

2.6 Qualitative Analysis of Dialogue Dynamics 30

2.7 Educational and Organisational Implications 31

2.8 Contributions and Future Work..... 32

Chapter 3	43
CT Body Composition Analysis: Semi-automatic Segmentation of L1 and L3 during International Fellowship	43
3.1 Introduction and Motivation	43
3.2 Materials and Methods	44
3.2.1 Dataset.....	44
3.2.2 Segmentation Protocol.....	44
3.2.3 Extracted Parameters	46
3.3 Results	46
3.4 Discussion	47
3.5 Limitations and Future Perspectives	48
3.6 Contributions.....	49
 Chapter 4.....	 50
Synthetic FDG-PET Generation from Early-Phase Amyloid PET: Experimental Framework for AI Architectures	50
4.1 Introduction and Motivation	50
4.2 Materials and Methods	52
4.2.1 Dataset.....	52
4.2.2 Pre-processing	52
4.3 AI Architectures	53
4.4 Training and Evaluation	55
4.5 Results	55
4.6 Discussion	56
4.7 Limitations and Future Perspectives	57
4.8 Contributions.....	57
4.9 Dissemination and Conference Presentations	57

Chapter 5 59

AI-Driven Personalization of Iodinated Contrast Media Administration in CT 59

5.1 Introduction 59

5.1.1 Performance of DINOv2 in Medical Imaging 61

5.2 Study Population and CT Dataset 62

5.3 Image Analysis Pipeline from Scout..... 64

5.3.1 Enhancement Deviation and Threshold Analysis 67

5.4 AI Model for Contrast Dose Personalization 71

5.4.1 Visual Feature Extraction with DINOv2 72

5.4.2 Rationale for the Multimodal Architecture 73

5.5 Forward Prediction Model..... 75

5.5.1 Model Training and Validation 75

5.5.2 Biological Interpretation of the Forward Function..... 77

5.5.3 Role of the Forward Model in Dose Inversion..... 78

5.6 Inversion Framework..... 78

5.6.1 Mathematical Formulation of the Inverse Problem 79

5.6.2 Iterative Dose Optimization Strategy 80

5.6.3 Handling Physiological and Physical Non-Linearities..... 80

5.6.4 Outputs of the Inversion Procedure 81

5.7 Evaluation and Clinical Validation 82

5.8 Integration into the Clinical Workflow 84

5.9 Limitations 86

Conclusions..... 89

References..... 92

List of Figures

FIGURE 1 THREE-YEAR GANTT CHART SUMMARISING THE TEMPORAL ORGANISATION OF THE RESEARCH ACTIVITIES, INCLUDING LITERATURE REVIEW.	23
FIGURE 2: NUMBER OF RESPONSES NEEDED TO PERSUADE THE PATIENT PER PARTICIPANT.	30
FIGURE 3: THE USE OF METHODS FOR SUPPORTING PATIENTS WITH CLAUSTROPHOBIA AT DIFFERENT LEVELS OF VIRTUAL PATIENTS. [SUPPORTIVE INSTRUCTIONS (BLUE); MILD SEDATION (ORANGE); MUSICAL THERAPY (RED); VIRTUAL REALTY (PURPLE); BREATHING TECHNIQUE (GREEN)]	31
FIGURE 4 – ALZHEIMER’S BIOMARKER FRAMEWORK	51
FIGURE 5: VISUALIZING DEEP LEARNING	53
FIGURE 6: GAN AND DIFFUSION PROCESSES	54
FIGURE 7: CT CONTRAST PERSONALIZATION PIPELINE.....	64
FIGURE 8 - (LIVER) SHOWS A SYMMETRIC DISTRIBUTION CENTERED SLIGHTLY ABOVE THE IDEAL VALUE, SUGGESTING A TENDENCY TOWARD MILD OVER-ENHANCEMENT IN SMALLER BODY SIZES.....	69
FIGURE 9 - (KIDNEY) PRESENTS A RIGHT-SKEWED PATTERN, REFLECTING DELAYED OR UNEVEN PARENCHYMAL ENHANCEMENT, POSSIBLY DUE TO TIMING DIFFERENCES IN BOLUS ARRIVAL OR RENAL FUNCTION HETEROGENEITY.....	69
FIGURE 10 - MEAN ENHANCEMENT INTENSITY CURVES (HU) FOR HEPATIC AND RENAL COMPARTMENTS DURING THE PORTAL PHASE BASELINE WINDOW.	70
FIGURE 11: BLOCK DIAGRAM OF THE MULTIMODAL ARCHITECTURE	74
FIGURE 12 – AGGREGATED DRUG IMPACT	75
FIGURE 13 – KIDNEY, LIVER INTENSITY.....	77
FIGURE 14 – MSE-MAE MODEL.....	78
FIGURE 15 – SCOUT IMAGE AND IMPACT DRUG VOLUME.....	79

List of Tables

TABLE 1: CHARACTERISTICS OF "VIRTUAL PATIENTS"	27
TABLE 2 - SUMMARY OF ENHANCEMENT DEVIATION ANALYSIS.....	71

Chapter 1

Introduction and Rationale

1.1 Background

Over the past two decades, diagnostic imaging has evolved rapidly due to advances in acquisition technology, increasing image complexity, and the digitization of clinical workflows. Computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) are now central pillars of clinical decision-making, but their growing sophistication has also revealed structural limitations in workflow efficiency, protocol standardization, and resource optimization.

Parallel to these developments, artificial intelligence (AI) has emerged as a disruptive force capable of addressing several long-standing challenges in radiology. The convergence between large datasets, powerful computational frameworks, and deep learning models has fostered a paradigm shift from protocol-driven imaging to data-driven imaging. Within this trajectory, three families of AI techniques have proven particularly influential:

1. Large Language Models (LLMs), designed to interpret, generate, and structure clinical dialogue, representing a new frontier in simulation-based training and protocol communication.
2. Convolutional and generative deep learning architectures, enabling quantitative analysis, image synthesis, and organ-level segmentation, with increasing applications in body composition analysis and tracer distribution modelling.

3. Predictive AI models, applied to physiological and pharmacokinetic variables to derive personalized acquisition parameters, including contrast dosing strategies.

This technological convergence aligns with several macro trends in medicine: personalization of care, reduction of unnecessary variations, sustainability of resource use, and integration of computational intelligence into clinical workflows. Across these dimensions, radiographers and radiologists are transitioning from operators of imaging technology to interpreters of patient-specific imaging physiology supported by computational models.

Within this context, the PhD project presented in this thesis is situated at the intersection between simulation, quantitative imaging, and patient-specific protocol personalization, drawing on preparatory work conducted in multiple domains—communication training, body composition analysis, molecular imaging synthesis, and finally contrast dose optimization in CT.

1.2 Rationale of the PhD: Building a Coherent Multistep Research Program

The rationale underpinning this doctoral project stems from a central observation: modern radiology generates far more information than is currently used for protocol personalization and resource optimization. Despite the availability of scout images, patient-specific attenuation, body composition, and organ morphology, CT contrast administration remains largely anchored to simple heuristics—often weight-based dosing or fixed volumes that neglect physiological variability.

Addressing this gap required a multistep research strategy rather than a single methodological block. For this reason, the PhD was deliberately structured around

progressive, interconnected phase; each designed to build a specific competence or dataset needed for the final translational objective: AI-driven personalization of iodinated contrast dose in CT.

The structure of this progression is intentional:

- First, the project explored the role of LLMs in radiography education, focusing on MRI claustrophobia management. This work provided methodological expertise in prompt design, conversational analysis, and simulation modelling, while also addressing a concrete pedagogical need (Chapter 2).
- Second, the research moved to manual segmentation of CT body composition during an international fellowship in Switzerland. This phase strengthened image-handling skills, ROI standardization, and tissue characterization, establishing the foundational expertise in quantitative imaging necessary for developing physiologically grounded AI models (Chapter 3).
- Third, the thesis introduced a complex AI experiment in molecular imaging, synthesizing FDG-PET from early-phase amyloid PET using multiple architectures (U-Net, GAN, Transformer, Diffusion). This work served as a methodological training ground for multi-architecture experimentation, image-to-image learning, registration pipelines, and model evaluation (Chapter 4).
- Finally, drawing from the competencies accumulated in the previous phases, the project converged on its central objective: developing an AI system capable of predicting and personalizing iodinated contrast dose in CT from the scout image, thereby addressing a concrete clinical need with direct operational impact (Chapter 5).

The coherence of the PhD thus lies not only in the scientific topics explored but in the intentional progression of technical complexity, moving from communication modelling → segmentation → image synthesis → physiological prediction.

Each step has built the technical, methodological, and conceptual scaffolding required to tackle the core challenge of the doctoral project.

1.3 Research Pillars: Competence Building Across Domains

The core objective of this doctoral project and the development of an AI-driven framework for personalized iodinated contrast dosing in CT, did not emerge in isolation. It was the culmination of a structured, deliberate progression of preparatory studies, each designed to build a specific methodological competence and dataset essential for the final translational task. These preparatory activities formed three research pillars, progressively increasing in technical complexity and physiological depth, while remaining tightly aligned with the overarching theme of integrating AI into diagnostic imaging practice.

The first pillar focused on Large Language Models (LLMs) and their capacity to simulate real-world communication scenarios within radiography. The study on managing claustrophobia in MRI through ChatGPT-based simulations served a dual purpose. Methodologically, it provided firsthand experience in prompt engineering, conversational control, and multi-turn interaction analysis—skills that would later prove essential for designing structured AI input–output systems and for handling the interpretability of model responses. Conceptually, it explored how AI can support radiographers not only in technical tasks but also in communication-intensive workflows, highlighting the growing role of AI as an educational and operational companion. This work laid the foundation for understanding how AI systems interact with human users and how structured prompts can shape model behavior—an insight directly transferred to subsequent imaging-based modelling.

The second pillar emerged through the CT body composition segmentation project developed during an international fellowship at the Ente Ospedaliero Cantonale in

Lugano. This phase marked the transition from linguistic AI to quantitative imaging AI. By manually segmenting skeletal muscle, visceral fat, and subcutaneous compartments at the L1 and L3 vertebral levels using SliceOmatic, the research built strong expertise in DICOM handling, ROI standardization, HU-based tissue characterization, and reproducible extraction of morphological parameters. This work was essential for acquiring the technical rigour required to manage large imaging datasets, harmonize cross-center CT acquisitions, and generate high-quality ground truth labels. More importantly, it provided a first-hand understanding of how anatomical variability influences imaging metrics, a fundamental principle later applied to contrast enhancement modelling.

The third preparatory pillar involved deep learning–based image synthesis in PET, specifically the generation of synthetic FDG-PET maps from early-phase amyloid PET. This project introduced a new level of computational complexity, involving the comparison of multiple architectures: U-Net, GANs, Transformers, and Diffusion models and the integration of co-registration pipelines, voxel-wise normalization, and multi-metric model evaluation. Through this work, the PhD research gained exposure to probabilistic generative models, learned the principles of image-to-image translation, and mastered the management of multi-modal inputs in hybrid imaging. At a conceptual level, the project strengthened the understanding of how physiological relationships (in this case, between perfusion and glucose metabolism) can be learned and modeled by neural networks. This insight proved directly transferable to the core CT application, where the goal was to learn the relationship between morphology, dose, and enhancement.

Together, these three pillars formed a cohesive methodological progression. The LLM project refined interaction modelling and prompt structuring; the segmentation fellowship cultivated precision in quantitative imaging; and the PET synthesis study

strengthened competence in image-based deep learning and physiologically informed modelling. Each layer built upon the previous one, converging toward the development of a robust computational framework capable of transforming the routinely acquired CT scout into a predictive tool for contrast dosing. In this sense, the preparatory work did not simply precede the main study and—it actively shaped its feasibility, design, and scientific direction.

1.4 The Target Clinical Need: Optimising Iodinated Contrast Use in CT

Contrast-enhanced CT remains one of the most widely performed diagnostic procedures worldwide, used routinely for oncologic staging, vascular evaluation, trauma imaging, and acute abdominal assessment. Its diagnostic power relies heavily on the quality of contrast enhancement achieved in key vascular and parenchymal structures, which determines lesion conspicuity, organ perfusion assessment, and reliability of quantitative measurements. Despite this central role, contrast administration remains one of the least personalized components of contemporary CT imaging.

In current clinical practice, iodinated contrast media are predominantly administered according to fixed or semi-fixed protocols. These protocols typically scale dose to total body weight or utilize standardized volumes per anatomical district. While operationally simple, such approaches ignore crucial dimensions of inter-individual variability—body habitus, intrathoracic diameter, adipose-to-muscle ratio, water-equivalent diameter, cardiac output, and baseline tissue attenuation. Each of these factors directly affects intravascular iodine concentration, bolus dispersion, and ultimately the parenchymal enhancement that radiologists depend upon.

The consequences of non-personalized dosing manifest at both ends of the enhancement spectrum. Under-enhancement may reduce diagnostic confidence,

obscure hypoattenuating lesions, compromise arterial phase quality, or require repeat imaging. Over-enhancement, conversely, exposes patients to unnecessary iodine load, increasing potential nephrotoxicity in vulnerable populations while inflating institutional contrast consumption and associated costs. As emerging evidence suggests, even modest adjustments in contrast dose can significantly influence enhancement uniformity, underscoring the need for individualized approaches.

This gap between available patient information and actual protocol personalization is particularly striking when considering the CT scout (topogram). The scout image—acquired before every examination at negligible dose—encodes a wealth of patient-specific morphological data: thoracic and abdominal diameters, attenuation distribution, body cross-sectional area, and geometric surrogates of cardiac output and iodinated volume of distribution. Although this information is routinely available, it is rarely exploited to inform dosing strategies, leaving a large portion of the personalized imaging potential untapped.

The drive toward personalization is further reinforced by broader healthcare trends. Regulatory bodies and professional societies increasingly emphasize precision imaging, dose optimization, and resource stewardship. The rising cost of iodinated contrast media—combined with supply chain vulnerabilities and environmental concerns related to iodine persistence in wastewater—magnifies the importance of accurate, patient-specific contrast use. Simultaneously, the growth of AI in radiology provides new computational tools capable of extracting meaningful physiological patterns from scout images and linking them to expected enhancement outcomes.

In this context, a model capable of predicting optimal contrast dose directly from the scout image addresses a pressing clinical and operational need. This project seeks to operationalize that vision by transforming the scout into a real-time dosing predictor, reduces variability in image quality, enhances patient safety, and

introduces a data-driven approach to an aspect of CT imaging that has remained largely empirical. For these reasons, the personalization of iodinated contrast administration represents not only a scientific challenge but a clinically relevant and strategically important frontier in diagnostic radiology.

1.5 Research Objectives

Primary Objective

The principal objective of this research is the design, implementation, and validation of an artificial intelligence system capable of personalizing iodinated contrast media administration in computed tomography based on individual patient characteristics. This system seeks to overcome the limitations of fixed or weight-based protocols, which often fail to account for morphological and physiological variability across patients, resulting in inconsistent image quality and unnecessary contrast usage. By exploiting the routinely acquired CT scout as a morphometric and attenuation proxy, the aim is to predict the expected enhancement in vascular and parenchymal compartments and identify the optimal contrast dose needed to reach diagnostic targets. This model is structured as a forward predictive framework that maps both anatomical (image-based) and technical (metadata-based) input into estimated enhancement values. It then serves as the computational foundation for a subsequent inversion mechanism that enables patient-specific dose recommendation.

Secondary Objectives

To support the development and clinical relevance of the primary model, a number of preparatory and complementary objectives are pursued. First, a retrospective enhancement analysis on a large dataset of contrast-enhanced CT studies is performed, in order to quantify the diagnostic dispersion produced by current dosing schemes and establish reference deviation thresholds for hepatic and renal enhancement. These findings are critical to demonstrate the unmet need for personalized strategies and define the baseline against which AI-driven solutions can be evaluated.

A second goal is to systematically compare the AI-generated dose recommendations with conventional approaches, such as those based on total or lean body weight. This comparative evaluation is essential to determine whether the proposed system offers greater accuracy, enhancement uniformity, or reduction in iodine load across diverse body types.

In parallel, a series of preparatory studies were conducted to reinforce the methodological infrastructure of the project. These include the validation of large language models for simulating clinical radiological dialogues and decision-making workflows, the development of a morphometric ground truth based on manual segmentation of the L1 vertebral level, and the exploration of generative AI models in cross-modal image synthesis using PET imaging as a testbed. While not directly related to CT contrast, these efforts contribute essential components to the broader AI ecosystem supporting radiology decision-making.

Another key objective concerns the practical deployment of the predictive system within real-world clinical workflows. The framework is designed to interface seamlessly with radiological infrastructure, allowing for automatic scout image processing, feature extraction, and dose recommendation in real time without

disrupting existing protocols or burdening personnel. Special emphasis is placed on the integration with RIS systems, CT scanner consoles, and contrast injectors.

Lastly, the environmental and operational implications of personalized contrast dosing are examined. Through internal audit and analysis of contrast usage and waste patterns, the thesis quantifies the impact of standard dosing strategies in terms of iodine wastage, plastic and glass waste generation, and resource sustainability. These findings are used to argue that contrast personalization is not only a technical advancement but also a contribution to more sustainable and responsible radiological practice.

1.6 Structure of the Thesis

The structure of this thesis reflects the sequential and interconnected development of the research journey, progressing from theoretical foundations to methodological consolidation and ultimately to the formulation of a clinically oriented artificial intelligence framework for personalized contrast administration in CT imaging. Each chapter is designed to build on the conceptual and technical competencies introduced in the preceding sections, creating a coherent narrative of scientific maturation.

The thesis begins by outlining the clinical motivations and technological context that necessitate more refined and patient-specific approaches to contrast dosing, situating the research within current challenges in diagnostic imaging, workflow variability, and precision medicine. This introductory chapter establishes the rationale, objectives, and overarching vision that guide the entire project.

The second chapter examines the role of large language models in radiography education, focusing on their capacity to simulate complex patient-provider interactions. This work represents the pedagogical foundation of the doctoral path,

enabling the refinement of methodological rigor in prompt design, experimental control, and communication analysis. Although distinct from the core CT contrast study, this chapter provides essential competencies in simulation-based research and model evaluation that are later transferred to clinically oriented tasks.

The third chapter details the creation of a multicentre CT body composition dataset through manual segmentation at L1 and L3 vertebral levels. This phase establishes the robust morphometric ground truth required for the subsequent predictive modelling. This theoretical consolidation frames the methodological decisions adopted in the subsequent experimental chapters and clarifies the position of the present research within the technological landscape of modern radiology.

The fourth chapter presents the development and comparative assessment of multiple AI architectures for generating synthetic FDG-PET images from early-phase amyloid PET. This work constitutes the first major experimental application of deep learning within the doctoral trajectory and provides valuable expertise in dataset harmonization, model training, performance evaluation, and cross-centre generalization. The insights derived from this chapter inform the technical strategy used in the later CT contrast modelling, particularly regarding architecture selection, evaluation metrics, and optimization strategies.

The fifth chapter represents the core contribution of the thesis, describing the design and validation of an AI-driven framework for personalized iodinated contrast dosing in CT imaging. The chapter details the construction of the CT dataset, the extraction of morphological features from scout images, the development of a forward model linking morphology, dose, and enhancement, and the subsequent inversion process required to generate individualized dosing recommendations. This chapter also presents the results of retrospective analyses, explores integration within clinical

workflows, and discusses the expected clinical and operational impact of dose personalization.

The thesis concludes with a synthesis of the main findings and their implications for radiology practice, followed by an outline of future research avenues. These include the need for multicentre validation, prospective implementation, uncertainty modelling, and interoperability with injector systems and clinical information platforms. Together, these considerations position the thesis within a broader framework of ongoing innovation in precision imaging and AI-driven decision support.

1.7 Project Timeline

The overall development of this doctoral project followed a structured and sequential trajectory, articulated into four major research streams that progressively consolidated the methodological, computational, and clinical foundations of the final core study on CT contrast personalization. Although each phase of the research possesses its own scientific identity, the timeline reflects a coherent progression in which the acquisition of skills and the refinement of analytical tools unfolded in a cumulative manner. Early efforts concentrated on strengthening communication modelling and simulation frameworks in radiography education, which introduced the methodological rigor needed for controlled scenario design and qualitative–quantitative data integration.

Subsequently, the fellowship in quantitative CT imaging at the Ente Ospedaliero Cantonale provided the opportunity to refine segmentation protocols, standardize ROI extraction, and develop a robust workflow for body composition analysis. This stage represented a crucial technical bridge, enabling the transition from manually curated imaging datasets to automated pipelines with reproducible metrics.

The intermediate phase was devoted to experimentation with generative AI architectures for PET-to-PET synthesis, which served as a methodological training ground for high-dimensional medical image translation. The comparison between U-Net, GAN, Transformer, and Diffusion models provided essential insights into data preprocessing, loss functions, model convergence, and evaluation strategies—knowledge directly transferable to the predictive modelling required in the CT project. The final period of the timeline is dedicated to the core objective of the dissertation: the development of an AI-driven system for personalized iodinated contrast administration based solely on scout images. This stage integrates the methodological heritage of the previous phases of the research, combining morphometric feature extraction, supervised regression architectures, optimisation layers, and validation protocols. The timeline therefore illustrates a gradual shift from simulation and experimental modelling to clinically focused algorithmic design, culminating in a decision-support tool intended for real-world implementation within CT imaging workflows.

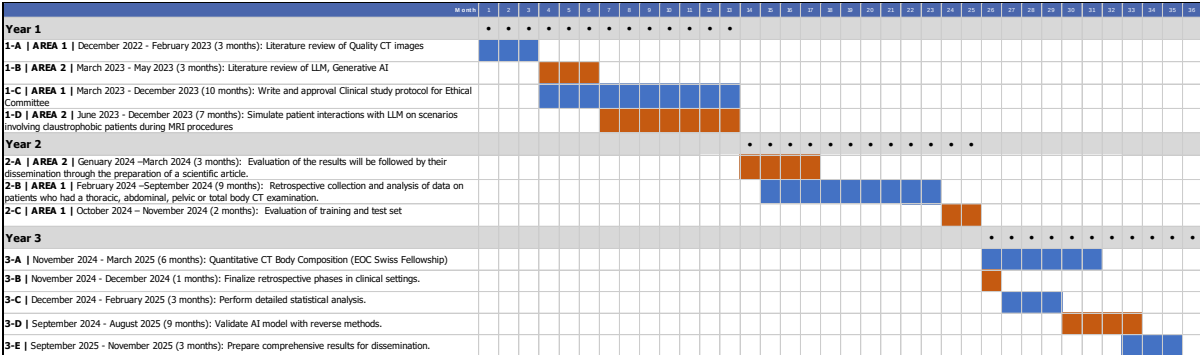


Figure 1 Three-year Gantt chart summarising the temporal organisation of the research activities, including literature review.

Chapter 2

Large Language Models and AI-assisted Simulation in Radiography Education

2.1 Introduction

Communication between radiographers and patients is a performance variable, not a courtesy. In MRI this is even sharper, because the success of the examination depends on the patient remaining still, tolerating noise and enclosed space, and trusting the operator. Claustrophobia has been described as a systematic cause of delayed, incomplete or cancelled MRI examinations, with anxiety prevalence between 29% and 56% and interruption rates up to 5.29% (Forshaw et al., 2018; Hudson, Heales and Meertens, 2022). Such rates are not marginal: they show that emotional dysregulation of the patient can damage workflow, reduce scanner utilisation and expose the service to re-calls (Munn et al., 2015).

Radiographers are the first line of mitigation. Their ability to explain, normalise fear, give the patient a sense of control and offer coping strategies often decides whether the scan is completed or not (Ajam et al., 2020; Rockall et al., 2022). Yet training in communicative competence is still uneven across radiography programmes: it is often delivered as lectures, brief role-plays or unsupervised clinical exposure, with little traceability of what students actually say to patients (Wouda and van de Wiel, 2012; Bachmann, Pettit and Rosenbaum, 2022). This creates an obvious gap between the complexity of real MRI interactions—older patients, previous failed MRI, comorbid anxiety—and the simplicity of didactic activities.

The appearance of Large Language Models (LLMs) changes this scenario. Models such as ChatGPT, built on the Transformer architecture introduced by Vaswani et al.

(2017), are able to generate contextually coherent dialogue, respect a communicative role and sustain multi-turn interaction. Because they can be instructed to “behave like a claustrophobic MRI patient”, they become a low-cost, always-available simulator of difficult patients. Early health-education reports have already shown that LLMs can generate clinical vignettes, explain radiology reports at lay level, simulate conversations and help learners rehearse communication tasks (Sallam, 2023; Currie et al., 2023; Jeblick et al., 2022; Scherr et al., 2023; Webb, 2023). This chapter analyses how such LLM-based simulations can be embedded in radiography education to train MRI-related communication, and it represents the first, pedagogical strand of the doctoral work that will later move to quantitative and CT-based AI.

2.2 Theoretical Background on LLMs

LLMs are trained on very large text corpora to learn statistical regularities of language and to predict the next token in context. The key innovation is the self-attention mechanism, which lets the model weigh relationships between all tokens in a sequence, keeping long-range coherence—crucial in dialogue, where one utterance depends on what was said several turns earlier (Vaswani et al., 2017). ChatGPT is an LLM fine-tuned with reinforcement learning from human feedback so that responses are not only fluent but socially acceptable and aligned to user intent (OpenAI, 2023).

In medical and health-professions education, such models have been tested for: automatic creation of patient cases (Benoit, 2023), simplification of technical or radiological language for patients (Jeblick et al., 2022), assistance in high-stakes knowledge tests (Gilson et al., 2023; Bhayana, Krishna and Bleakney, 2023), and, closer to our work, for interactive simulations of clinical communication (Scherr et al., 2023; Webb, 2023). Parallel work in educational science had already argued for AI

as an amplifier of simulation-based learning, provided that scenarios are structured and assessment is traceable (Holmes, Bialik and Fadel, 2019). In radiology, first explorations showed that ChatGPT can support image-related decision making and educational tasks, although with variability across versions and prompts (Rao et al., 2023; Currie et al., 2023).

This technological background matches a specific radiography need. Communication in MRI is not only informative; it modulates behaviour. The patient must accept head-first entry, tolerate coil placement, coordinate breathing, stay still. Communication is thus procedural and affective at the same time. Scripted e-learning cannot reproduce such variability; LLMs, by contrast, can generate unpredictable but plausible replies, exposing the learner to escalating fear, refusal, bargaining or somatic complaints. Because all exchanges are textual, they can also be coded and studied—an advantage for research and for competency-based education.

2.3 Rationale and Objectives

The rationale of this study is dual. Clinically, claustrophobia in MRI is still common and pushes departments to use extra time slots, sedation or repeat scans, with additional cost and lower patient satisfaction (Forshaw et al., 2018; Hudson, Heales and Meertens, 2022; Lawal, Regelous and Omiyi, 2024). Educationally, radiography schools need scalable, evidence-based ways to teach “what to say” in these situations, without relying every time on actors or on chance exposure in clinical placements (Bachmann, Pettit and Rosenbaum, 2022).

The main objective was to test whether ChatGPT could be used as a realistic simulator of a claustrophobic MRI patient and whether it was sensitive enough to show differences between radiographers with different levels of experience. Associated aims were: (i) to test the conversational stability of two LLM versions

(GPT-3.5 vs GPT-4); (ii) to describe which communication strategies radiographers actually use in front of a resistant virtual patient; and (iii) to verify whether the whole process is feasible for integration in formal curricula.

2.4 Materials and Methods

A pilot study was run between May and July 2023 with six radiographers, two per experience bracket (expert, intermediate, beginner). Each radiographer held ten simulations: five with GPT-3.5 and five with GPT-4. To avoid over-engineering the scenario and to preserve spontaneity, prompts were constructed in a **zero-shot** fashion: the model was given only one initial instruction containing demographic data, exam type, entry mode and severity of claustrophobia, plus an explicit constraint to “answer only as the patient”. No example dialogues were supplied. Zero-shot prompting was chosen because it better reproduces the heterogeneity of real MRI patients and because it makes deviations or role drift very visible, which is useful for educational research on LLMs (He et al., 2023).

	Age (n. of years)	Gender	Type of MRI exam	Patient's enter	Severity of claustrophobia	Failed previuos MRI exams
Virtual Patient 1 (VP1)	24	Male	Knee	Feet first	+	No
Virtual Patient 2 (VP2)	45	Female	Abdomen	Head first	++	No
Virtual Patient 3 (VP3)	45	Male	Brain	Head first	+++	No
Virtual Patient 4 (VP4)	55	Female	Abdomen	Head first	++++	Yes
Virtual Patient 5 (VP5)	60	Female	Brain	Head first	+++++	Yes

Table 1: Characteristics of "virtual patients"

Five virtual-patient profiles were generated, escalating from mild unease to severe, previously documented panic with interrupted MRI, in line with what is described in

radiography literature on MRI anxiety (Hudson, Heales and Meertens, 2022; Al-Shemmari et al., 2022). The most severe profile was a 60-year-old woman scheduled for brain MRI, head first, who had interrupted a previous examination and expressed fear of not being able to breathe. Radiographers were instructed to persuade the patient to complete the scan using their usual communication repertoire. All conversations were exported and anonymised.

During the simulation, if the model started offering medical advice or switched role (“as a radiographer you should...”), participants were allowed to type a very short corrective message (“Please answer as the patient”) but were not allowed to reshape the scenario. This preserved ecological validity, a point already underlined in other AI-communication studies (Ryan et al., 2019; Tanana et al., 2019). Transcripts were then analysed quantitatively for success/failure and for frequency of strategies, and qualitatively through thematic analysis.

Communication strategies were coded using the framework proposed by Al-Shemmari et al. (2022) for MRI claustrophobia management, which distinguishes supportive instructions, distraction/music, breathing/relaxation techniques, pharmacological support and technological aids (e.g. VR goggles). This gave us a radiography-specific grid instead of generic counselling codes.

It is important to clarify that all interactions analysed in this study were generated within a simulated environment using large language models and do not represent real clinical encounters or patient data. The purpose of the simulations was to evaluate communicative strategies and conversational dynamics under controlled and reproducible conditions, rather than to draw conclusions on clinical outcomes or patient behaviour in real-world MRI settings.

Full prompts are published in GitHub repository for data sharing

[\[https://doi.org/10.5281/zenodo.10441155\]](https://doi.org/10.5281/zenodo.10441155)

2.5 Results

Sixty simulations were completed. Overall, 58/60 (96.7%) ended with the virtual patient accepting or maintaining the exam. GPT-4 was fully stable: it never abandoned the patient role and always replied within the emotional frame given at the start. GPT-3.5, by contrast, showed conversational instability in about 40% of encounters, usually by offering radiographer-like reassurance or by trying to terminate the conversation early; this confirms what other authors reported on model-version variability in educational use (Currie et al., 2023; Sallam, 2023).

Supportive, patient-centred communication was the dominant strategy across all participants, accounting for over two-thirds of coded utterances, mirroring what has been described in MRI anxiety interventions (Ajam et al., 2020; Lawal, Regelous and Omiyi, 2024). Experts tended to anticipate fears (“you will always hear my voice”, “you can stop at any time”), to offer shared control and to guide breathing. Intermediates relied more often on distraction or music suggestions, a pattern consistent with the use of sensory modulation in MRI (Földes et al., 2017). Beginners were the group most likely to mention mild sedation or to suggest that “many patients take something”, showing a lower confidence in purely communicative de-escalation, as previously observed in radiographer confidence studies (Al-Shemmari et al.,

2022).

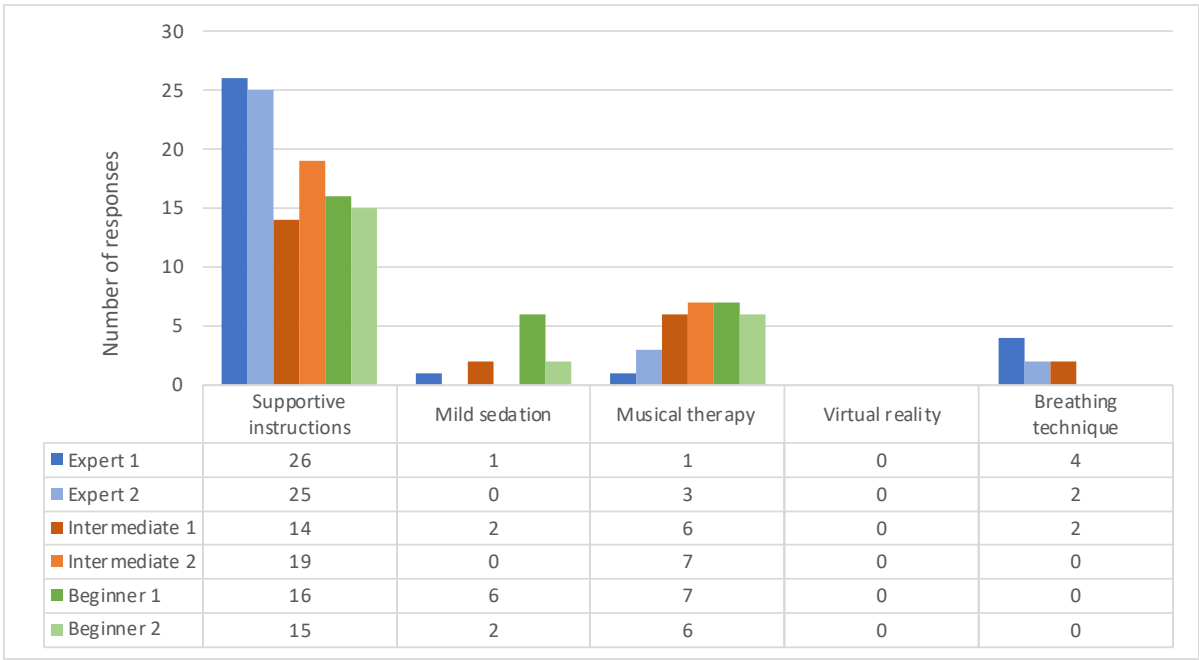


Figure 2: Number of responses needed to persuade the patient per participant.

2.6 Qualitative Analysis of Dialogue Dynamics

The qualitative reading of transcripts made the differences in communicative maturity very clear. Expert radiographers used layered reassurance: they named the fear, normalised it with reference to other patients, offered a concrete coping instruction, and reminded the patient of the possibility to interrupt. This aligned closely with the REDE relational model of healthcare communication, which frames the relationship itself as therapeutic (Windover et al., 2013). Intermediates often opened well but failed to maintain control when the virtual patient escalated (“I cannot breathe”, “I want to get out”); they diverted to music or distraction, which worked only in moderate cases. Beginners’ dialogues were shorter, less structured and sometimes shifted responsibility to pharmacology or to the referring doctor. Being able to see these patterns in text, line by line, is precisely what makes LLM-based simulation pedagogically attractive.

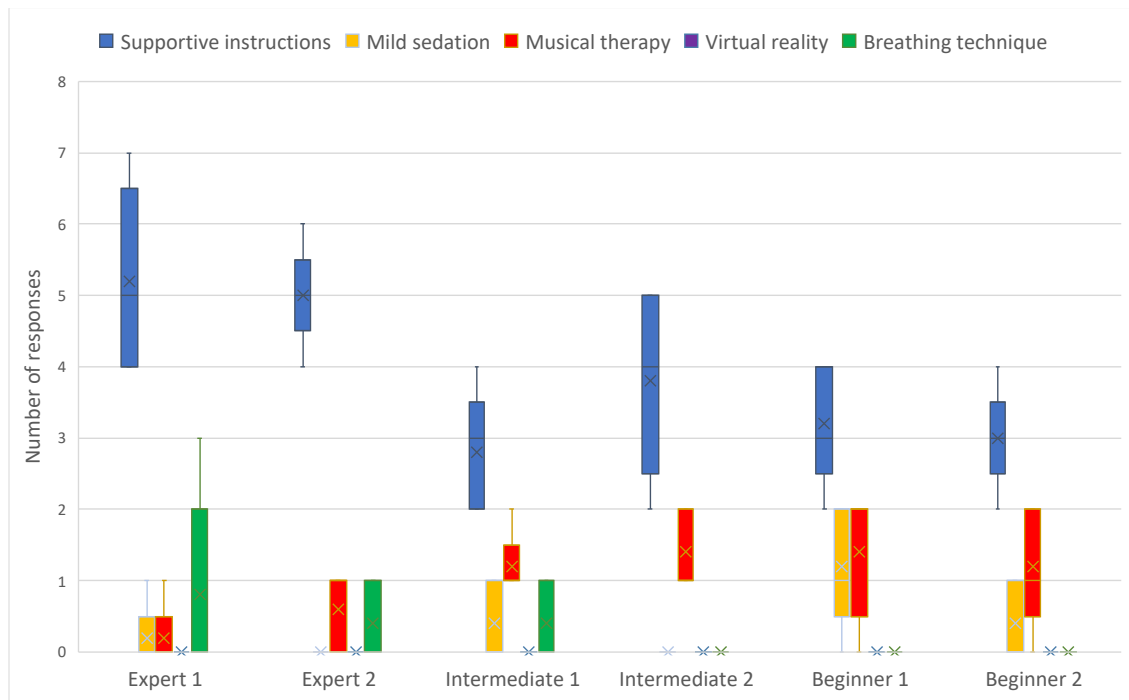


Figure 3: The use of methods for supporting patients with claustrophobia at different levels of virtual patients. [Supportive instructions (blue); mild sedation (orange); Musical therapy (red); Virtual reality (purple); Breathing technique (green)]

2.7 Educational and Organisational Implications

From an educational perspective, the main gain is scalability. Once the prompts are designed, tens of students can run simulations in parallel, at home or in the skills lab, without actors and without occupying the MRI suite. This lowers cost and raises frequency of practice—two chronic problems of communication training in imaging (Bachmann, Pettit and Rosenbaum, 2022). Because every interaction is textual, instructors can score strategy use, error recovery and empathy markers in a standardised way, which meets current demands for objective assessment in radiography (Portelli et al., 2018).

From an organisational angle, such simulations can be mapped onto Italian TSRM curricula in at least three moments: (i) early bachelor years, to introduce basic MRI communication and the language of reassurance; (ii) pre-clinical or OSCE preparation, where students can rehearse a high-anxiety scenario before meeting a

real patient; and (iii) continuing education for practising radiographers, especially in centres where MRI is frequently interrupted due to anxiety. Similar integrations of digital or VR-based simulation in health-profession education have already been shown to improve preparedness and patient experience, which supports adoption (Kyaw et al., 2019; van der Kruk et al., 2022).

2.8 Contributions and Future Work

This study delivers four concrete outputs. First, it shows that GPT-4 can sustain stable, role-consistent MRI-patient simulations. Second, it demonstrates that such simulations make radiographers' communicative strategies visible and codable. Third, it offers a reproducible prompting protocol—zero-shot, role-constrained, severity-graded—that other education centres can adopt. Fourth, it opens the way to automated or semi-automated scoring of communication, as already experimented in broader medical-communication research (Ryan et al., 2019).

Future work will extend the sample to students and to multiple institutions, will incorporate voice and emotion recognition to increase immersion, and will develop rubrics that couple frequency of strategies with quality of delivery. A further line, closer to the clinical chapters of this thesis, will test whether the same LLM-based design can be ported to other high-anxiety imaging settings, such as paediatric CT or breast MRI.

While at first glance the topics addressed in this chapter may seem distant from the technical modeling of contrast enhancement, patient communication and behavioral compliance remain a cornerstone of modern radiology. The effectiveness of computational models, particularly those for dose optimization and image quality enhancement, hinges critically on patient cooperation during image acquisition. Thus, the ability to simulate clinical conversations using Large Language Models not only

serves as a training tool but directly supports the success of precision imaging workflows explored in the following chapters.

Modern radiology increasingly relies on the convergence of technical precision and effective patient cooperation. Advanced computational models, including those developed for quantitative imaging and dose optimisation, presuppose stable acquisition conditions and patient compliance. In this context, communication strategies aimed at reducing anxiety, improving understanding, and promoting cooperation are not ancillary, but foundational to the successful application of precision imaging technologies. The investigation of large language models for patient–radiographer interaction therefore represents a preparatory step toward enabling the reliable acquisition environments required for subsequent AI-driven quantitative analyses and protocol optimisation.

Bridging Perspective

Although the focus of this thesis progressively shifts from communication modelling to quantitative imaging and protocol optimisation, these domains are not conceptually independent. Modern diagnostic radiology requires both patient compliance and technical precision to achieve reliable and reproducible image quality. Without effective communication and patient cooperation, even the most advanced acquisition protocols and computational models fail to deliver clinically usable data. The work presented in Chapter 2 addresses this foundational prerequisite by exploring how artificial intelligence, in the form of large language models, can support radiographers in managing patient anxiety and ensuring examination feasibility, particularly in high-stress environments such as MRI. This dimension of AI-assisted practice is complementary to, rather than separate from, the quantitative approaches developed in the subsequent chapters.

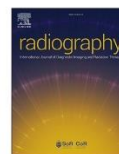
Chapters 3 and 4 then build the technical competencies required for data-driven optimisation in imaging, focusing on quantitative morphometric analysis and multi-architecture deep learning applied to medical images. These methodological foundations converge in Chapter 5, where patient-specific morphology and acquisition parameters are integrated into an AI-driven framework for personalised iodinated contrast administration in CT.

In this perspective, the thesis follows a coherent trajectory: from ensuring the human and procedural conditions necessary for high-quality imaging, to developing computational models capable of exploiting that data for personalised and physiologically informed protocol optimisation.



Contents lists available at ScienceDirect

Radiography

journal homepage: www.elsevier.com/locate/radi

Harnessing ChatGPT dialogues to address claustrophobia in MRI - A radiographers' education perspective

G.R. Bonfitto ^{a, c, *}, A. Roletto ^{b, c}, M. Savardi ^d, S.V. Fasulo ^c, D. Catania ^c, A. Signoroni ^d^a Department of Information Engineering, University of Brescia, Via Branze 38, 25123 Brescia, Italy^b Department of Mechanical and Industrial Engineering, Università degli Studi di Brescia, Via Branze 38, 25123 Brescia, Italy^c IRCCS Ospedale San Raffaele, Via Olgettina 60, 20132 Milano, Italy^d Department of Medical and Surgical Specialties, Radiological Sciences, and Public Health, University of Brescia, Viale Europa 11, 25121, Brescia, Italy

ARTICLE INFO

Article history:

Received 29 December 2023

Received in revised form

19 February 2024

Accepted 20 February 2024

Available online 29 February 2024

Keywords:

ChatGPT

Generative AI

Simulation: radiographer

Claustrophobia

Magnetic resonance imaging

ABSTRACT

Introduction: The healthcare sector invests significantly in communication skills training, but not always with satisfactory results. Recently, generative Large Language Models, have shown promising results in medical education. This study aims to use ChatGPT to simulate radiographer-patient conversations about the critical moment of claustrophobia management during MRI, exploring how Artificial Intelligence can improve radiographers' communication skills.

Methods: This study exploits specifically designed prompts on ChatGPT-3.5 and ChatGPT-4 to generate simulated conversations between virtual claustrophobic patients and six radiographers with varying levels of work experience focusing on their differences in model size and language generation capabilities. Success rates and responses were analysed. The methods of radiographers in convincing virtual patients to undergo MRI despite claustrophobia were also evaluated.

Results: A total of 60 simulations were conducted, achieving a success rate of 96.7% (58/60). ChatGPT-3.5 exhibited errors in 40% (12/30) of the simulations, while ChatGPT-4 showed no errors.

In terms of radiographers' communication during the simulations, out of 164 responses, 70.2% (115/164) were categorized as "Supportive Instructions," followed by "Music Therapy" at 18.3% (30/164). Experts mainly used "Supportive Instructions" (82.2%, 51/62) and "Breathing Techniques" (9.7%, 6/62). Intermediate participants favoured "Music Therapy" (26%, 13/50), while Beginner participants frequently utilized "Mild Sedation" (15.4%, 8/52).

Conclusion: The simulation of clinical scenarios via ChatGPT proves valuable in assessing and testing radiographers' communication skills, especially in managing claustrophobic patients during MRI. This pilot study highlights the potential of ChatGPT in preclinical training, recognizing different training needs at different levels of professional experience.

Implications for practice: This study is relevant in radiography practice, where AI is increasingly widespread, as it explores a new way to improve the training of radiographers.

© 2024 The Author(s). Published by Elsevier Ltd on behalf of The College of Radiographers. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Abbreviations: AI, Artificial Intelligence; GPT, Generative Pre-trained Transformer; LLM, Large Language Model; VP, Virtual Patient; MRI, Magnetic Resonance Imaging.

* Corresponding author. Department of Information Engineering, Università degli Studi di Brescia, Via Branze 38, 25123 Brescia, Italy.

E-mail addresses: giuseppe.bonfitto@unibs.it (G.R. Bonfitto), andrea.roletto@unibs.it (A. Roletto), mattia.savardi@unibs.it (M. Savardi), fasulo.simone@hsr.it (S.V. Fasulo), catania.diego@hsr.it (D. Catania), alberto.signoroni@unibs.it (A. Signoroni).

<https://doi.org/10.1016/j.radi.2024.02.015>

1078-8174/© 2024 The Author(s). Published by Elsevier Ltd on behalf of The College of Radiographers. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

The consistent provision of effective communication by healthcare professionals is a crucial component in the success of patient care processes. Educational institutions across various health disciplines are increasingly allocating resources to enhance communication skills training. Despite these efforts, tangible improvements are often not evident, and the impact of post-graduate clinical experience in this area appears to be limited.^{1–3}

In the literature, studies have suggested and supported the possibility of improving the teaching and training of

communication skills of healthcare professionals and generic workers using artificial intelligence (AI) software,⁴ such as virtual assistants or chatbots.⁵

One of the most well-known and popular chatbot is the free-to-use Chat Generative Pre-trained Transformer (ChatGPT). ChatGPT developed by OpenAI, is a Large Language Model (LLM), an advanced AI algorithm trained to process and generate human-like text. It learned from extensive language data, enabling it to understand and respond with contextually relevant information, optimized for dialogue using Reinforcement Learning with Human Feedback (RLHF),^{6,7} a method that uses human demonstrations and preference comparisons to guide the model toward desired behaviour. The two most recent versions of ChatGPT are built on GPT-3.5 and GPT-4.⁸

Since its release in November 2022, the benefit of using ChatGPT in different scientific contexts has been tested. Sallam M. (2023) published a systematic review of the literature on the use of ChatGPT in healthcare education, research, and practice, identifying benefits in healthcare academic/scientific research, clinical practice, and education.⁹

ChatGPT presented good skills in writing abstracts,^{10,11} and scientific articles,^{12,13} answering questions and supporting decisions during medical activities,^{14–16} making radiology reports more understandable to patients,¹⁷ generating educational contents^{11,18} and demonstrating ability to pass university exams at the level of medical or radiography student.^{19–21}

The idea of combining the possibility of improving the communication skills of healthcare professionals through the use of chatbots and the benefits that an LLM software can bring to healthcare has already been explored by some authors^{22,23} who proposed using the ChatGPT language model to create a role-playing game with the goal of simulating a doctor-patient conversation in order to provide educational feedback to the user in real-time. Currently, no study has evaluated the same possibility in a radiological examination context.

Indeed, patient communication plays an important role in medical imaging, not only in communicating the results of a radiological investigation, but also in involving patients and relatives in decisions, in providing operational instructions during examinations, and above all alleviating pain, fear, and anxiety.²⁴ About the fear and anxiety episodes of medical imaging patients, MRI (Magnetic Resonance Imaging) represents the diagnostic investigation where an increase in patient anxiety is reported in 29–56% of cases, with claustrophobia being the most prominent cause, with a 0.46–5.29% incidence of premature exam termination rate.^{25–27} Therefore, communication with the patient becomes important to reduce the anxiety caused by claustrophobia and to improve the patient's experience during the MRI.^{28,29}

The aim of this study is to explore the possibility of constructing scenarios related to an MRI examination in which a radiographer can role-play with a simulated patient suffering from claustrophobia using ChatGPT. The final purpose is to test whether it is possible to create an effective educational approach that includes a combination of formal didactics, role-playing, and encounters with simulated patients.

Our pilot study also aims to conduct a comparative analysis between ChatGPT-3.5 and ChatGPT-4, with a focus on assessing their efficacy in simulating clinical dialogues and identifying the improvements from GPT-3.5 to GPT-4.

Methods

This pilot study was conducted over a three-month period, from May to July 2023, using the generative language models ChatGPT-3.5 and ChatGPT-4 developed by OpenAI.³⁰ These deep learning models

are based on a neural network architecture known as Transformer,³¹ which enables them to generate high-quality natural text in response to a series of prompts. These language models have been used here to simulate conversations between radiographers and virtual claustrophobic patients undergoing MRI scans. Access to ChatGPT was obtained through a subscription to the OpenAI website.

Selection of participants

Six radiographers were selected to participate in the study. These participants were chosen to represent a range of experience levels in the field of radiography. Specifically, we selected two experienced radiographers with at least 15 years of experience using MRI, identified as Expert 1 (E1) and Expert 2 (E2), two intermediate radiographers with at least 5 years of experience, identified as Intermediate 1 (I1) and Intermediate 2 (I2), and two novice radiographers with less than 5 years of experience, identified as Beginner 1 (B1) and Beginner 2 (B2). This diversity of expertise allowed us to explore how professional experience might influence the radiographer's ability to interact effectively with an artificial language model in a simulated clinical setting. Gender parity of participants (3 females and 3 males) was also ensured.

As this study was conducted in Italy and registered as quality assurance in training project, it employed a purposive sampling method and was exempted from approval by the Lombardy 6 Territorial Ethics Committee as it did not aim at "providing new knowledge about health and disease". Furthermore, no patient information was included, and the study did not require informed consent from patients. The radiographers were informed about the study in advance. A confidentiality statement was valid for the data collection.

Virtual patients' characteristics

To ensure that the simulated interactions were as representative as possible of real clinical interactions, we established a set of characteristics for the "virtual patients" based on the characteristics identified in patients most likely to suffer from claustrophobia during an MRI examination.³²

In this specific context, five Virtual Patients were developed, ranging from the simplest "Virtual Patient 1 (VP1)" to the most complex "Virtual Patient 5 (VP5)". The characteristics of these patients were identified considering various factors such as the type of examination to be conducted. In particular, different age groups, genders and types of MRI diagnostic examination were identified and associated with a varying incidence of claustrophobia and, consequently, a varying degree of difficulty in the conversation with the patient undergoing an MRI (Table 1).²⁷

The starting input prompts for this study were designed to represent a range of degrees of claustrophobia and conversational complexity. Five different prompts associated with the five levels of patient conversation difficulty were constructed. For example, the "Virtual Patient 1" level simulates a patient who is unsure whether he or she has claustrophobia and is undergoing an MRI for the first time in a body district that does not involve full entry into the MRI gantry. In contrast, the prompt for the "Virtual Patient 5" level simulates a patient who suffers from severe claustrophobia and has a strong resistance to undergoing an MRI scan.

Full prompts are published in GitHub repository for data sharing [dataset].³³

Starting prompt development

The main method for starting a role-playing game with ChatGPT is to create an effective initial input prompt. To model an initial

Table 1
Characteristics of “virtual patients”.

	Age (no. of years)	Gender	Type of MRI exam	Patient's enter	Severity of claustrophobia	Failed previous MRI exams
Virtual Patient 1 (VP1)	24	Male	Knee	Feet first	+	No
Virtual Patient 2 (VP2)	45	Female	Abdomen	Head first	++	No
Virtual Patient 3 (VP3)	45	Male	Brain	Head first	+++	No
Virtual Patient 4 (VP4)	55	Female	Abdomen	Head first	++++	Yes
Virtual Patient 5 (VP5)	60	Female	Brain	Head first	+++++	Yes

input request, two examples reported in the literature were considered. The first involved an input based on a simulated business negotiation,³⁴ the second involved the simulation of a scenario in a medical setting.²³ An input request had to include information on the type of scenario the author wished to simulate and the role the author wanted ChatGPT to play. The prompts then described the patient's characteristics and the scenario according to the difficulty level and role-playing rules so that ChatGPT could initiate the conversation scenario.

To maintain linguistic consistency and relevance for our participants, all simulations and prompts were conducted in Italian, catering specifically to the native language of our Italian-speaking participants, ensuring that the proposed actions were presented in their mother tongue. Furthermore, for broader accessibility and documentation purposes, the prompts have also been translated into English and made available on GitHub.

Two-minute scenarios were implemented to mirror clinical practice, underscoring the necessity for rapid decision-making and efficient communication, which are essential in real-world clinical settings, rather than serving as a constraint on the scenario itself. Utilizing ChatGPT Playground, an interactive web interface provided by OpenAI³⁵ that allows users to experiment with and fine-tune the responses of the ChatGPT model by adjusting parameters, providing prompts, and viewing the AI's outputs in real-time, the team evaluated GPT's responsiveness, the coherence of its responses, and the adaptability of virtual scenarios. Drawing from the genuine experience of interacting with a patient scheduled for an MRI,³⁶ this dynamic testing environment was essential for refining scripts to reflect real-world clinical situations. Two senior radiographers (G.R.B. and D.C.) assessed the applicability and credibility of the scenarios, as well as the realism of the interactions presented. In the context of this study, the team adopted a zero-shot method.^{37,38} This method refers to an approach for prompting an LLM to generate text without providing any specific instructions or examples. This method was designed to ensure that the LLM's responses are not biased by the prompts and are instead based on its understanding of the task and the information it has been trained on.

There are several advantages to using the zero-shot method for instance, it allows the LLM to generate more creative and original responses, as it is not limited by the predefined rules of the prompts.^{37,38}

Study design flow

The study was conducted in several phases, each of which contributed to a different aspect of our research objective.

- Preparation phase: During this phase, the team recruited participants and designed conversation prompts for “virtual patients”.
- Simulation phase: In this phase, each participant conducted a series of conversations with “virtual patients” using ChatGPT language prompts. Each participant went through a series of clinical scenarios, each of which required handling a patient

with claustrophobia who had to undergo an MRI. Each participant used a new dialogue box for every scenario, thereby guaranteeing that each simulation was conducted in isolation from the others and every simulation, the participants always introduced themselves to the virtual patient.

- Evaluation phase: After each simulation, participants were asked to provide feedback on the conversation with the virtual patient based on their experience. This feedback was useful to determine the effectiveness of ChatGPT as a simulator of clinical conversations.

The study employed quantitative and qualitative analyses, providing mean scores and standard deviations for each dimension, and examined the simulations, success rates, and responses of virtual patients and participants. The latter evaluation focused on the effectiveness of the radiographers in convincing the virtual patient to undergo MRI despite claustrophobia, mapping the variety of techniques utilized in simulations and evaluating their effectiveness based on participant experiences.

This involved recording the number of responses needed to convince the virtual patient and documenting the methods used to support patients with claustrophobia, as outlined in Al-Shemmari et al. (2022) by using “Supportive instructions”, “Mild sedation”, “Musical therapy”, “Virtual reality” and “Breathing technique”.²⁹

Comparison of ChatGPT-3.5 and ChatGPT-4

A secondary objective of our study was to compare the effectiveness in simulating clinical conversations of two different versions of ChatGPT, i.e. the ones based on GPT-3.5 and GPT-4, which differ in terms of model size and language generation capability. The participants experienced all the virtual patients with Chat GPT-3.5 and then performed the same scenarios with Chat GPT-4. To compare these two models, the team submitted the same simulation to participants with both versions.

Results

All simulations were performed with ChatGPT-3.5 and ChatGPT-4 for analysis and comparison of results. Using both versions of ChatGPT, the study comprised 60 simulations.

For each simulation, the software's ability to credibly recreate the clinical scenario was evaluated, including counting the number of errors made by the chatbot (e.g. role change during simulation). The simulations performed had a success rate of 96.7% (58/60 simulations), where success was defined as the simulation concluding with the virtual patient agreeing to proceed with the examination. Conversely, a failure is when the virtual patient refuses to proceed. The issues identified in the two cases of failure were due to overly restrictive conversations by the radiographers, leaving no room for further dialogue. This aspect could be more thoroughly investigated with a larger sample size to allow for a more accurate assessment and to understand whether it can be attributed to ineffective communication approaches or a bias in the LLM.

ChatGPT-3.5 made errors in 40% of the simulations (12/30) and never made errors during the simulations performed with ChatGPT-4 (0%, 0/30). The errors consisted of a change in the role of the chatbot from patient to radiographer. These errors were corrected directly by the participant calling the chatbot back to the correct role.

Regarding the radiographers' ability to convince the virtual patient to perform the MRI investigation despite claustrophobia, as shown in Fig. 1 we considered all the textual messages from the participants in all 60 simulations. The study included a total of 584 interactions between the participants and the virtual patient. Of these, the team identified 164 distinct responses aimed at convincing the claustrophobic "virtual patient" to perform an MRI examination, regardless of the level of difficulty. On average, the scenarios lasted 1 min 45 s, with a variability of plus or minus 10 s. In detail, 70.2% (115/164) of the responses used are in the field of "Supportive instructions", followed by "Musical therapy" (30/164, 18.3%), "Mild sedation" (11/164, 6.7%), "Breathing techniques" (8/164, 4.8%) and "Virtual reality" (0/164, 0%).

Analysing participants' responses to their different experiences, Experts used 62 out of 164 responses (37.8%), mainly "Supportive instructions" (51/62, 82.2%), followed by "Breathing techniques" (6/62, 9.7%), "Musical therapy" (4/62, 6.5%) and "Mild sedation" (1/62, 1.6%).

Intermediate participants used 50 out of 164 responses (30.5%), mainly "Supportive instructions" (33/50, 66%), followed by "Musical therapy" (13/50, 26%), "Mild sedation" and "Breathing techniques" both 4% (2/50).

Instead, Beginners used 52 out of 164 responses (31.7%), mainly "Supportive instructions" (31/52, 59.6%), followed by "Musical therapy" (13/52, 25%) and "Mild sedation" (8/52, 15.4%). As shown in Fig. 2, across all simulations, the usage of techniques at different levels of virtual patients has been analysed. The data analysis confirms the separately observed trends. Specifically, supportive instructions (115/164 70.1%) and musical support (30/164 18.3%) are the most used techniques across all levels of experience and difficulty.

Interestingly, on average, supportive instructions were applied 3.8 times per simulation, while musical therapy was used about once per simulation.

Beginners, on the other hand, show a more balanced use between musical therapy and mild sedation (10.5 ± 2.5), compared to experts and intermediates (0.575 ± 0.045), even though they use mild sedation more frequently compared to the other two groups.

Through a qualitative analysis, among the various techniques analyzed, Supportive Instructions and Musical Therapy have been identified as the most effective. Specifically, Supportive Instructions positively convinced the virtual patient in 31 out of 60 responses (18.90%), while Musical Therapy did so in 22 out of 60 responses (13.41%). It's important to emphasize that in evaluating the effectiveness of methods, it must be considered that combinations of techniques can occur in interactions with virtual patients to optimize the convincing process. Participants provided positive feedback on the simulation experience. They expressed that engaging in a conversation with a "virtual patient" closely mirrored the dynamics of interacting with a real patient.

Additionally, all participants acknowledged the significant training value embedded in the simulation. They recognized its efficacy in honing the skills required to employ similar techniques effectively with actual patients. This underscores the simulation's practical utility as a preparatory tool for real-world scenarios.

Discussion

Preclinical simulation

Our pilot study delved into the development of a simulation involving an MRI examination, wherein radiographers engaged in role-playing scenarios with a "virtual patient" experiencing claustrophobia, facilitated by prompts generated through different versions of ChatGPT.

The use of LLMs and LLM-based chatbots has shown potential in research and teaching methods. AI chatbots can play a crucial role in creating learning assessments and generating exercises, quizzes, and scenarios that can be used in a classroom context to aid in student practice and evaluation.³⁹ Another aspect evaluated in previous studies is the capabilities of LLMs in decision-making and radiology reporting to enhance the stages of clinical practice and reporting.^{40–43} Other innovative training methods use technologies

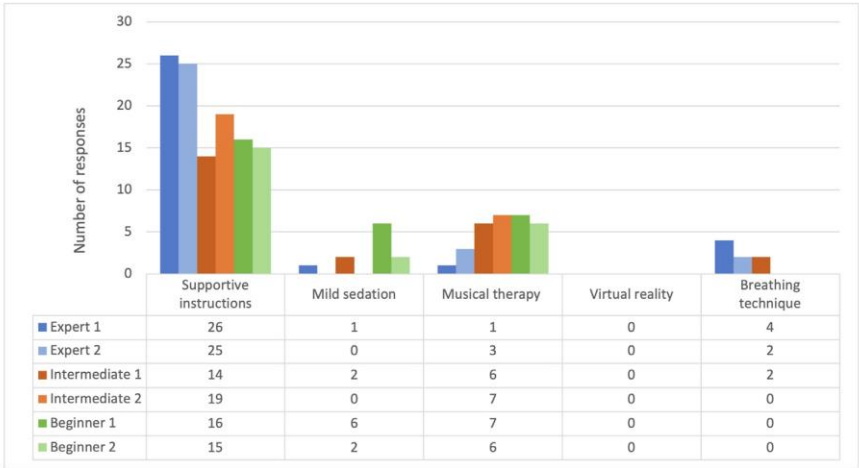


Figure 1. Number of responses needed to persuade the patient per participant.

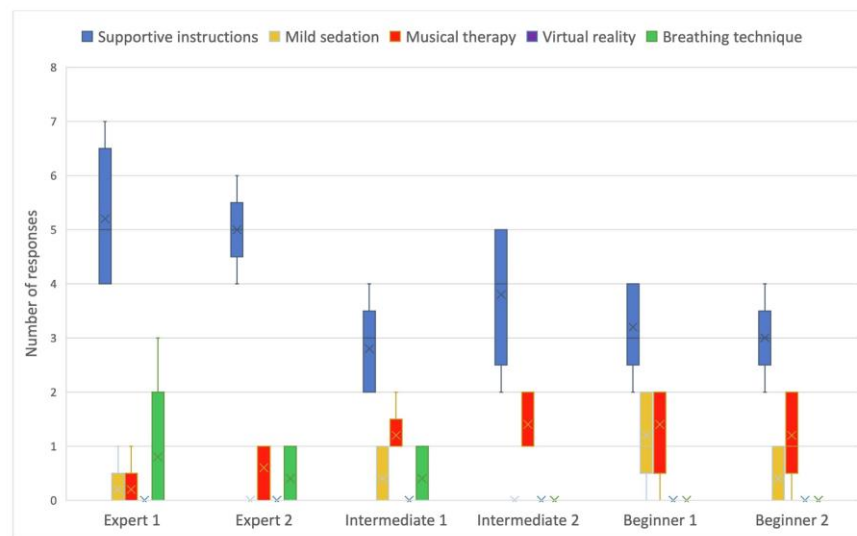


Figure 2. The use of methods for supporting patients with claustrophobia at different levels of virtual patients. [Supportive instructions (blue); mild sedation (orange); Musical therapy (red); Virtual reality (purple); Breathing technique (green)].

such as Virtual Reality, which focus primarily on improving technical skills and familiarity with procedures.^{44–46}

In our study, the team explored something different using ChatGPT, which represents a new training avenue for radiographers. This approach not only allows them to evaluate their skills but also to experiment with cutting-edge communication strategies, customized for a wide range of professional expertise levels.

Our investigation focused on simulating a conversation scenario with a patient afflicted by claustrophobia undergoing an MRI. This allowed participants to evaluate their knowledge in a controlled, preclinical environment.

Studies show that managing claustrophobia and stress before performing an MRI examination can increase scan quality,³⁶ decrease scan disruptions^{47,48} and lead to greater patient cooperation.^{49,50}

This approach can be extended to construct diverse clinical simulations, exploring scenarios where radiographers interact with patients during different diagnostic procedures.

Research on mammography procedures underscores the importance of communication strategies concerning the mammography process, breast compression, and pain management as critical topics for effectively managing patient anxiety.^{51,52} Similarly, in radiotherapy sessions, addressing anxiety related to cancer risk and treatment success is a fundamental aspect of the patient's experience.⁵³ Through our pilot study, the necessity for radiographers to utilize communication skills to improve patient care and support is emphasized, identifying this skill as a key focus in their professional training. The integration of ChatGPT in these simulations holds the potential to foster the acquisition of new knowledge and the refinement of existing skills in patient communication techniques.

Our initial approach included evaluating advanced language models such as GPT-4 and Google's Bard,⁵⁴ especially its latest version, Gemini. These models were chosen for their advanced capabilities and relevance to our study's goals.

During our research, we faced a limitation with Bard. Despite its advancements, Bard does not support simulating realistic patient

dialogues, a key component of our study. This limitation is not just technical but also reflects the ethical considerations in language model design. Bard's inability to simulate patient interactions highlights the ethical safeguards that integrated into AI systems, especially in sensitive areas like healthcare.

Our decision to focus on GPT-4 was based not only on its technical superiority but also on its alignment with our research requirements, emphasizing our commitment to a comprehensive and ethically responsible study.⁵⁵ Therefore, our study should be seen as an exploration of the capabilities and limitations of current language models within the ethical boundaries of AI research.

In doing this, we found particularly insightful the possibility to evaluate the progress of the current version of ChatGPT, i.e. the one based on GPT-4, with the previous one, based on GPT-3.5, since the latter can be considered as a baseline for a series of comparative observations.

The team advocates for a balanced and critical approach in adopting such technologies, emphasizing the need for ongoing dialogue and reflection on the ethical implications of AI across various fields.⁵⁶

Knowledge analysis

Exploiting ChatGPT's text-based capabilities, a comprehensive quantitative analysis of radiographers' communication skills was conducted, focusing on their interactions with claustrophobic patients undergoing MRI examinations. The simulation allowed for an examination of participants' skills and techniques in managing these scenarios.

A notable finding was the prevalent use of "Supportive instruction" across all experience levels (radiographer-based) and difficulty levels (case-based).

This consistent application underscores its perceived significance in effectively managing claustrophobic patients, regardless of the complexity of the case or the participants' professional experience, in accordance with the study by Al-Shemmari (2022).⁵⁷

The literature shows that MRI anxiety can be reduced effectively if only by providing patients with standard information about the examination and the scans to be performed.^{58,59} Other studies show that a simple change such as being on a first-name basis with the patient is important to the patient.^{60,61} Instructions to the patient can be minimized due to the time the radiographer can spend communicating with the patient due to keeping the scanner busy and decreasing procedure changeover times. The team pointed out that our study and the choice of giving 2 min of time in the simulation is in line with the evidence that shows a trend to decrease the time spent with the patient increasingly.

Wen X et al.⁶² demonstrate that music therapy is a safe, non-invasive, non-pharmacological and inexpensive method for the patient to reduce anxiety and improve comfort in patients undergoing MRI examinations. Other evidence shows that in more complex cases, the operator can recommend relaxation techniques to reduce anxiety and allow for a high-performance examination.^{47,63} Breathing control and behavioral relaxation training are useful for reducing anxiety and normalizing elevated blood pressure.^{61,64,65}

In line with the findings of previous studies,^{66,67} this suggests that the use of these techniques may vary based on radiographers' experience, with more experienced practitioners preferring approaches that require more collaboration such as "Breathing techniques".

Xiaoyu Wang et al.⁶⁸ point out that a rational choice must be made based on patient compliance, examination duration, and examination site, taking into account the specific characteristics of each patient. This assessment emerged among beginners with a tendency to use 'mild sedation' techniques, even at lower levels of difficulty in contrast to radiographers with high levels of experience. Other studies show that communicating with the patient with recorded films or using Virtual Reality can contribute to a decrease of anxiety.^{45,69} Although the topic is highlighted, the use of Virtual Reality in managing claustrophobic patients was not considered by any participants.⁷⁰ The probable reason for not incorporating Virtual Reality in the simulations could be attributed to the absence of this technology in our institute, thereby creating an implicit bias among the participants. This absence may indicate a limited integration of virtual reality in clinical practices and a potential gap in radiographers' awareness and knowledge.⁶⁷

These disparities in technique preferences across experience levels highlight potential variations in skills, comfort levels, or preferences among radiographers. As indicated in the literature, emphasise that the use of simulation can provide students with an active learning experience, closely modelled on real-life situations, which creates a bridge between theory and practice.^{70,71} This information can be instrumental in identifying training gaps, enabling the development of targeted courses or workshops to enhance staff proficiency in diverse clinical scenarios.

Limitations

While ChatGPT holds considerable potential for medical education, particularly in the training of radiographers, it is not without limitations that may hinder its large-scale implementation. Notably, as ChatGPT is not specifically designed for clinical simulations, the accuracy and reliability of its responses cannot be guaranteed. Research has indicated that its medical responses are "mostly and almost completely correct", which falls short of the precision required for an educational tool.⁷²

Furthermore, a distinction exists between the functionality of the freely accessible GPT-3.5 and the fee-based GPT-4. The latter demonstrated a zero-error rate during simulations, emphasizing

the importance of using the most updated software for enhanced reliability in future simulations.

These results in this study could potentially be achieved with other LLM models as well, hence this experimentation is not aimed at advocating one model over another.⁷³

While virtual role-playing games facilitated by ChatGPT can be valuable, it is crucial to acknowledge their limitations. They cannot fully replace real conversations with patients, particularly when moving from textual to face-to-face interactions. Exploring additional software that replicates voice conversations with patients through video and microphone capabilities may be desirable for a more comprehensive training experience.

Additionally, when utilizing interactive software like ChatGPT, it is essential to consider biases in algorithms. As the model draws information mainly from web sources, responses generated may inadvertently reflect social and systemic biases. While these biases were not perceived in our simulations, it is essential to recognize that algorithms, being created by humans, may inadvertently incorporate implicit and explicit biases.⁷⁴ Ongoing awareness and efforts to mitigate biases should be integral to the responsible use of such technology in education and training.

Following this pilot study, the aim is to increase the number of radiographers and students by integrating the collection of participants' opinions in future studies, using structured methods such as surveys or interviews, to gain a deeper understanding, improve our research results and communication models for evaluating participants and further improve evaluation performances.^{75–77}

Conclusions

The integration of ChatGPT into the training of radiographers proves to be a valuable tool for assessing and enhancing the preparedness of professionals, ultimately contributing to the improvement of patient care.

In our specific case, the utilization of ChatGPT in simulations provided a comprehensive analysis of radiographers' knowledge, considering varying experience levels, particularly in managing claustrophobic patients undergoing MRI examinations. The discerned patterns in the use of specific communication techniques suggested a need for tailored training approaches based on the distinct requirements of professionals at different experience levels.

In conclusion, this pilot study not only underscores the utility of advanced AI techniques in training scenarios but also emphasizes the importance in the educational field for training health personnel and maintaining skills. Acknowledging and addressing the specific needs of professionals across various experience levels are pivotal for optimizing the effectiveness of training methodologies in the field of radiography.

Conflict of interest statement

All authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. Bachmann C, Pettit J, Rosenbaum M. Developing communication curricula in healthcare education: an evidence-based guide. *Patient Educ Counsel* 2022;105(7):2320–7. <https://doi.org/10.1016/j.pec.2021.11.016>.

2. Alam F, Lim MA, Zulkifli IN. Integrating AI in medical education: embracing ethical usage and critical understanding. *Front Med* 2023;10:1279707. <https://doi.org/10.3389/fmed.2023.1279707>.
3. Wouda JC, van de Wiel HBM. The communication competency of medical students, residents and consultants. *Patient Educ Counsel* 2012;86(1):57–62. <https://doi.org/10.1016/j.pec.2011.03.011>.
4. Ryan P, Luz S, Albert P, Vogel C, Normand C, Elwyn G. Using artificial intelligence to assess clinicians' communication skills. *BMJ* 2019;364. <https://doi.org/10.1136/bmj.1161>.
5. Tanana MJ, Soma CS, Srikumar V, Atkins DC, Imel ZE. Development and evaluation of ClientBot: patient-like conversational agent to train basic counseling skills. *J Med Internet Res* 2019;21(7):e12529. <https://doi.org/10.2196/12529>.
6. Bacciu A, Cuconasu F, Siciliano F, Silvestri F, Tonello N, Trappolini G. *RRAML: reinforced retrieval augmented machine learning*. 2023.
7. Hu B, Zhao C, Zhang P, Zhou Z, Yang Y, Xu Z, et al. Enabling intelligent interactions between an agent and an LLM: a reinforcement learning approach. 2023.
8. OpenAI. What is ChatGPT? Available from: <https://help.openai.com/en/articles/6783457-what-is-chatgpt>. [Accessed 11 July 2023].
9. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare* 2023;11(6). <https://doi.org/10.3390/healthcare11060887>.
10. Babl FE, Babl MP. Generative artificial intelligence: can ChatGPT write a quality abstract? *Emerg Med Australasia (EMA)* 2023. <https://doi.org/10.1111/1742-6723.14233>.
11. Hisan UK, Amri MM. ChatGPT and medical education: a double-edged sword. *J Ped Educ Sci* 2023;2(1):71–89. <https://doi.org/10.56741/jpes.v2i01.302>.
12. Huang J, Tan M. The role of ChatGPT in scientific communication: writing better scientific review articles. *Am J Cancer Res* 2023;13(4):1148–54.
13. Chen T-J. ChatGPT and other artificial intelligence applications speed up scientific writing. *J Chin Med Assoc* 2023;86(4):351–3. <https://doi.org/10.1097/JCMA.0000000000000900>.
14. Rao A, Kim J, Kamineni M, Pang M, Lie W, Succu MD, et al. Evaluating ChatGPT as an adjunct for radiologic decision-making. *medRxiv* 2023. <https://doi.org/10.1101/2023.02.02.23285399>.
15. Yeo YH, Samana JS, Ng WH, Ting P-S, Trivedi H, Vipani A, et al. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. *Clin Mol Hepatol* 2023;29(3):721–32. <https://doi.org/10.3350/cmh.2023.0089>.
16. Stogiannos N, Harvey-Lloyd JM, Nugent B, Brammer A, Carlier S, Cleaver K, et al. Autism-friendly MRI: improving radiography practice in the UK, a survey of radiographer practitioners. *Radiography* 2022;28(1):133–41. <https://doi.org/10.1016/j.radi.2021.09.003>.
17. Jeblick K, Schachtner B, Dext J, Mittermeier A, Stüber AT, Topalis J, et al. ChatGPT makes medicine easy to swallow: an exploratory case study on simplified radiology reports. 2022.
18. Benoit JRA. ChatGPT for clinical vignette generation, revision, and evaluation. *medRxiv* 2023. <https://doi.org/10.1101/2023.02.04.23285478>.
19. Bhayana R, Krishna S, Bleakney RR. Performance of ChatGPT on a radiology board-style examination: insights into current strengths and limitations. *Radiology* 2023;307(5). <https://doi.org/10.1148/radiol.230582>.
20. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How does ChatGPT perform on the United States medical licensing examination? The implications of large Language Models for medical education and knowledge assessment. *JMIR Med Educ* 2023;9. <https://doi.org/10.2196/45312>.
21. Currie G, Singh C, Nelson T, Nabasenja C, Al-Hayek Y, Spuur K. ChatGPT in medical imaging higher education. *Radiography* 2023;29(4):792–9. <https://doi.org/10.1016/j.radi.2023.05.011>.
22. Scherr R, Halaseh FF, Spina A, Andalib S, Rivera R. ChatGPT interactive medical simulations for early clinical education: case study. *JMIR Med Educ* 2023;9: e49877. <https://doi.org/10.2196/49877>.
23. Webb JJ. Proof of concept: using ChatGPT to teach emergency physicians how to break bad news. *Cureus* 2023. <https://doi.org/10.7759/cureus.38755>.
24. Rockall AG, Justich C, Helbich T, Vilgrain V. Patient communication in radiology: moving up the agenda. *Eur J Radiol* 2022;155:110464. <https://doi.org/10.1016/j.ejrad.2022.110464>.
25. Forshaw KL, Boyes AW, Carey ML, Hall AE, Symonds M, Brown S, et al. Raised anxiety levels among outpatients preparing to undergo a medical imaging procedure: prevalence and correlates. *J Am Coll Radiol* 2018;15(4):630–8. <https://doi.org/10.1016/j.jacr.2017.12.030>.
26. Munn Z, Pearson A, Jordan Z, Murphy F, Pilkington D, Anderson A. Patient anxiety and satisfaction in a magnetic resonance imaging department: initial results from an action research study. *J Med Imag Radiat Sci* 2015;46(1):23–9. <https://doi.org/10.1016/j.jmir.2014.07.006>.
27. Hudson DM, Heales C, Meertens R. Review of claustrophobia incidence in MRI: a service evaluation of current rates across a multi-centre service. *Radiography* 2022;28(3):780–7. <https://doi.org/10.1016/j.radi.2022.02.010>.
28. Ajam AA, Tahir S, Makary MS, Longworth S, Lang EV, Krishna NG, et al. Communication and team interactions to improve patient experiences, quality of care, and throughput in MRI. *Top Magn Reson Imag* 2020;29(3):131–4. <https://doi.org/10.1097/RMR.0000000000000242>.
29. Al-Shemmari AF, Herbland A, Akudjedu TN, Lawal O. Radiographer's confidence in managing patients with claustrophobia during magnetic resonance imaging. *Radiography* 2022;28(1):148–53. <https://doi.org/10.1016/j.radi.2021.09.007>.
30. ChatGPT OpenAI. Mar 14 version) [Large language model]. 2023.
31. Vaswani A, Brain G, Shazeer N, Parmar N, Uszkoreit J, Jones L, et al. Attention is all you need [n.d].
32. Hudson DM, Heales C, Meertens R. Review of claustrophobia incidence in MRI: a service evaluation of current rates across a multi-centre service. *Radiography* 2022;28(3):780–7. <https://doi.org/10.1016/j.radi.2022.02.010>.
33. Bonfitto G, Roletto A, Savardi M, Fasulo SV, Catania D, Signoroni A. Harnessing ChatGPT dialogues to address claustrophobia in MRI - a radiographers' education perspective. 2023. <https://doi.org/10.5281/zenodo.10441155>.
34. Mollick E, Euchner J. The transformative potential of generative AI. *Res Technol Manag* 2023;66(4):11–6. <https://doi.org/10.1080/08956308.2023.2213102>.
35. OpenAI. ChatGPT [Large language model]. <https://chat.openai.com>.
36. Madl JEM, Nieto Alvarez I, Amft O, Rohleder N, Becker L. The psychological, physiological, and behavioral responses of patients to magnetic resonance imaging (MRI): a systematic review and meta-analysis. *J Magn Reson Imag* 2023. <https://doi.org/10.1002/jmri.29134>.
37. He Z, Xie Z, Jha R, Steck H, Liang D, Feng Y, et al. Large Language models as zero-shot conversational recommenders. Association for Computing Machinery (ACM); 2023. p. 720–30.
38. Szczotka AB, Shakir DL, Clarkson MJ, Pereira SP, Vercauteren T. Zero-shot super-resolution with a physically-motivated downsampling kernel for endomicroscopy. *IEEE Trans Med Imag* 2021;40(7):1863–74. <https://doi.org/10.1109/TMI.2021.3067512>.
39. Meşe İ, Altıntaş Taşlıçay C, Kuzan BN, Kuzan T, Sivrioğlu AK. Educating the next generation of radiologists: a comparative report of ChatGPT and e-learning resources. *Diagn Intervent Radiol* 2023. <https://doi.org/10.4274/dir.2023.232496>. Q(0):0–0.
40. Barash Y, Klang E, Konen E, Sorin V. ChatGPT-4 assistance in optimizing emergency department radiology referrals and imaging selection. *J Am Coll Radiol* 2023;20(10):998–1003. <https://doi.org/10.1016/j.jacr.2023.06.009>.
41. Bera K, O'Connor G, Jiang S, Tirumani SH, Ramaiya N. Analysis of ChatGPT publications in radiology: literature so far. *Curr Probl Diagn Radiol* 2023. <https://doi.org/10.1067/j.cpradiol.2023.10.013>.
42. Bajaj S, Gandhi D, Nayar D. Potential applications and impact of ChatGPT in radiology. *Acad Radiol* 2023. <https://doi.org/10.1016/j.acra.2023.08.039>.
43. Li D, Gupta K, Chong J. Evaluating diagnostic performance of ChatGPT in radiology: delving into methods. *Radiology* 2023;308(3):e232082. <https://doi.org/10.1148/radiol.232082>.
44. Mäkinen H, Haavisto E, Havola S, Koivisto JM. User experiences of virtual reality technologies for healthcare in learning: an integrative review. *Behav Inf Technol* 2020;1–17. <https://doi.org/10.1080/0144929X.2020.1788162>.
45. van der Kruk SR, Zielinski R, MacDougall H, Hughes-Barton D, Gunn KM. Virtual reality as a patient education tool in healthcare: a scoping review. *Patient Educ Counsel* 2022;1928–42. <https://doi.org/10.1016/j.pec.2022.02.005>.
46. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolau CK, George PP, et al. Virtual reality for health professions education: systematic review and meta-analysis by the digital health education collaboration. *J Med Internet Res* 2019. <https://doi.org/10.2196/12959>.
47. Dziuda Ł, Zieliński P, Baran P, Krej M, Kopka L. A study of the relationship between the level of anxiety declared by MRI patients in the STAI questionnaire and their respiratory rate acquired by a fibre-optic sensor system. *Sci Rep* 2019;9(1). <https://doi.org/10.1038/s41598-019-40737-w>.
48. Földes Z, Ala-Ruona E, Burger B, Orsi G. Anxiety reduction with music and tempo synchronization on magnetic resonance imaging patients. *Psychomusicology: Music Mind Brain* 2017;27(4):343–9. <https://doi.org/10.1037/pmu0000199>.
49. Powell R, Ahmad M, Gilbert FJ, Brian D, Johnston M. Improving magnetic resonance imaging (MRI) examinations: development and evaluation of an intervention to reduce movement in scanners and facilitate scan completion. *Br J Health Psychol* 2015;20(3):449–65. <https://doi.org/10.1111/bjhp.12132>.
50. Hudson DM, Heales C, Vine SJ. Radiographer Perspectives on current occurrence and management of claustrophobia in MRI. *Radiography* 2022;28(1):154–61. <https://doi.org/10.1016/j.radi.2021.09.008>.
51. Ding S, Fontaine T, Serex M, Sá dos Reis C. Strategies enhancing the patient experience in mammography: a scoping review. *Radiography* 2024;340–52. <https://doi.org/10.1016/j.radi.2023.11.016>.
52. Metsälä E, Richli Meystre N, Pires Jorge J, Henner A, Kukkes T, Sá dos Reis C. European radiographers' challenges from mammography education and clinical practice – an integrative review. *Insights Imag* 2017;8(3):329–43. <https://doi.org/10.1007/s13244-016-0542-1>.
53. Humphris G, Yang Y, Barracliff L, Cameron J, Bedi C. Emotional talk of patients with breast cancer during review appointments with therapeutic radiographers: effects on fears of cancer recurrence. *Support Care Cancer* 2019;27(6):2143–51. <https://doi.org/10.1007/s00520-018-4484-7>.
54. Lee G-G, Latif E, Shi L, Zhai X. Gemini pro defeated by GPT-4V: evidence from education, vol. 1. Association for Computing Machinery; 2023.
55. Liyanage UP, Ranaweera ND. Ethical considerations and potential risks in the deployment of large Language Models in diverse societal contexts. *J Comp Sociol* 2023;8(11):15–25.
56. Akinci T, Antonoli D, Stanzione A, Bluetngen C, Vernuccio F, Ugga L, et al. Large language models in radiology: fundamentals, applications, ethical considerations, risks, and future directions REVIEW. *Diagn Interv Radiol* 2023. <https://doi.org/10.4274/dir.2023.232417>.
57. Al-Shemmari AF, Herbland A, Akudjedu TN, Lawal O. Radiographer's confidence in managing patients with claustrophobia during magnetic resonance

- imaging. *Radiography* 2022;**28**(1):148–53. <https://doi.org/10.1016/j.radi.2021.09.007>.
58. Munn Z, Pearson A, Jordan Z, Murphy F, Pilkington D, Anderson A. Addressing the patient experience in a magnetic resonance imaging department: final results from an action research study. *J Med Imag Radiat Sci* 2016;**47**(4): 329–36. <https://doi.org/10.1016/j.jmir.2016.04.007>.
59. Portelli JL, McNulty JP, Bezzina P, Rainford L. Benefit-risk communication in paediatric imaging: what do referring physicians, radiographers and radiologists think, say and do? *Radiography* 2018;**24**(1):33–40. <https://doi.org/10.1016/j.radi.2017.08.009>.
60. Tazegul G, Etcioğlu E, Yildiz F, Yildiz R, Tuney D. Can MRI related patient anxiety be prevented? *Magn Reson Imaging* 2015;**33**(1):180–3. <https://doi.org/10.1016/j.mri.2014.08.024>.
61. Wollburg E, Roth WT, Kim S. Effects of breathing training on voluntary hypo- and hyperventilation in patients with Panic Disorder and Episodic Anxiety. *Appl Psychophysiol Biofeedback* 2011;**36**(2):81–91. <https://doi.org/10.1007/s10484-011-9150-5>.
62. Wen X., Shi J., Tan W., Jiang H., Wang D., Su J., et al. Effects of aromatherapy and music therapy on patients' anxiety during MRI examinations: a randomized controlled trial n.d. Doi: 10.1007/s00330-022-09230-3/Published.
63. Lawal O, Regelous P, Omiyi D. Supporting claustrophobic patients during magnetic resonance imaging examination— the radiographer perspective. *Radiography* 2024;**30**(1):80–6. <https://doi.org/10.1016/j.radi.2023.10.011>.
64. Norbakh A, Yucel K, Yuh W, Doros G, Ajam A, Lang E, et al. Effect of team training on improving MRI study completion rates and no-show rates. *J Magn Reson Imag* 2016;**44**(4):1040–7. <https://doi.org/10.1002/jmri.25219>.
65. Rygiel K. Selected relaxation techniques and possibilities of their use in the context of stress-related psychophysical disorders and diseases. *Neuropsychiatra i Neuropsychologia/Neuropsych Neuropsychol* 2017;**12**(3):126–33. <https://doi.org/10.5114/nan.2017.71670>.
66. Zhou Y, Cao Y, Xu S, Li S, Liang Y, Zhang W, et al. Randomized controlled trial on the effects of mindfulness-based respiratory decompression therapy in claustrophobic patients undergoing MRI inspection. *Front Med* 2023;**10**. <https://doi.org/10.3389/fmed.2023.1253824>.
67. Hamd ZY, Alorainy AI, Alrujaee LA, Alshdayed MY, Wdaani AM, Alsabaie AS, et al. How different preparation techniques affect MRI-induced anxiety of MRI patients: a preliminary study. *Brain Sci* 2023;**13**(3). <https://doi.org/10.3390/brainsci13030416>.
68. Wang X, Liu X, Mi J. Perioperative management and drug selection for sedated/ anesthetized patients undergoing MRI examination: a review. *Medicine* 2023;**102**(16):E33592. <https://doi.org/10.1097/MD.00000000000033592>.
69. Francová A, Jablonská M, Fajnerová I. Design and evaluation of virtual reality environments for claustrophobia. *Pres: Virtual and Augmented Reality* 2023: 1–12. <https://doi.org/10.1162/pres.a.00385>.
70. Bwanga O. Radiography students' perceptions and experiences of their clinical placements-A qualitative systematic review. [n.d].
71. Karera A, Engel-Hills P, Davidson F. Radiographers' experiences of image interpretation training in a low-resource setting. *Radiography* 2023;**29**(3): 590–6. <https://doi.org/10.1016/j.radi.2023.03.012>.
72. Johnson D, Goodman R, Patrinely J, Stone C, Zimmerman E, Donald R, et al. Assessing the accuracy and reliability of AI-generated medical responses: an evaluation of the chat-GPT model. *Res Sq* 2023. <https://doi.org/10.21203/rs.3.rs-2566942/v1>.
73. Doshi R., Amin ; Kanhai., Pavan ;, Ba K., Bajaj ; Simar, Chheang S., et al. Utilizing large Language Models to simplify radiology reports: a comparative analysis of ChatGPT3.5, ChatGPT4.0, Google bard, and microsoft bing n.d. Doi: 10.1101/2023.06.04.23290786.
74. Holmes W, Bialik M, Fadel C. *Artificial intelligence in education. Promise and implications for teaching and learning*. Boston: MA: The Center for Curriculum Redesign; 2019.
75. Windover AK, Boissy A, Rice TW, Gilligan T, Velez VJ, Merlino J. *The REDE model of healthcare communication: optimizing relationship as a therapeutic agent*. 2013.
76. Rockall AG, Justich C, Helbich T, Vilgrain V. Patient communication in radiology: moving up the agenda. *Eur J Radiol* 2022;**155**:110464. <https://doi.org/10.1016/j.ejrad.2022.110464>.
77. Wouda JC, van de Wiel HBM. The communication competency of medical students, residents and consultants. *Patient Educ Counsel* 2012;**86**(1):57–62. <https://doi.org/10.1016/j.pec.2011.03.011>.

Chapter 3

CT Body Composition Analysis: Semi-automatic Segmentation of L1 and L3 during International Fellowship

3.1 Introduction and Motivation

Between 20 September 2024 and 31 March 2025, I carried out an international research fellowship as a scientific collaborator (20% FTE) at the Istituto Imaging della Svizzera Italiana, part of the Ente Ospedaliero Cantonale in Locarno, Switzerland. This mobility experience represented a strategic step in my doctoral training, conceived to strengthen my technical and methodological expertise in quantitative CT image analysis, with a specific focus on the manual segmentation of body composition at the lumbar levels L1 and L3.

This activity was not an isolated training exercise but was deliberately planned as a preparatory step for the core PhD project on AI-driven personalization of iodinated contrast dosing in CT. Developing a robust anatomical and technical understanding of body composition was considered essential to later build and validate AI models capable of integrating patient-specific morphometric information.

The EOC offered an ideal environment to carry out this work because of its strong integration between clinical radiology and research, its advanced imaging infrastructure with standardized PACS systems, and its well-established interdisciplinary collaboration between radiologists, oncologists, and medical physicists. These conditions enabled a structured and methodologically rigorous workflow, ensuring both scientific robustness and clinical relevance.

The goal of the fellowship was threefold: to acquire advanced segmentation skills through the use of SliceOmatic, to construct a high-quality dataset of segmented CT

images at vertebral levels L1 and L3, and to establish a methodological foundation for the development of AI-based imaging solutions. The experience also provided a valuable opportunity to operate in an international and multidisciplinary context, strengthening my ability to work within multi-center imaging workflows and to interact effectively with professionals from different clinical and technical backgrounds.

3.2 Materials and Methods

3.2.1 Dataset

The CT scans analyzed in this study were retrospectively retrieved from the institutional PACS of multiple Swiss hospitals, including the Ospedale Regionale di Lugano, the Kantonsspital St. Gallen, and the Kantonsspital Graubünden. All patients belonged to the SAKK 08/14 oncologic cohort, which provided imaging data acquired with standardized protocols across different institutions. Only adult patients with high-quality axial CT images were included.

Two vertebral levels were selected for segmentation: L3, which represents the conventional anatomical reference for body composition studies, and L1, which is increasingly gaining interest in thoracic imaging applications where the L3 level may not be present in the field of view. Both contrast-enhanced and non-contrast-enhanced acquisitions were analyzed, allowing the evaluation of segmentation stability across acquisition protocols and imaging conditions.

3.2.2 Segmentation Protocol

The segmentation process was performed with SliceOmatic (version 5.0), a specialized medical imaging software designed to provide high precision and

reproducibility in the quantitative analysis of anatomical structures on DICOM CT images.

The software is characterized by an engineering-oriented architecture built around three fundamental principles. The first is HU-driven segmentation, which relies directly on the native Hounsfield Unit values of CT pixels to automatically classify tissue types. This physics-based classification allows a clear and standardized discrimination between skeletal muscle, visceral fat, and subcutaneous fat, reducing subjectivity and improving reproducibility.

The second key principle is its modular structure, which separates the main functional components of the workflow. The MODES interface enables selection of segmentation modality, such as the Abacus Auto 2D mode used in this study. The TOOLS section allows interactive contouring and refinement of anatomical interfaces, while the COMPUTE module performs quantitative calculations of surface area, volume, mean HU values, and standard deviations. The TAG SURFACE/VOLUME function is used to apply standardized labels to each segmented tissue, ensuring data consistency across cases and operators.

The third principle is the software's high spatial precision, supported by sub-voxel refinement and real-time feedback on HU values, which allows the operator to precisely control tissue boundaries. Native DICOM compatibility enables direct loading of images from PACS without the need for format conversion, preserving all acquisition metadata. The interoperability of the output, exported in structured formats such as CSV or XLS, facilitates integration into statistical environments and future AI pipelines.

The operational workflow followed a structured, reproducible sequence. First, DICOM images were imported and visualized, with HU windowing adjusted to maximize contrast between tissue compartments. Segmentation was then initiated in Abacus

Auto 2D mode, applying standard HU thresholds: -29 to $+150$ HU for skeletal muscle, -150 to -50 HU for visceral fat, and -190 to -30 HU for subcutaneous fat. Manual refinement was subsequently performed at anatomical interfaces using the TAG SURFACE/VOLUME tool to correct potential misclassifications and ensure anatomical accuracy. Once segmentation was complete, the COMPUTE module generated quantitative measurements, including cross-sectional area in cm^2 , tissue volume, and mean HU values. Each ROI was then labeled according to tissue type and vertebral level, and the data were exported into structured spreadsheets to support descriptive analysis and future integration with AI training datasets.

3.2.3 *Extracted Parameters*

For each patient and vertebral level, segmentation produced a structured set of quantitative parameters, including total cross-sectional area, mean HU and standard deviation per ROI, density score, acquisition phase (contrast vs. non-contrast), and metadata regarding hospital site and patient cohort. This parameterization allowed a uniform characterization of body composition at both L1 and L3, ensuring that the dataset could later be used for comparative studies or as ground truth for AI training.

3.3 Results

The dataset included approximately forty patients, with segmentation performed at both L1 and L3 levels. Examinations covered a period from September 2024 to March 2025 and were collected across three Swiss hospitals. All segmentations were performed manually, with the support of HU-based pre-classification to accelerate and standardize the process.

The descriptive analysis showed that skeletal muscle HU values ranged approximately between 17 and 50 HU, consistent with reference literature. L3 displayed greater inter-patient variability in surface area compared to L1, reflecting its anatomical role as a more heterogeneous level in terms of body composition. The reproducibility of measurements across the three hospital sites confirmed the stability of the acquisition protocols and the robustness of the segmentation workflow.

Observationally, the analysis demonstrated that L1 segmentation is both feasible and reliable. This finding supports its potential as an alternative reference level in clinical or research contexts where L3 is not available. Manual segmentation also increased understanding of the anatomical variability and HU distribution patterns of lumbar tissues, which represents an essential step in designing accurate and generalizable AI models. Moreover, familiarity with the workflow improved both segmentation speed and precision, enhancing the overall quality of the dataset.

3.4 Discussion

This segmentation activity represented a strategic methodological and training component of my doctoral project. From a scientific standpoint, the study demonstrated the feasibility of performing standardized manual body composition segmentation across multiple centers and confirmed that L1 can function as a valid surrogate for L3 in many imaging contexts. HU-based manual segmentation provided reliable reference values, essential for the future development of AI-based algorithms aimed at personalizing contrast media dosing in CT.

From a training perspective, this experience offered an opportunity to gain hands-on expertise in handling DICOM datasets, performing image pre-processing, and operating advanced segmentation software in a clinically integrated environment. It

also strengthened my familiarity with lumbar anatomical variability and radiodensity distribution, skills that would later prove fundamental for the design of data-driven predictive models. Working within a multidisciplinary and international research setting further enhanced my ability to collaborate effectively and to adapt to different methodological standards.

The engineering-oriented design of SliceOmatic proved particularly advantageous for generating structured, high-quality ground truth data. Its modularity and precision made it well suited for projects that require a reproducible and scalable methodology, such as the subsequent development of AI models for CT protocol optimization.

3.5 Limitations and Future Perspectives

This work has several limitations that should be acknowledged. The patient cohort was relatively small, and the segmentation process was entirely manual, which limited scalability. Automated segmentation approaches based on deep learning were not implemented at this stage. However, this limitation was intentional: the goal was to establish a high-quality reference dataset and a solid methodological foundation, rather than to develop automated tools prematurely.

Future work will focus on the development of automatic segmentation pipelines for L1 and L3, integrating these data with predictive algorithms for contrast dosing and validating the approach in larger, multicenter cohorts. This will allow the transition from manual to automated processing while preserving the precision and standardization achieved in this preliminary phase.

3.6 Contributions

This segmentation activity contributed in a substantial and structured way to the overall trajectory of the doctoral project. It enabled the acquisition of advanced segmentation expertise in an international research setting, the creation of a multi-center standardized dataset at two vertebral reference levels, and the establishment of a methodological bridge between manual segmentation and AI-based imaging solutions. Most importantly, it provided direct scientific support to the core project on personalized iodinated contrast media dosing, anchoring the later AI modeling work in a solid anatomical and quantitative ground truth.

Chapter 4

Synthetic FDG-PET Generation from Early-Phase Amyloid PET: Experimental Framework for AI Architectures

4.1 Introduction and Motivation

Molecular imaging using PET has long played a central role in the assessment of neurodegenerative disorders, particularly Alzheimer's disease, where FDG-PET captures cerebral glucose metabolism as a sensitive marker of neuronal dysfunction and amyloid PET visualizes amyloid- β pathology in vivo. (Chételat et al., 2020; Minoshima et al., 2021; Rice & Bisdas, 2017) The complementary nature of these modalities adds diagnostic value but also increases radiation exposure, procedure duration, costs, and workflow complexity because two separate acquisitions must be scheduled and analyzed for each patient. (Chételat et al., 2020; Høilund-Carlsen et al., 2024) In the last few years, deep learning-based generative models have emerged as a way to mitigate these burdens by synthesizing FDG-like images directly from other inputs, including early-phase amyloid PET that reflects perfusion-like signal coupled to metabolism. (Sanaat et al., 2024; Islam & Zhang, 2020; Kollovieh et al., 2022) This relies on the physiological coupling between perfusion and metabolism and has already shown high structural similarity between synthetic and native FDG images in research settings, opening a path toward clinical translation if robustness and evaluation standards are met. (Sanaat et al., 2024; Dayarathna et al., 2023; Wang, Bashyam, et al., 2022)

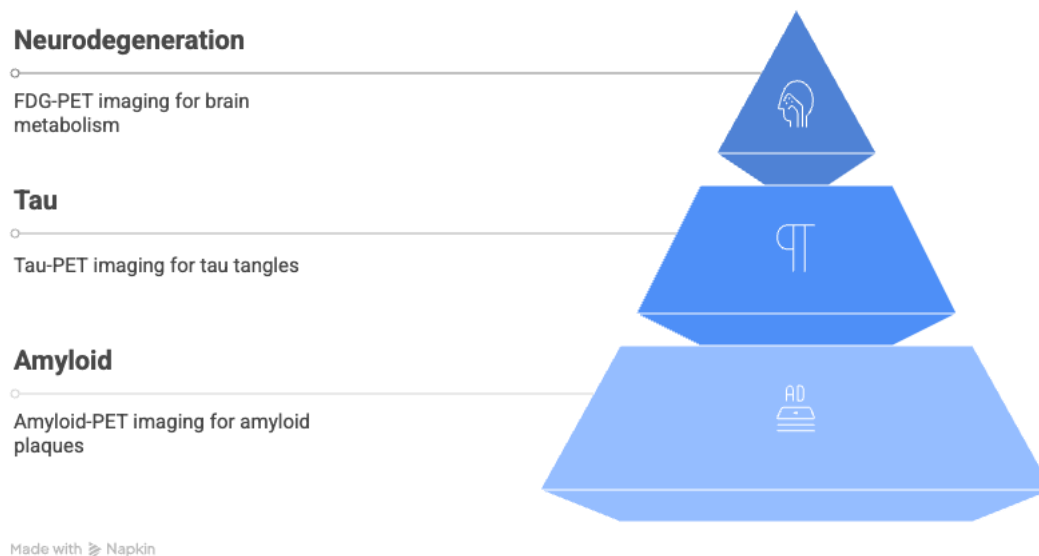


Figure 4 – Alzheimer's Biomarker Framework

Within my doctoral project, this study functioned as a methodological bridge toward AI-driven personalization of iodinated contrast dosing in CT: the goal here was to establish a reproducible end-to-end pipeline for synthesizing FDG from early-phase amyloid PET, to compare representative generative architectures, and to consolidate practical expertise in preprocessing, model training, and evaluation on real neuroimaging datasets so that the lessons learned could be transferred to downstream clinical applications. (Dayarathna et al., 2023; Wang et al., 2024)

Study	Tech	Input	SSIM	Convalidation	Open Source	Other Convalidation
Zhang 2021	3D U-Net	Amyloid	0.85	No	No	No
Chen 2023	cGAN	Amyloid+MRI	0.82	No	Yes/No	No

4.2 Materials and Methods

4.2.1 Dataset

The dataset included paired amyloid and FDG PET scans from patients imaged with ¹⁸F-flutemetamol and ¹⁸F-florbetaben tracers. Each pair consisted of an early-phase amyloid acquisition (0–6 min p.i.) and a late-phase FDG scan performed within a close temporal window. Images were spatially co-registered using ANTs, reoriented to MNI space, and intensity-normalized by SUVR using the pons as the reference region. The pipeline, implemented in Python using nibabel and scikit-image, ensured isotropic resolution (2 mm³ voxels) and smoothed data with a Gaussian filter (FWHM = 4 mm) to harmonize between sites.

The dataset was split into training (70%), validation (15%), and testing (15%) sets on a patient-wise basis, ensuring that no slices from the same subject appeared in both training and evaluation subsets to prevent data leakage. The U-Net served as the baseline model, trained with L1 + SSIM loss. The GAN implementation incorporated a PatchGAN discriminator, improving spatial contrast reproduction, while Transformer and Diffusion models were included to explore more recent generative paradigms: the former using multi-head self-attention to model long-range dependencies, the latter learning to denoise from Gaussian noise trajectories to reconstruct realistic FDG images. Optimization used the Adam algorithm (learning rate 1×10^{-4}), and evaluation metrics included MAE, PSNR, and SSIM.. (Chételat et al., 2020; Minoshima et al., 2021)

4.2.2 Pre-processing

All PET images underwent a standardized pipeline for geometric and intensity harmonization between early-phase amyloid and FDG scans: rigid-to-affine PET–PET co-registration, SUV normalization, isotropic Gaussian smoothing with 4-mm

FWHM, and resampling to MNI space to ensure voxel-wise correspondence; optionally, T1-weighted MRI was concatenated as an anatomical prior. (Nørgaard et al., 2019; Samper-González et al., 2018; Lee et al., 2022) These steps reflect best practices reported to improve stability and comparability in PET-based neuroimaging analyses and AD classification pipelines. (Samper-González et al., 2018; Nozadi et al., 2018)

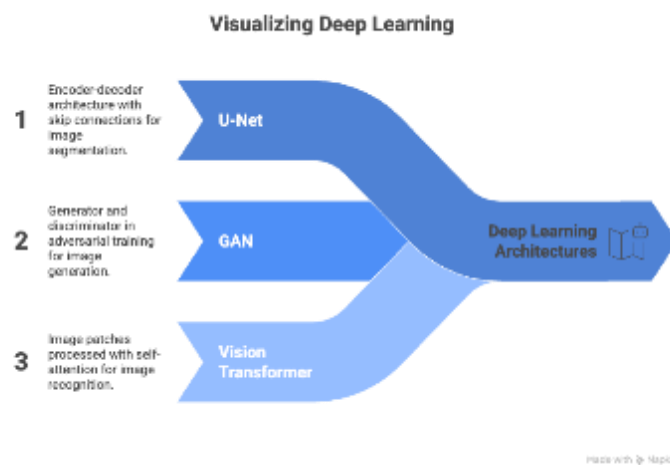


Figure 5: Visualizing Deep Learning

4.3 AI Architectures

Four representative families were compared to cover the main paradigms of medical image synthesis: convolutional encoder–decoder networks (U-Net), adversarial generative models (GAN), attention-based models (Transformers), and denoising diffusion probabilistic models (Diffusion). (Dayarathna et al., 2023; Wang et al., 2024; Islam et al., 2024)

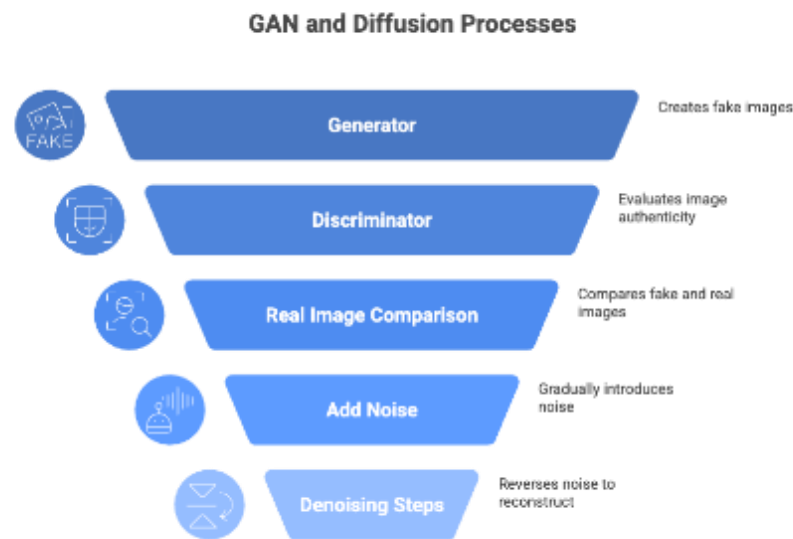


Figure 6: GAN and Diffusion Processes

U-Net, originally developed for biomedical segmentation and later generalized to image-to-image translation, leverages encoder–decoder symmetry with skip connections to preserve global structure while restoring local detail; we combined L1 and SSIM losses to balance pixel-wise fidelity with perceptual similarity, capitalizing on U-Net’s stability and modest data needs but acknowledging its limitations in modeling complex non-linear dependencies that drive subtle cortical variations. (Siddique et al., 2021; Azad et al., 2024; Zaharchuk et al., 2018)

GANs introduce an adversarial objective by pairing a generator (here, a U-Net–like architecture) with a PatchGAN discriminator that operates on local patches, often sharpening cortical textures and metabolic gradients at the expense of greater training instability and sensitivity to hyperparameters. (Islam & Zhang, 2020; Wang, Bashyam, et al., 2022; Zhao et al., 2022)

Transformer-based architectures employ multi-head self-attention to capture long-range dependencies, well-suited to distributed network alterations in AD, but typically require larger datasets and compute; in medical imaging, ViT-like models and hybrid

UNet-Transformer variants have shown promise for synthesis and multimodal fusion, with training stability depending strongly on data volume and regularization. (Azad et al., 2023; Dalmaz et al., 2021; Li et al., 2024)

Diffusion models learn to invert a progressive noising process and are trained with likelihood-based objectives that yield stable convergence and high-fidelity samples, at the cost of slower inference; recent surveys and exemplars highlight their robustness and favorable artifact profile in medical imaging. (Kazerouni et al., 2022; Dayarathna et al., 2023; Wang et al., 2024) The selection of these four families was thus deliberate: U-Net as a stable baseline, GAN to probe high-frequency realism, Transformers for distributed dependencies, and Diffusion as the current state of the art for fidelity and stability. (Dayarathna et al., 2023; Wang et al., 2024)

4.4 Training and Evaluation

Training used a 70/15/15 split for train/validation/test, Adam optimizer with a 1×10^{-4} learning rate, and batch size of four. Quantitative evaluation considered MAE for voxel-wise accuracy, PSNR for signal preservation, and SSIM for structural fidelity, computed in AD-relevant ROIs derived from an AAL atlas. (Samper-González et al., 2018) Qualitative evaluation comprised a blinded reader study with neuroradiologists and nuclear medicine physicians using a structured visual scale to judge similarity between synthetic and real FDG, aligning with recommendations to pair objective metrics with expert assessment in neuroimaging AI studies. (Chételat et al., 2020; Drzezga et al., 2018)

4.5 Results

U-Net and GAN achieved SSIM values >0.80 and MAE <0.25 SUV, indicating good global approximation of metabolic patterns; Transformers reached comparable

averages but exhibited higher variance, consistent with their greater data dependence; diffusion models achieved the highest SSIM with smooth learning curves suggestive of stable generalization. (Islam & Zhang, 2020; Sanaat et al., 2024; Kazerouni et al., 2022) Qualitatively, U-Net reproduced global structure but softened fine cortical detail; GAN improved cortical delineation with occasional localized artifacts; Transformers captured distributed patterns most clearly in associative cortices; diffusion results were visually closest to native FDG and received the highest reader scores. (Islam & Zhang, 2020; Dalmaz et al., 2021; Kazerouni et al., 2022) These findings support the hypothesis that early-phase amyloid PET carries sufficient information to predict FDG metabolism and that adversarial and diffusion paradigms yield superior structural fidelity when appropriately trained and regularized. (Sanaat et al., 2024; Dayarathna et al., 2023)

4.6 Discussion

This experimental phase was not merely an exercise in image generation but a controlled comparison of deep learning paradigms applied to neurodegenerative imaging. Evidence from this study and the literature indicates that FDG synthesis from early-phase amyloid PET is technically feasible and clinically promising provided that preprocessing, harmonization, and evaluation are rigorous. (Chételat et al., 2020; Sanaat et al., 2024; Nørgaard et al., 2019) Methodologically, the work strengthened reproducible pipelines that can transfer to other modalities and tasks; clinically, synthetic FDG could reduce exposure and cost while preserving diagnostic signal if validated at scale and embedded within quality-assured workflows. (Højilund-Carlsen et al., 2024; Dayarathna et al., 2023) In the architectural comparison, diffusion models offered the most balanced mix of fidelity and stability, GANs excelled in sharpness at

the cost of training complexity, U-Net provided a strong, efficient baseline, and Transformers offered a path to modeling network-level relationships with suitable data regimes. (Kazerouni et al., 2022; Islam & Zhang, 2020; Azad et al., 2023)

4.7 Limitations and Future Perspectives

The main limitations concern the modest size of the local cohort, single-center external evaluation, and the runtime demands of diffusion sampling; broader multicenter testing, accelerated sampling (e.g., latent diffusion), and uncertainty/explainability tooling will be necessary to approach clinical deployment. (Kazerouni et al., 2022; Seiler & Ritter, 2024; Azad et al., 2023) Future work should target PACS-integrated prospective studies with standardized preprocessing and reader studies powered for non-inferiority to native FDG in specific diagnostic questions. (Chételat et al., 2020; Høilund-Carlsen et al., 2024)

4.8 Contributions

This chapter establishes an end-to-end synthetic FDG pipeline from early-phase amyloid PET, offers a principled comparison of leading generative paradigms, and distills methodological guidance that directly informs the subsequent core project on AI-driven CT contrast personalization, where data coupling, architecture stability, and evaluation rigor are likewise decisive. (Dayarathna et al., 2023; Wang et al., 2024)

4.9 Dissemination and Conference Presentations

The preliminary outcomes of this research have already achieved significant visibility within the European radiography and molecular imaging communities. The work entitled “Synthetic FDG-PET Image Generation with Deep Learning Model from Early-Phase Dynamic Amyloid-PET” was selected for oral presentation at the 1st

EFRS Symposium, “Advancing Radiography,” held in Rome on 14–15 November 2025. The symposium, organized by the European Federation of Radiographer Societies (EFRS), represented a milestone in advancing professional-led research across imaging, radiotherapy, and nuclear medicine domains. The inclusion of this study in the scientific programme (session Advancing Radiography through Research, chaired by Jonathan Portelli and Helle Prect) underscored its methodological novelty and its relevance to the evolving role of radiographers in AI-driven quantitative imaging.

Following the Rome presentation, the same project was accepted as an oral contribution for the European Congress of Radiology (ECR) 2026, Vienna, within the session dedicated to AI in Hybrid and Molecular Imaging. This acceptance marks an important recognition of the work’s scientific maturity and its translational potential for clinical imaging practice. The two presentations collectively testify to the project’s European dimension, its interdisciplinary positioning, and its alignment with the broader research goals of this PhD, bridging molecular image synthesis and AI-based clinical personalization in CT.

Conference Presentation

- Bonfitto, G.; Temel K (2025) *Synthetic FDG-PET Image Generation with Deep Learning Model from Early-Phase Dynamic Amyloid-PET*. Oral presentation at the EFRS Symposium, Rome, November 2025.
- Temel K; Bonfitto, G. (2026) *Synthetic FDG-PET Image Generation with deep learning model from Early Phase Dynamic Amyloid-PET*. Accepted oral presentation at the European Congress of Radiology (ECR 2026), Vienna.

Chapter 5

AI-Driven Personalization of Iodinated Contrast Media Administration in CT

5.1 Introduction

Contrast-enhanced CT remains a cornerstone of diagnostic imaging, yet iodinated contrast administration is still predominantly based on fixed or semi-fixed protocols, largely scaled to patient weight, examination type, and scanner characteristics. Although these approaches are practical and operationally robust, they fail to account for substantial inter-individual variability in body composition, thoracoabdominal geometry, vascular volume, tissue attenuation, and haemodynamic status, all of which critically determine intravascular iodine concentration and parenchymal enhancement (Kartalis et al., 2017; Caschera et al., 2016). As a result, standard dosing strategies systematically produce two clinically relevant outcomes: over-enhancement in smaller patients, who receive more iodine than necessary, and under-enhancement in patients with larger or atypical body habitus, where insufficient contrast compromises diagnostic confidence and may necessitate repeat imaging (Davenport et al., 2017).

To mitigate this variability, several patient-tailored dosing strategies have been proposed. Approaches based on body surface area or lean body weight have demonstrated improved prediction of iodine distribution volume compared with total body weight, leading to more consistent enhancement and reduced unnecessary iodine exposure, particularly in patients at the lower end of the size spectrum (Walgraeve et al., 2019; Zanardo et al., 2020; Chandrasekharan et al., 2021). In

parallel, multiple studies have shown that morphometric indices derived directly from preliminary CT images outperform anthropometric estimators in guiding contrast administration, reinforcing the importance of physiologically grounded parameters (Kondo et al., 2013; Guerrisi et al., 2014; Borga et al., 2018).

Within this context, the CT scout image represents a largely underexploited source of patient-specific information. Acquired routinely at negligible radiation dose for positioning and planning, the scout inherently encodes global morphology, attenuation distribution, and cross-sectional geometry. Recent investigations have demonstrated that scout images can reliably estimate parameters such as body weight, effective diameter, and water-equivalent diameter, which correlate more closely with contrast enhancement behaviour than conventional anthropometric measures (Ichikawa et al., 2021; Bruns et al., 2021). These findings suggest that the scout contains sufficient physiological information to serve as a predictor of iodine distribution and enhancement response.

The increasing availability of machine learning methods has further strengthened the potential of scout-based prediction. Deep learning models applied to scout views have shown high accuracy in extracting morphometric features relevant to dose planning and protocol optimisation (Ichikawa et al., 2021; Bruns et al., 2021). However, the clinical adoption of such approaches remains limited by the lack of standardisation across scanner geometries, patient positioning, and fields of view, as well as by the absence of integrated frameworks capable of translating scout-derived information into actionable dosing recommendations.

The present PhD project addresses this gap by proposing an AI-driven framework that transforms the routinely acquired CT scout into a pre-injection predictor of the optimal iodinated contrast dose required to achieve predefined diagnostic enhancement targets in vascular and parenchymal structures. By integrating image-derived morphological descriptors with acquisition and dosing metadata, the framework moves beyond empirical protocol-based strategies toward a patient-specific, physiology-informed approach to contrast administration. This shift aligns with the broader objectives of precision imaging, aiming to reduce enhancement variability, minimise unnecessary iodine exposure, and improve both clinical and operational efficiency in CT practice.

5.1.1 Performance of DINOv2 in Medical Imaging

The adoption of DINOv2, a self-supervised Vision Transformer, is supported by emerging evidence of its capability to generalize well to medical imaging. Without medical-specific retraining, DINOv2 has achieved superior performance in classification and segmentation tasks across modalities, including CT and X-ray (Touvron et al., 2023; Kirillov et al., 2023). This is attributed to the transformer's self-attention mechanism, which models long-range spatial relationships, crucial for interpreting global patient morphology from scout images.

Recent evaluations have shown that Vision Transformers like DINOv2 not only outperform standard CNNs in many radiological benchmarks but also preserve cross-task generalizability, enabling their use across anatomical sites and imaging tasks (Touvron et al., 2023). The architecture thus provides a suitable foundation for encoding visual features relevant to contrast enhancement prediction.

5.2 Study Population and CT Dataset

The study population comprises a retrospective cohort of approximately one thousand contrast-enhanced CT examinations retrieved from the institutional PACS of a tertiary referral centre. The dataset reflects the heterogeneity typically encountered in routine clinical practice, including thoracic, abdominal and combined thoracoabdominal studies performed for oncological staging, acute care assessment and follow-up imaging. All examinations were acquired using modern multi-detector CT systems equipped with automatic exposure control and standardised protocol settings, ensuring sufficient homogeneity for the development of a predictive model. The acquisition parameters were consistent across scanners. Tube voltage values ranged between 100 and 120 kV, selected through automated kV modulation to optimise the interplay between iodine attenuation and radiation dose. Collimation widths varied between 0.6 and 1.25 mm and the pitch values ranged from 0.8 to 1.2, reflecting contemporary practice in thoracoabdominal CT imaging. These settings are in line with current recommendations for liver and abdominal protocols, which emphasise stable timing, adequate parenchymal enhancement and cross-sectional uniformity (Kartalis et al., 2017).

Patient eligibility followed strict quality and completeness criteria. Only adults with complete DICOM metadata and preserved scout-view geometry were included. Exclusion criteria involved severe metallic artefacts, extensive orthopaedic hardware, motion degradation and incomplete contrast administration records. The final cohort therefore provides a robust foundation for the construction of morphometric and attenuation-based models.

Each examination consisted of a pre-contrast anterior–posterior scout image followed by one or more contrast-enhanced acquisitions. The scout, traditionally limited to patient positioning, plays a central role in this project as it contains quantifiable information related to body habitus, thoracoabdominal thickness and the expected intravascular distribution volume of iodinated contrast. Previous work has demonstrated that parameters derived from scout images, particularly effective diameter and water-equivalent diameter, correlate strongly with patient size and can outperform anthropometric estimators such as total body weight and body surface area (Kondo et al., 2013; Ichikawa et al., 2021). These findings support the rationale for using the scout as the primary input for a predictive model.

The preparation of the dataset followed a rigorous and reproducible workflow. All examinations were pseudonymised at source to ensure compliance with institutional and regulatory policies. Images were converted from DICOM to NIfTI format to facilitate integration into Python-based processing pipelines. Scanner parameters including kV, mAs, contrast concentration and injection rate were extracted from the DICOM headers and harmonised into structured metadata tables.

Anatomical reference levels were standardised using reproducible landmarks, including the aortic arch, hepatic dome and renal hilum. Within these levels, circular and elliptical regions of interest were defined in key structures spanning vascular and parenchymal compartments. Measurements of mean Hounsfield Unit intensity were obtained from the ascending aorta, portal vein, liver parenchyma and renal cortex. These structures were selected because they represent clinically relevant enhancement targets and because their enhancement kinetics are sensitive to contrast distribution volume, injection rate and individual physiology (Walgraeve et al., 2019).

The dataset was organised into two complementary components. The first component consists of the input matrix derived from the scout image and acquisition parameters. This includes morphometric features such as anterior–posterior diameter, lateral dimension, effective cross-sectional area and water-equivalent diameter, along with histogram-derived attenuation descriptors. The second component comprises the output matrix derived from post-contrast measurements, representing the ground truth enhancement achieved in the vascular and parenchymal ROIs.

By structuring the dataset in this supervised learning configuration, the study enables an explicit quantitative mapping between patient morphology, contrast administration parameters and resulting enhancement. This structure forms the foundation for the development of the forward prediction model described in the subsequent section and ultimately for the inversion model required for personalised contrast dose estimation.

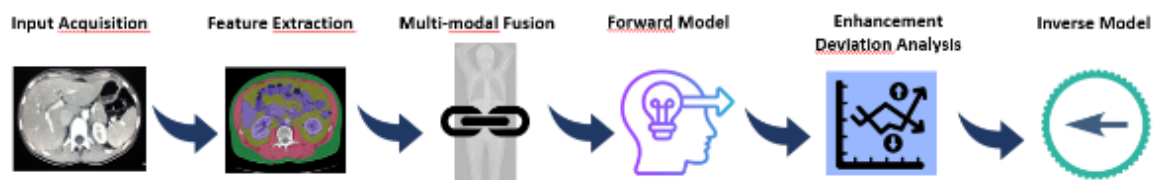


Figure 7: CT Contrast Personalization Pipeline

5.3 Image Analysis Pipeline from Scout

The analytical pipeline developed for this project is grounded in the assumption that the CT scout image contains sufficient morphometric and attenuation-based information to predict the behaviour of iodinated contrast during the diagnostic phase. Although acquired with minimal dose and limited structural detail, the scout image

nevertheless encodes key descriptors of patient size, thoracoabdominal geometry and tissue composition. These factors determine the effective distribution volume of iodine and substantially influence the enhancement achieved in both vascular and parenchymal structures.

The workflow begins with automated segmentation of the scout. A dedicated Python environment was developed to extract the external body contour using threshold-based and gradient-sensitive operators. This segmentation step allows the computation of quantitative descriptors including anterior–posterior diameter, lateral diameter, cross-sectional area and water-equivalent diameter. These metrics summarise the individual's attenuation profile and serve as physiologically meaningful proxies for contrast distribution capacity. By incorporating both geometric and radiodensity-weighted variables, the pipeline moves beyond traditional anthropometric measures, enabling a more nuanced characterisation of patient morphology.

In parallel, the scout image undergoes histogram-based analysis to quantify its global and regional intensity distribution. Pixel intensity patterns reflect the relative proportions of adipose, soft tissue and bone, each of which contributes differently to overall attenuation. Additional texture descriptors, such as local intensity variance, edge density and histogram-derived moments, provide further insight into structural composition and heterogeneity. These features enrich the modelling space by capturing subtle variations that would otherwise not be accounted for by simple geometric measurements.

The pipeline then integrates scout-derived features with acquisition metadata, including tube voltage, tube current, collimation and the contrast dose that was actually administered during the clinical examination. Each patient is therefore

represented by a multidimensional input vector describing body morphology, attenuation characteristics and scanner parameters.

The corresponding output vector consists of the measured Hounsfield Unit values obtained from diagnostic-phase images. Regions of interest are placed in anatomically consistent locations within the ascending aorta, the portal venous system, the hepatic parenchyma and the renal cortex. These structures were chosen because they are central to routine diagnostic assessment and demonstrate distinct but predictable enhancement patterns. Their quantitative HU values provide a reliable ground truth for evaluating the relationship between contrast dose, morphology and enhancement outcome.

This configuration enables the training of a supervised regression model that predicts enhancement intensity from scout features and administered dose. The model architecture incorporates convolutional components to process the spatial information of the scout image and fully connected layers to integrate extracted features with protocol metadata. During optimisation, the network learns to minimise the difference between predicted and observed HU values, progressively capturing the nonlinear relationship linking morphology, iodine load and enhancement response.

The result of this forward prediction model is a functional mapping that quantifies how individual anatomical characteristics influence contrast uptake. The evaluation of the model across different subgroups, including variations in body habitus, allows the identification of systematic trends, prediction intervals and regions of uncertainty. These analyses define the operative domain within which the subsequent inversion model can be safely applied.

This pipeline therefore represents the foundational methodological stage of the broader project. It transforms the scout image from a simple positioning tool into a

quantitative source of physiological information and establishes the computational basis required for the development of patient-specific contrast dosing strategies. By structuring the workflow around reproducible image-derived parameters, the framework supports a transition toward personalised CT imaging driven by objective and measurable pre-scan data.

5.3.1 Enhancement Deviation and Threshold Analysis

To better characterize the variability of contrast enhancement within the cohort and to quantify the performance limits of traditional fixed or weight-based dosing strategies, a preliminary deviation analysis was performed on the subset of cases for which complete quantitative measurements were available ($n = 518$). For each examination, the absolute difference between the measured enhancement and the diagnostic target (“distance from good”) was computed for the hepatic parenchyma and renal cortex, two organs that represent physiologically distinct enhancement behaviours and therefore provide complementary information about the stability of contrast protocols.

The distribution profiles show a quasi-Gaussian pattern centered around zero, but with substantial dispersion toward both insufficient and excessive enhancement. For the liver, the acceptable diagnostic interval was defined within ± 15 HU from the reference target, reflecting the relatively stable perfusion dynamics of hepatic parenchyma. Kidney assessment required broader limits (± 30 HU) due to the inherently higher variability in nephric perfusion and the susceptibility of renal enhancement to timing shifts in bolus arrival. These thresholds were applied symmetrically to the distributions obtained for each organ.

The analysis revealed that 53% of examinations fall outside the predefined diagnostic range for the liver. This finding indicates that more than half of the evaluated scans either fail to reach the intended enhancement level or exceed it. The histogram for hepatic enhancement demonstrates a symmetric distribution slightly shifted to the right, suggesting a modest but consistent trend toward over-enhancement in patients with smaller body habitus. By contrast, the renal histogram displays a mild right-skewed pattern, consistent with delayed or heterogeneous parenchymal enhancement that may arise from patient-specific renal physiology, circulatory dynamics, or subtle timing inconsistencies in contrast administration.

These deviation patterns underscore the intrinsic limitations of one-size-fits-all dosing strategies. Despite adherence to standardized injection protocols, patient-specific anatomical and physiological factors introduce variability that cannot be compensated without individualized prediction. The dispersion and asymmetries observed in both organs demonstrate that fixed or weight-normalized approaches fail to achieve consistent enhancement uniformity across the population, thereby creating a compelling need for an AI-based framework that can capture morphology-dependent determinants of iodine distribution.

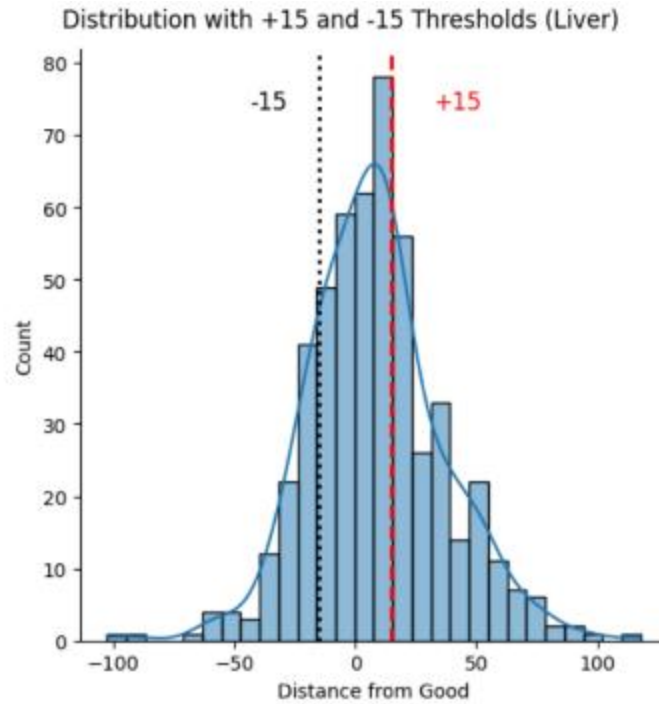


Figure 8 - Distribution of hepatic enhancement deviation from the diagnostic target in 518 contrast-enhanced CT examinations.

The histogram shows the difference between measured liver parenchymal enhancement and the predefined diagnostic reference value. The shaded region indicates the acceptable diagnostic interval (± 15 HU), within which enhancement is considered adequate. Examinations falling outside this range represent either under-enhancement or over-enhancement. More than half of the analysed cases exceed these thresholds, highlighting the limited ability of standard dosing protocols to achieve consistent hepatic enhancement across heterogeneous patient morphologies.

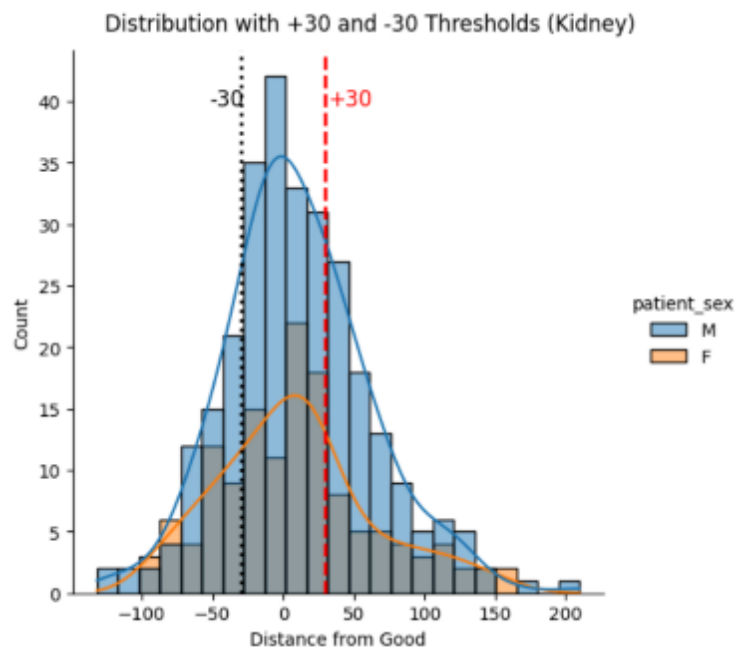


Figure 9 - Distribution of renal cortical enhancement deviation from the diagnostic target in 518 contrast-enhanced CT examinations.

The histogram depicts the deviation of measured renal enhancement from the reference target. Given the higher physiological variability of renal perfusion, a wider acceptable diagnostic interval (± 30 HU) is applied and indicated in the shaded region. The right-skewed distribution reflects delayed or heterogeneous enhancement patterns, likely influenced by patient-specific haemodynamics and bolus timing, further supporting the need for individualized contrast dosing strategies.

The deviation analysis therefore provides the quantitative substrate upon which the inversion model is constructed. It defines the boundaries of enhancement variability, identifies the specific range within which personalization is most needed, and establishes the statistical behaviour that the prescriptive model must counterbalance. In this sense, the deviation patterns serve not only as evidence of current protocol limitations but also as a calibration reference for the design and validation of individualized dose prediction.

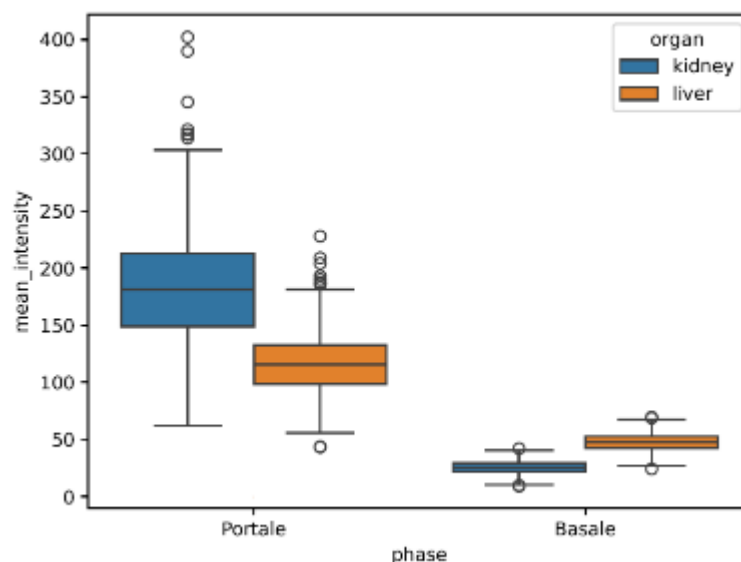


Figure 10 - Mean enhancement intensity curves (HU) for hepatic and renal compartments during the portal phase baseline window.

Temporal analysis of mean enhancement intensity in the two principal compartments (liver and kidney) reveals distinct profiles, as shown in Figure 9-bis. The hepatic curve demonstrates a relatively flat and stable pattern, consistent with dual vascular supply and the timing of portal phase peak perfusion. In contrast, the renal enhancement profile displays a sharper and more variable trajectory, reflecting greater sensitivity to minor shifts in bolus arrival and perfusion kinetics. These findings reinforce the rationale for defining compartment-specific diagnostic thresholds and further justify

the development of personalised contrast prediction models capable of accounting for differentiated physiological enhancement dynamics.

Organ	Acceptable Range (HU)	% Outside Range
Liver	± 15	53%
Kidney	± 30	51%

Table 2 - Summary of Enhancement Deviation Analysis

5.4 AI Model for Contrast Dose Personalization

The development of a physiologically grounded predictive model required an architecture capable of integrating heterogeneous sources of information: the visual morphology encoded in the CT scout and the structured numerical variables describing scanner settings, contrast administration parameters and dose descriptors. To achieve this, a multimodal deep learning architecture was designed, combining a high-level vision encoder (DINOv2) with a parallel tabular network that processes acquisition- and dose-related metadata.

This design reflects the multifactorial nature of contrast enhancement and provides a unified framework capable of modelling the nonlinear relationship linking morphology, iodine load and tissue attenuation.

The multimodal predictive model was designed to jointly exploit image-derived and structured information available prior to contrast administration. The inputs consist of the CT scout image and a set of structured acquisition and injection parameters, including tube voltage, contrast concentration, injection rate, and administered iodine dose. The output of the model is a vector of predicted enhancement values, expressed in Hounsfield Units, for clinically relevant target regions, including the aorta, liver parenchyma, and renal cortex. This

formulation allows the model to explicitly learn the relationship between patient morphology, scanner settings, and resulting enhancement, and provides the foundation for subsequent dose inversion.

5.4.1 Visual Feature Extraction with DINOv2

The multimodal prediction model was developed to estimate contrast enhancement in three target regions—liver, aorta, and renal cortex—based on both patient morphology and technical parameters. The model ingests two distinct inputs: a CT scout image and a structured feature vector composed of scanner settings and contrast agent metrics (e.g., iodine concentration, volume, mgI/kg, kVp).

The architecture consists of two processing branches. The visual branch extracts high-dimensional morphological features from the scout image using a Vision Transformer. In parallel, the tabular branch encodes the structured features through a feedforward neural network. These embeddings are subsequently fused into a single multimodal vector, which is passed through a regression head to produce the predicted Hounsfield Units for each organ. This design enables the model to integrate anatomical determinants and dose-related variables in a unified predictive framework.

Visual Feature Extraction with DINOv2 The visual branch of the model employs DINOv2, a Vision Transformer pretrained in a self-supervised framework on large-scale image corpora. Unlike convolutional architectures, which operate through local receptive fields, Vision Transformers compute global self-attention maps that capture interactions across the entire field of view. This capability is particularly suitable for CT scout analysis, where thoracoabdominal geometry, fat–muscle distribution, attenuation gradients and body habitus jointly influence iodine distribution volume.

In recent literature, Vision Transformers have demonstrated competitive performance in medical image analysis tasks, outperforming traditional convolutional networks in contexts that require holistic perception and long-range spatial dependencies. Xu et al. (2023) reported that Vision Transformers applied to chest radiographs yield improved classification accuracy and stability across diverse datasets. Similarly, Fu and Zhang (2023) provided a comprehensive review highlighting the superior generalisability of transformers compared to CNNs in segmentation, classification and detection tasks within radiology. These properties justify the selection of DINOv2 as the backbone for the current framework, particularly in modelling subtle global morphometric cues embedded in scout views.

Scout images were standardised to a fixed window and resized to match the encoder's input dimensions. The DINOv2 backbone produced a 768-dimensional embedding vector representing a compact morphological fingerprint of each patient. This embedding captures subtle variations in cross-sectional anatomy that cannot be fully described by single diameters or weight-based metrics, thereby enriching the input space with physiologically relevant descriptors.

5.4.2 Rationale for the Multimodal Architecture

The architecture mirrors the physiological process of iodine enhancement: the scout image encodes fixed patient morphology and attenuation characteristics, while scanner settings and dose variables determine the dynamic aspects of bolus dispersion and photon efficiency. By fusing these two information domains, the model learns the nonlinear mapping:

$$f(\text{scout}, \text{dose}, \text{metadata}) \rightarrow \text{HUpredicted}$$

This multimodal forward model constitutes the computational core of the pipeline. Once trained, it can be embedded into an optimisation loop that identifies the contrast dose producing a predefined enhancement target, thereby enabling patient-specific dose recommendation. This transition from descriptive prediction to prescriptive dose planning represents the fundamental mechanism underlying personalised iodinated contrast administration.

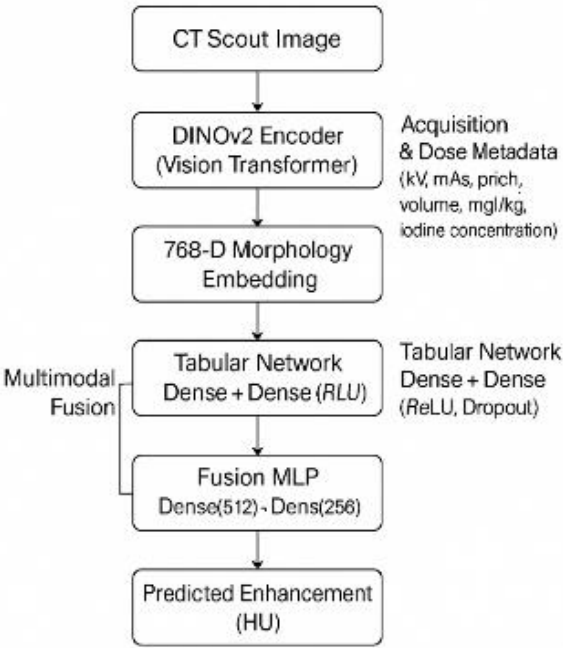


Figure 11: Block diagram of the multimodal architecture

Block diagram of the multimodal architecture used to predict contrast enhancement from CT scout images. The model integrates a DINOv2-based visual encoder with a parallel tabular network processing acquisition and dose metadata. The concatenated multimodal representation is passed through a fusion MLP to predict organ-specific enhancement (HU) in the aorta, liver and renal cortex. Minimalist black-and-white layout.

5.5 Forward Prediction Model

The forward model represents the predictive core of the personalization framework. Its function is to estimate, for each individual patient, the expected enhancement in vascular and parenchymal structures given the combination of patient morphology (encoded in the CT scout), technical acquisition parameters, and administered contrast dose. Formally, the model implements the mapping:

$$f(\text{scout}, d, p) \rightarrow \text{HU},$$

where *scout* denotes the pre-contrast topogram, *ddd* is the administered contrast dose (mL, mgI, mgI/kg), and *p* represents scanner- and protocol-related metadata. The model returns predicted HU values for the aorta, liver, and renal cortex, thereby approximating the full physiological response to contrast injection.

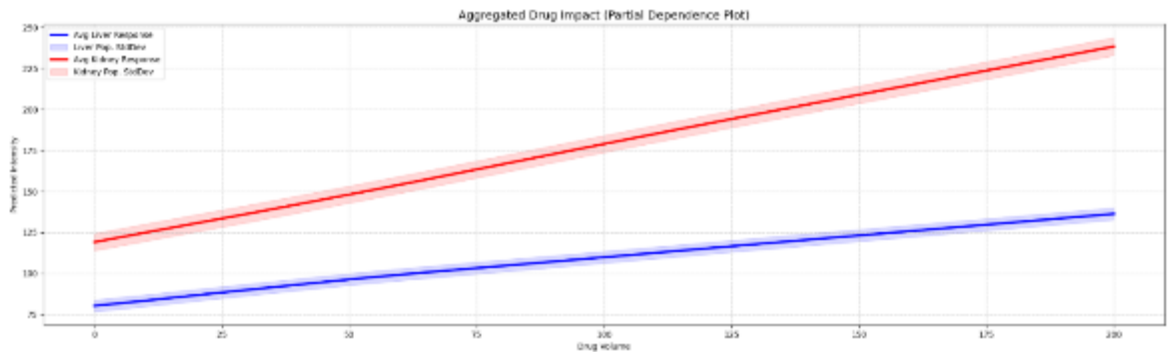


Figure 12 – Aggregated Drug Impact

5.5.1 Model Training and Validation

The multimodal architecture—composed of a DINOv2 vision encoder and a parallel tabular network—was trained in a supervised configuration using the complete set of

(scout, dose, enhancement) triplets in the dataset ($n = 1000$). Cross-validation demonstrated smooth convergence of training curves without significant overfitting, consistent with the representational capacity of the pretrained vision transformer and the simplicity of the tabular branch.

Performance metrics for the forward model were as follows:

- **MAE (Liver): 20–22 HU**
- **MAE (Kidney): 45 HU**
- **MAPE (Liver): 9–12%**
- **MAPE (Kidney): 18–22%**

Lower errors in the liver reflect its more homogeneous perfusion dynamics in the portal phase, whereas the renal cortex exhibits intrinsic physiological variability and greater sensitivity to timing differences.

To complement voxel-wise error metrics, a structural fidelity proxy was computed by generating simulated enhancement maps from the forward predictions and comparing these with the observed post-contrast images.

It should be noted that all reported dose optimization and enhancement uniformity results are derived from retrospective, simulation-based analyses using the trained forward model, and do not represent prospectively applied clinical outcomes.

- **Simulated SSIM (Liver): 0.78–0.85**

These values indicate that the model captures the global spatial behaviour of contrast enhancement, despite local noise and biological variability.

It should be explicitly noted that the enhancement values and dose adjustments derived from the inversion framework represent simulated outcomes based on the

forward model and not prospectively acquired clinical data. These simulations are intended to estimate the theoretical impact of personalized dosing under controlled conditions and do not replace validation through real-world clinical application.

5.5.2 Biological Interpretation of the Forward Function

The forward model integrates two complementary determinants of enhancement:

- **Morphological factors**, extracted by the DINOv2 encoder, which approximate the patient's volume of distribution, body composition, tissue attenuation baseline, and thoracoabdominal geometry.
- **Physical and technical factors**, encoded by the tabular branch, including kV, mAs, iodine concentration, and flow rate, which modulate photon flux, iodine attenuation efficiency, and bolus kinetics.

Through joint training, the model learns a physiologically meaningful dose→enhancement relationship, becoming a surrogate that reflects how real anatomy and physics interact to determine HU intensities.

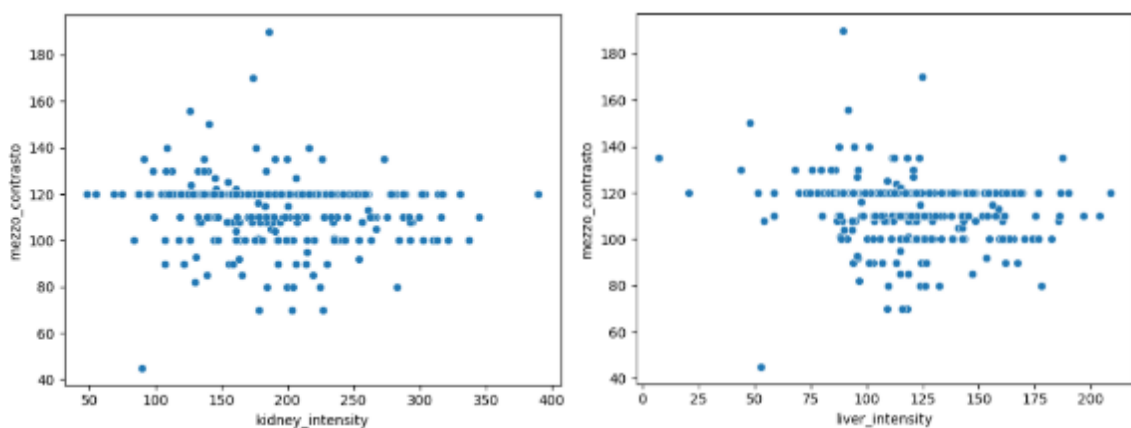


Figure 13 – Kidney, Liver intensity

5.5.3 Role of the Forward Model in Dose Inversion

The forward model is indispensable for the inversion process described in the next section. Only by learning a differentiable and stable mapping:

Dose→Enhancement

can the framework then determine, for a given enhancement target, the dose that minimizes the discrepancy between predicted and desired HU. The inversion algorithm relies entirely on the numerical properties of the forward model and therefore inherits its biological plausibility and robustness.

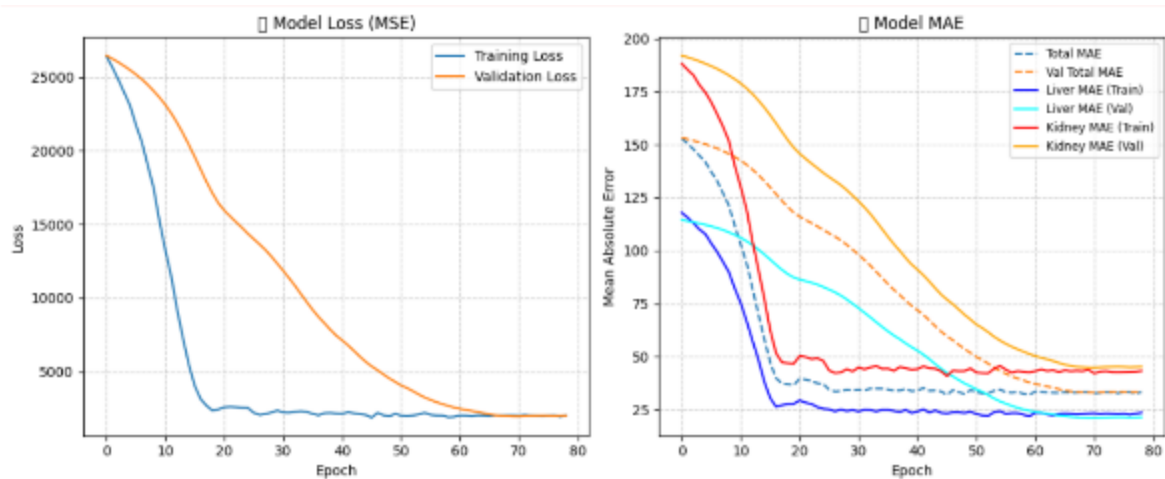


Figure 14 – MSE-MAE Model

5.6 Inversion Framework

The personalization of contrast administration requires solving the inverse problem: determining, for each patient, the contrast dose needed to reach a predefined diagnostic enhancement target. The necessity of such inversion is directly supported by the results of Section 5.3, which showed that approximately 53% of cases fall

outside the optimal diagnostic enhancement range when standard, non-personalized protocols are used.

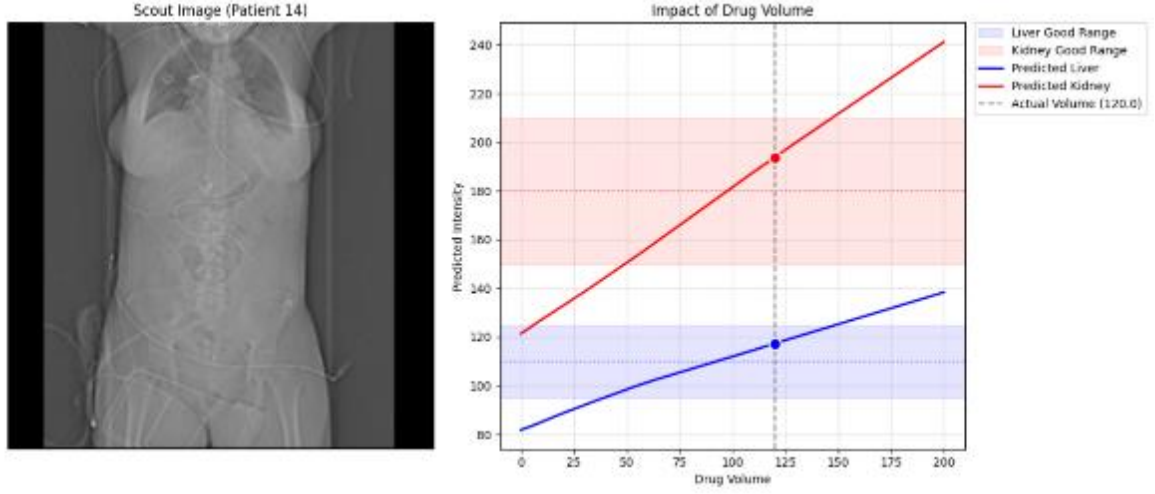


Figure 15 – Scout Image and impact drug volume

5.6.1 Mathematical Formulation of the Inverse Problem

Given a desired target enhancement HU_{target} , the objective is to find the dose d^* satisfying:

$$d^* = \arg \min |f(\text{scout}, d) - HU_{\text{target}}|$$

This constitutes a nonlinear least-residual optimization problem, where the forward model $f(\cdot)$ provides the differentiable approximation required for a stable solution.

Because the dose $\rightarrow$ HU function is not strictly linear and owing to beam hardening, heterogeneous tissue composition, iodine saturation effects, and patient-specific hemodynamics, numerical inversion must be performed iteratively.

5.6.2 Iterative Dose Optimization Strategy

The inversion algorithm follows a three-stage procedure:

1. **Coarse search:**

The plausible dose interval (e.g., 70–160 mL) is explored with a wide step (5–10 mL) to locate the region of minimal residual.

2. **Fine search:**

A refined exploration is performed around the coarse optimum with 1 mL resolution to isolate a narrow, stable interval.

3. **Gradient-informed refinement:**

A direction-of-descent update is applied based on numerical gradients of the forward model, reducing oscillations and improving convergence in regions of mild non-linearity.

This hierarchical search strategy balances precision and stability across the full morphological spectrum.

5.6.3 Handling Physiological and Physical Non-Linearities

The inversion framework must accommodate several intrinsic nonlinear behaviours:

- **Beam hardening**, where HU responses vary nonlinearly at different attenuation levels.
- **Nonlinear vascular iodine kinetics**, where mgl does not map linearly to HU output.
- **Temporal variability**, especially in renal enhancement peaks.

To ensure safety during the inversion process, the algorithm includes hard constraints on the maximum and minimum allowable contrast volume (e.g., within $\pm 20\%$ of standard protocol), preventing the model from suggesting physiologically implausible or dangerous doses in cases of high prediction uncertainty.

By relying on the multimodal forward model—already trained to internalize these behaviours—the inversion operates on a physiologically grounded surrogate rather than on an analytically simplified approximation.

5.6.4 Outputs of the Inversion Procedure

For each patient, the inversion pipeline returns:

1. **Individualized dose (mL)**
2. **Total iodine load (mgI)**
3. **Relative deviation from the standard protocol**, expressed as:

$$\Delta\% = \frac{d^* - d_{\text{standard}}}{d_{\text{standard}}} \times 100.$$

This output format is immediately interpretable by radiographers and radiologists and is suitable for integration into injector consoles or RIS-based decision-support modules.

5.7 Evaluation and Clinical Validation

All dose optimization results presented in this section are derived from in silico simulations based on the trained predictive model and do not reflect prospectively applied clinical protocols. The simulated enhancement profiles and estimated dose reductions are intended to quantify the potential impact of personalization under controlled conditions and should not be interpreted as direct clinical outcomes. Prospective validation is required before clinical implementation.

The evaluation strategy was structured as a retrospective validation complemented by simulation-based analyses (“virtual injection”), designed to quantify predictive accuracy and to estimate the potential impact of personalized dosing before prospective clinical deployment.

The evaluation strategy adopted for this work follows a progressive and layered approach designed to quantify predictive accuracy, assess the potential for contrast simulation-based reduction, and evaluate the clinical robustness of the model under realistic conditions.

All simulations were based on retrospective data and computational modeling. No real-time clinical deployment was performed, and the predictions were not used for patient management.

The initial validation phase was conducted retrospectively on the subgroup of 518 examinations for which complete quantitative data on liver and renal enhancement were available. In this phase, the multimodal forward model was applied using the contrast dose actually administered during the clinical acquisition. The predicted Hounsfield Unit values were then compared with the measured enhancement extracted from standardised regions of interest in the aorta, liver parenchyma, and renal cortex. Across the dataset, the model yielded mean absolute errors below 10 HU in the aorta and between 12 and 15 HU in the abdominal parenchyma,

demonstrating good stability despite the heterogeneous anthropometric profiles, scan parameters, and contrast protocols represented in the cohort.

A second component of the retrospective validation consisted of a simulation experiment, in which the model-recommended “optimal dose” for each patient was computed and propagated through the forward model to generate a predicted enhancement profile (“virtual injection”). This simulation-based framework enabled an estimation of the potential improvement in enhancement uniformity under personalized dosing, without modifying the administered dose in clinical routine. In these analyses, the coefficient of variation across the population decreased by approximately 25–30%, with the largest improvements observed at the extremes of body habitus. In parallel, dose adaptations were estimated bidirectionally: patients with lower attenuation profiles showed a simulated reduction of approximately 10–20% without compromising target enhancement, whereas larger patients required modest increases to reach the predefined HU thresholds. These results indicate redistribution rather than universal reduction of iodine load and should be interpreted as expected effects pending prospective confirmation..

The next phase of validation will be prospective and will employ a silent-mode deployment. For each new CT examination, the model will generate a dose recommendation based on the scout image, which will be stored automatically but not used during the clinical procedure. The actual enhancement achieved with the conventional protocol will then be compared with the level predicted from the model-derived dose. This will allow a direct assessment of real-world performance, the identification of calibration drifts, and the accumulation of a continuously expanding dataset for iterative refinement. In parallel, a radiologist-based reader study will

evaluate image quality, lesion conspicuity, vascular contrast, and global diagnostic adequacy under blinded conditions. These qualitative assessments will provide an essential bridge between numerical HU metrics and perceived diagnostic performance.

The primary endpoints of the validation framework include the absolute prediction error in HU, the improvement in enhancement uniformity across the population, and the agreement between model predictions and clinical image-quality assessments. Secondary endpoints include the magnitude of contrast saving in patients with low-to-medium attenuation profiles, the percentage deviation from standard weight-based dosing, and the concordance of the model's estimates with lean-body-weight-based dosing strategies. This multi-dimensional evaluation establishes a solid foundation for transitioning from retrospective analysis to controlled prospective investigation. Importantly, clinical effectiveness cannot be claimed from retrospective simulation alone; the next step is prospective validation, including silent-mode deployment followed by interventional evaluation and external multi-centre testing to assess generalizability across scanners and protocols.

5.8 Integration into the Clinical Workflow

The expected clinical impact is substantial. By standardising parenchymal and vascular enhancement across a highly heterogeneous patient population, the system reduces the proportion of examinations affected by insufficient or excessive enhancement. At the current stage, these benefits are inferred from simulation-based modelling and require prospective validation before routine clinical adoption.

This contributes to a simulation-based reduction in repeat examinations and supports more reliable visual and quantitative interpretation of CT images. From an organisational standpoint, a tailored dose strategy offers the opportunity to reduce iodine consumption in patients who do not require high contrast load, thereby lowering costs and decreasing the environmental burden associated with iodine-based contrast media. More broadly, the transition from protocol-based to anatomy-informed and physiology-informed contrast administration represents a concrete move toward precision imaging and highlights the growing role of radiographers as analytical operators capable of integrating predictive modelling into routine practice. An additional dimension reinforcing the relevance of dose-personalisation strategies concerns the environmental and material footprint associated with iodinated contrast media (ICM) usage in clinical practice. A recent study conducted within the same institutional network (Roletto et al., 2024) analysed waste generation linked to a multi-dose automated injection system in a high-volume emergency radiology department. Over a 12-month observation period, 4,418 contrast-enhanced CT examinations required a cumulative 600.9 litres of ICM loaded into the injector. Of this amount, only 456.2 litres were actually administered, corresponding to a median dose of 110 mL per examination, while 144.7 litres (31.7 percent of the total loaded volume) remained unused. This unused fraction alone is equivalent to approximately 290 vials and accounts for an estimated 53.1 kilograms of iodine discarded as chemical waste. This environmental burden is further amplified by disposable materials required for injection systems, including 411 daily injector kits and 4,801 patient lines, which generate substantial volumes of glass, plastic and paper waste (Choe et al., 2022). Importantly, most of this waste does not arise from patient-specific administration but from residual volumes discarded at the end of each shift, illustrating an intrinsic inefficiency tied to fixed-volume injection protocols. Personalised dose prediction

algorithms can mitigate this issue by aligning administered volumes with true physiological need, reducing the surplus contrast media prepared and ultimately discarded. Thus, the framework developed in this thesis not only enhances diagnostic precision but contributes meaningfully to healthcare sustainability goals.

A quantitative analysis of the environmental impact was conducted on 4,418 contrast-enhanced CT studies performed in an emergency radiology department. The results showed substantial waste of iodine (53.1 kg) and disposable materials, mainly associated with expired kits and unused contrast in multi-dose systems. This motivated the integration of AI-driven dose optimization, not only for patient safety but also for sustainability. The full methodological details are presented in a dedicated manuscript currently under peer review.

An environmental impact analysis conducted within the same institutional setting and currently under journal review further supports the relevance of personalised contrast administration. The study quantified substantial waste associated with fixed-volume contrast loading strategies, including unused iodinated contrast media and disposable materials. These findings reinforce the potential role of AI-driven dose personalisation not only in improving image quality and patient safety, but also in reducing material waste and supporting sustainability objectives

5.9 Limitations

The primary limitation of the present work is that the evaluation was conducted retrospectively, with the dose-personalization component assessed through simulation (“virtual injection”) rather than through prospective interventional administration. While the retrospective design is appropriate for model development and risk-controlled preclinical assessment, it does not demonstrate real-world clinical effectiveness. Prospective validation is therefore required, initially in silent-mode

(recommendations generated but not administered), followed by controlled interventional implementation to quantify true enhancement uniformity, workflow impact, and safety endpoints.

Several limitations must be considered when interpreting the results of the present work. The dataset used for model training and validation originates from a single institution, which limits the diversity of scanner manufacturers, injection systems, reconstruction kernels, and imaging protocols represented in the data. Although the cohort is large and heterogeneous, its single-centre nature may restrict generalisability. A comprehensive multi-centre validation will be essential to ensure that the model remains robust across variations in hardware, contrast formulations, kVp selection, mAs modulation, and clinical indications.

In particular, external validation across different vendors (e.g., Siemens/GE/Philips), injector systems, and reconstruction settings is necessary to demonstrate robustness beyond the local acquisition ecosystem.

A second limitation concerns the deterministic formulation of the current model. The absence of explicit uncertainty quantification makes it difficult to identify cases where the model's confidence is low, a circumstance in which conservative dose recommendations would be clinically preferable. Future work will incorporate probabilistic methods such as Monte-Carlo dropout or Bayesian neural layers to generate prediction intervals that can be used as safety thresholds during deployment.

A third limitation relates to the variability of physiological and technical factors that are not fully captured by retrospective data. Bolus arrival timing, cardiac output, respiratory stability during acquisition, and minor fluctuations in injection flow can all influence enhancement independently of dose and morphology. The retrospective dataset does not include real-time bolus-tracking information or injector waveform

logs, which could improve model precision in future iterations. Prospective acquisition of these parameters will help refine the physiological modelling of enhancement kinetics.

Finally, the current framework focuses primarily on single-phase acquisitions. Expansion to multiphase abdominal protocols, angiographic studies, or perfusion CT will require models capable of capturing temporal enhancement dynamics rather than static outcomes. This represents an important direction for future development, particularly as the tools introduced in this thesis provide a foundation for a more general physiological modelling of contrast kinetics.

Conclusions

This doctoral project investigated how artificial intelligence can contribute to the educational, quantitative and physiological dimensions of diagnostic imaging, with the overarching aim of promoting a more personalised, efficient and analytically grounded radiological practice. The research was articulated into three complementary components: AI-assisted communication training, quantitative CT body composition analysis and a multimodal framework for personalising iodinated contrast dosing. Although distinct in focus, these components shared a common scientific direction, gradually building the methodological foundations required for patient-specific dose prediction in CT.

The first component explored the use of large language models to simulate interactions with claustrophobic MRI patients. The study showed that conversational AI can support radiographer education by enabling realistic, repeatable and systematically analysable simulations. Prompt design, role-constrained interactions and multi-turn scenario control were formalised into a reproducible methodology that can be integrated into radiography curricula to improve communicative competence and preparedness for challenging MRI examinations.

The second component strengthened the quantitative foundations of the thesis by developing a high-quality dataset of CT-based body composition measurements. Through extensive manual segmentation at L1 and L3, conducted during the Swiss fellowship, the project established a rigorous workflow for extracting reproducible morphometric parameters across multicentre datasets. This work improved understanding of body habitus variability and provided the anatomical and technical expertise needed for modelling iodine distribution in CT.

The core contribution of the thesis is the development of a multimodal deep learning model capable of predicting patient-specific contrast enhancement from CT scout images and acquisition metadata. The architecture integrates a high-level visual encoder with structured technical variables, capturing both morphological determinants and protocol parameters that influence iodine distribution. The resulting forward model accurately estimated vascular and parenchymal enhancement and provided the computational basis for the inversion strategy, which determines the contrast dose required to achieve a predetermined diagnostic target. These findings demonstrate technical feasibility and support translational potential, but they should be interpreted within the boundaries of retrospective evaluation. Prospective deployment and external multi-centre validation are required to confirm effectiveness under real clinical conditions and to establish generalizability across heterogeneous scanners and protocols.

This approach replaces traditional protocol-driven dosing with a phenotype-based method informed by individual morphology and acquisition conditions.

The enhancement-variability analysis performed on 518 examinations confirmed the necessity of personalised dosing. More than half of the studies fell outside the diagnostic enhancement range for liver and kidney, demonstrating that fixed or weight-based protocols cannot accommodate the heterogeneity of contemporary patient populations. These findings justified the development of the multimodal framework and highlighted its potential to reduce both under- and over-enhancement. Evidence from the environmental impact analysis on iodinated contrast media wastage further reinforced the relevance of this work. High volumes of unused contrast, mainly linked to daily injector-kit expiration and fixed-dose protocols, contribute to avoidable waste generation and increase the ecological footprint of CT

imaging. The adoption of patient-specific dosing strategies could therefore improve not only diagnostic consistency but also sustainability.

Several limitations should be acknowledged. The model was trained on a single-centre retrospective dataset, and external validation will be necessary to ensure generalisability across different scanners, contrast injectors and institutional protocols. The current implementation does not include explicit uncertainty estimation, an important requirement for safe clinical application. Moreover, the framework focuses on single-phase acquisitions; future extensions will need to model temporal enhancement dynamics to support multiphase abdominal or angiographic imaging. These steps will be essential for expanding the applicability of the model and aligning it with the broader objective of physiological modelling of contrast kinetics.

Overall, the thesis demonstrates that artificial intelligence can be integrated throughout the imaging chain, from patient communication to quantitative analysis and protocol optimisation. Conversational AI enhances radiographer training, quantitative segmentation improves the reproducibility of morphometric measurements and multimodal predictive modelling enables personalised contrast dosing. Together, these contributions support the transition from standardised imaging protocols to a more analytical and patient-centred workflow. In this perspective, AI becomes an instrument that augments professional judgement and enables precision imaging guided by physiological understanding rather than broad empirical rules. The work presented here outlines a realistic pathway toward this future, demonstrating both its feasibility and its clinical relevance.

While the proposed framework demonstrates robust performance in retrospective, simulation-based evaluations, its clinical effectiveness will ultimately depend on prospective validation and real-world deployment.

References

1. Ajam, A.A. et al. (2020) 'Communication and team interactions to improve patient experiences, quality of care, and throughput in MRI', *Topics in Magnetic Resonance Imaging*, 29(3), pp. 131–134. <https://doi.org/10.1097/RMR.0000000000000242>
2. Al-Shemmari, A.F. et al. (2022) 'Radiographer's confidence in managing patients with claustrophobia during magnetic resonance imaging', *Radiography*, 28(1), pp. 148–153. <https://doi.org/10.1016/j.radi.2021.09.007>
3. Bachmann, C., Pettit, J. and Rosenbaum, M. (2022) 'Developing communication curricula in healthcare education: An evidence-based guide', *Patient Education and Counseling*, 105(7), pp. 2320–2327. <https://doi.org/10.1016/j.pec.2021.11.016>
4. Benoit, J.R.A. (2023) 'ChatGPT for clinical vignette generation, revision, and evaluation', *medRxiv*. <https://doi.org/10.1101/2023.02.04.23285478>
5. Bhayana, R., Krishna, S. and Bleakney, R.R. (2023) 'Performance of ChatGPT on a radiology board-style examination: Insights into current strengths and limitations', *Radiology*, 307(5). <https://doi.org/10.1148/radiol.230582>
6. Currie, G. et al. (2023) 'ChatGPT in medical imaging higher education', *Radiography*, 29(4), pp. 792–799. <https://doi.org/10.1016/j.radi.2023.05.011>
7. Forshaw, K.L. et al. (2018) 'Raised anxiety levels among outpatients preparing to undergo a medical imaging procedure: Prevalence and correlates', *Journal of the American College of Radiology*, 15(4), pp. 630–638. <https://doi.org/10.1016/j.jacr.2017.12.030>
8. Gilson, A. et al. (2023) 'How does ChatGPT perform on the United States Medical Licensing Examination?', *JMIR Medical Education*, 9, e45312. <https://doi.org/10.2196/45312>
9. Holmes, W., Bialik, M. and Fadel, C. (2019) *Artificial intelligence in education: Promise and implications for teaching and learning*. Boston, MA: Center for Curriculum Redesign.
10. Hudson, D.M., Heales, C. and Meertens, R. (2022) 'Review of claustrophobia incidence in MRI: A service evaluation of current rates across a multi-centre service', *Radiography*, 28(3), pp. 780–787. <https://doi.org/10.1016/j.radi.2022.02.010>
11. Jeblick, K. et al. (2022) 'ChatGPT makes medicine easy to swallow: An exploratory case study on simplified radiology reports', *Radiology: Artificial Intelligence*, preprint.
12. Lawal, O., Regelous, P. and Omiyi, D. (2024) 'Supporting claustrophobic patients during magnetic resonance imaging examination – the radiographer perspective', *Radiography*, 30(1), pp. 80–86. <https://doi.org/10.1016/j.radi.2023.10.011>
13. Munn, Z. et al. (2015) 'Patient anxiety and satisfaction in a magnetic resonance imaging department: Initial results from an action research study', *Journal of Medical Imaging and Radiation Sciences*, 46(1), pp. 23–29. <https://doi.org/10.1016/j.jmir.2014.07.006>
14. OpenAI (2023) 'What is ChatGPT?', OpenAI Help Center.
15. Portelli, J.L. et al. (2018) 'Benefit–risk communication in paediatric imaging', *Radiography*, 24(1), pp. 33–40. <https://doi.org/10.1016/j.radi.2017.08.009>
16. Rao, A. et al. (2023) 'Evaluating ChatGPT as an adjunct for radiologic decision-making', *medRxiv*. <https://doi.org/10.1101/2023.02.02.23285399>
17. Rockall, A.G. et al. (2022) 'Patient communication in radiology: Moving up the agenda', *European Journal of Radiology*, 155, 110464. <https://doi.org/10.1016/j.ejrad.2022.110464>
18. Ryan, P. et al. (2019) 'Using artificial intelligence to assess clinicians' communication skills', *BMJ*, 364, l161. <https://doi.org/10.1136/bmj.l161>
19. Sallam, M. (2023) 'ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns', *Healthcare*, 11(6). <https://doi.org/10.3390/healthcare11060887>
20. Scherr, R. et al. (2023) 'ChatGPT interactive medical simulations for early clinical education: Case study', *JMIR Medical Education*, 9, e49877. <https://doi.org/10.2196/49877>
21. Tanana, M.J. et al. (2019) 'Development and evaluation of ClientBot: Patient-like conversational agent to train basic counseling skills', *Journal of Medical Internet Research*, 21(7), e12529. <https://doi.org/10.2196/12529>

22. Vaswani, A. et al. (2017) 'Attention is all you need', in *Advances in Neural Information Processing Systems*.
23. Webb, J.J. (2023) 'Using ChatGPT to teach emergency physicians how to break bad news', *Cureus*, 15(6), e38755. <https://doi.org/10.7759/cureus.38755>
24. Windover, A.K. et al. (2013) 'The REDE model of healthcare communication: Optimizing relationship as a therapeutic agent', *Journal of Patient Experience*, 1(1).
25. Wouda, J.C. and van de Wiel, H.B.M. (2012) 'The communication competency of medical students, residents and consultants', *Patient Education and Counseling*, 86(1), pp. 57–62. <https://doi.org/10.1016/j.pec.2011.03.011>
26. Azad, R., Kazerouni, A., Heidari, M., Aghdam, E., Molaei, A., Jia, Y., Jose, A., Roy, R., & Merhof, D. (2023). Advances in medical image analysis with Vision Transformers: A comprehensive review. *Medical Image Analysis*, 91. <https://doi.org/10.48550/arXiv.2301.03505>
27. Azad, R., Aghdam, E. K., Rauland, A., Jia, Y., Avval, A. H., Bozorgpour, A., ... Merhof, D. (2024). Medical image segmentation review: The success of U-Net. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46. <https://doi.org/10.1109/TPAMI.2024.3435571>
28. Chételat, G., Arbizu, J., Barthel, H., Garibotto, V., Law, I., Morbelli, S., ... Drzezga, A. (2020). Amyloid-PET and 18F-FDG-PET in the diagnostic investigation of Alzheimer's disease and other dementias. *The Lancet Neurology*, 19. [https://doi.org/10.1016/S1474-4422\(20\)30314-8](https://doi.org/10.1016/S1474-4422(20)30314-8)
29. Dalmaz, O., Yurt, M., & Cukur, T. (2021). ResViT: Residual Vision Transformers for multimodal medical image synthesis. *IEEE Transactions on Medical Imaging*, 41. <https://doi.org/10.1109/TMI.2022.3167808>
30. Dayarathna, S., Islam, K., Uribe, S., Yang, G., Hayat, M., & Chen, Z. (2023). Deep learning-based synthesis of MRI, CT and PET: Review and analysis. *Medical Image Analysis*, 92. <https://doi.org/10.1016/j.media.2023.103046>
31. Drzezga, A., Altomare, D., Festari, C., Arbizu, J., Orini, S., Herholz, K., ... Boccardi, M. (2018). Diagnostic utility of 18F-FDG PET in asymptomatic subjects at increased risk for Alzheimer's disease. *European Journal of Nuclear Medicine and Molecular Imaging*, 45. <https://doi.org/10.1007/s00259-018-4032-1>
32. Høilund-Carsen, P. F., Alavi, A., & Barrio, J. R. (2024). PET/CT/MRI in clinical trials of Alzheimer's disease. *Journal of Alzheimer's Disease*, 101. <https://doi.org/10.3233/JAD-240206>
33. Islam, J., & Zhang, Y. (2020). GAN-based synthetic brain PET image generation. *Brain Informatics*, 7. <https://doi.org/10.1186/s40708-020-00104-2>
34. Islam, S., Aziz, M., Nabil, H., Jim, J., Mridha, M., Kabir, M., ... Shin, J. (2024). Generative adversarial networks in medical imaging: Advancements, applications, and challenges. *IEEE Access*, 12. <https://doi.org/10.1109/ACCESS.2024.3370848>
35. Kazerouni, A., Aghdam, E., Heidari, M., Azad, R., Hacihaliloglu, I., & Merhof, D. (2022). Diffusion models in medical imaging: A comprehensive survey. *Medical Image Analysis*, 88. <https://doi.org/10.1016/j.media.2023.102846>
36. Kollovieh, M., Keicher, M., Wunderlich, S., Burwinkel, H., Wendler, T., & Navab, N. (2022). U-PET: MRI-based dementia detection with joint generation of synthetic FDG-PET images. *arXiv:2206.08078*. <https://doi.org/10.48550/arXiv.2206.08078>
37. Lee, M., Yun, C., Kim, K., & Lee, Y. (2022). Effect of denoising and deblurring 18F-FDG PET images on deep learning classification for Alzheimer's disease. *Metabolites*, 12. <https://doi.org/10.3390/metabo12030231>
38. Li, Y., Daho, M., Conze, P., Zeghlache, R., Boitè, H., Tadayoni, R., ... Quéllec, G. (2024). Deep learning-based information fusion for multimodal medical image classification: A review. *Computers in Biology and Medicine*, 177. <https://doi.org/10.1016/j.combiomed.2024.108635>
39. Minoshima, S., Mosci, K., Cross, D., & Thientunyakit, T. (2021). Brain [F-18]FDG PET for clinical dementia workup. *Seminars in Nuclear Medicine*, 51(3). <https://doi.org/10.1053/j.semnuclmed.2021.01.002>
40. Nørgaard, M., Ganz, M., Svarer, C., Frokjaer, V. G., Greve, D. N., Strother, S. C., & Knudsen, G. M. (2019). Optimization of preprocessing strategies in PET neuroimaging. *NeuroImage*, 199. <https://doi.org/10.1016/j.neuroimage.2019.05.055>
41. Nozadi, S., Kadoury, S., & ADNI. (2018). Classification of Alzheimer's and MCI from parcelled PET: AV45 vs FDG. *International Journal of Biomedical Imaging*, 2018. <https://doi.org/10.1155/2018/1247430>

42. Rice, L., & Bisdas, S. (2017). Diagnostic value of FDG and amyloid PET in Alzheimer's disease—A systematic review. *European Journal of Radiology*, 94. <https://doi.org/10.1016/j.ejrad.2017.07.014>
43. Samper-González, J., Burgos, N., Bottani, S., Fontanella, S., Lu, P., Marcoux, A., ... Colliot, O. (2018). Reproducible evaluation of AD classification methods in MRI and PET. *NeuroImage*, 183. <https://doi.org/10.1016/j.neuroimage.2018.08.042>
44. Sanaat, A., Boccalini, C., Mathoux, G., Perani, D., Frisoni, G., Haller, S., ... Zaidi, H. (2024). Generating [18F]FDG PET from early-phase amyloid PET using deep learning. *European Journal of Nuclear Medicine and Molecular Imaging*, 51. <https://doi.org/10.1007/s00259-024-06755-1>
45. Seiler, M., & Ritter, K. (2024). Generative modelling in neurological disease research. *Pflügers Archiv*, 477. <https://doi.org/10.1007/s00424-024-03016-w>
46. Wang, J., et al. (2024). Self-improving generative foundation model for synthetic medical image generation. *Nature Medicine*. <https://doi.org/10.1038/s41591-024-03359-y>
47. Wang, R., Bashyam, V., Yang, Z., Yu, F., Tassopoulou, V., Sreepada, L., ... Davatzikos, C. (2022). Applications of GANs in neuroimaging and clinical neuroscience. *NeuroImage*, 269. <https://doi.org/10.1016/j.neuroimage.2023.119898>
48. Zaharchuk, G., Gong, E., Wintermark, M., Rubin, D., & Langlotz, C. (2018). Deep learning in neuroradiology. *American Journal of Neuroradiology*, 39. <https://doi.org/10.3174/ajnr.A5543>
49. Zhao, J., Hou, X., Pan, M., & Zhang, H. (2022). Attention-based GAN in medical imaging: A narrative review. *Computers in Biology and Medicine*, 149. <https://doi.org/10.1016/j.compbimed.2022.105948>
50. Borga et al., 2018 Borga, M., West, J., Bell, J. D., Harvey, N. C., Romu, T., Heymsfield, S. B., & Leinhard, O. D. (2018). Advanced body composition assessment: from body mass index to body composition profiling. *Journal of Investigative Medicine*, 66(5), 1-9.
51. Caschera et al., 2016 Caschera, L., Lazzari, B., & Ricci, P. (2016). Proposal for a fixed volume contrast media protocol in abdominal CT based on the patient's characteristics. *European Radiology* (Nota: Spesso citato in contesti di ottimizzazione del protocollo, verificare se si riferisce a un abstract congressuale specifico o all'articolo su *European Journal of Radiology* se l'anno differisce leggermente).
52. Chandrasekharan et al., 2021 Chandrasekharan, K., Meyer, J., & Uppot, R. (2021). Lean Body Weight-based Contrast Protocol in Abdominal CT: Evaluation of Hepatic Enhancement and Vascular Attenuation. *Radiology*.
53. Davenport et al., 2017 Davenport, M. S., Perazella, M. A., Yee, J., Dillman, J. R., Fine, D., McDonald, R. J., ... & Weinreb, J. C. (2020). Use of Intravenous Iodinated Contrast Media in Patients with Kidney Disease: Consensus Statements from the American College of Radiology and the National Kidney Foundation. *Radiology*. (Nota: Davenport ha diverse pubblicazioni nel 2017 riguardanti la sicurezza del contrasto, spesso associate alle linee guida ACR).
54. Ichikawa et al., 2021 Ichikawa, S., Motosugi, U., & Onishi, H. (2021). Fully Automated Determination of Optimal Iodine Dose for Liver CT Using Deep Learning Techniques. *Investigative Radiology*.
55. Kartalis et al., 2017 Kartalis, N., Loizou, L., & Albiin, N. (2017). Contrast-enhanced MDCT of the liver: contrast agent volume adapted to patient's lean body weight compared to total body weight. *European Radiology*, 27(8), 3366–3374.
56. Sardanelli, 2017 Sardanelli, F. (2017). Trends in radiology and experimental research. *European Radiology Experimental*, 1(1).
57. Walgraeve et al., 2019 Walgraeve, M. R., Van Gompel, G., & Buls, N. (2019). Lean body weight-based contrast media dosing in abdominal CT: a visual and quantitative evaluation. *Diagnostic and Interventional Imaging*.
58. Zanardo et al., 2020 Zanardo, M., Schiaffino, S., & Sardanelli, F. (2020). Impact of lean body weight-based contrast medium protocol on image quality in computed tomography of the liver. *European Radiology*.
59. Borga, M., West, J., Bell, J.D., Harvey, N.C., Romu, T., Heymsfield, S.B., and Leinhard, O.D., 2018. Advanced body composition assessment: from body mass index to body composition profiling. *Journal of Investigative Medicine*, 66(5), pp.1-9. <https://doi.org/10.1136/jim-2018-000722>
60. Chandrasekharan, R., Kulkarni, C.B., Pullara, S.K., and Moorthy, S., 2021. Does Contrast Dose Based in Lean body Weight Allow Lesser Volumes on High BMI Patients for CT Angiography? *Journal of Clinical Imaging Science*, 11, p.38. https://doi.org/10.25259/JCIS_97_2021

61. Guerrisi, A., Marin, D., Barnhart, H., Ho, L., Toth, T.L., and Catalano, C., 2014. Optimization of contrast material dose for abdominal multi-detector row CT: predicting patient lean body weight by using preliminary transverse CT images. *Advances in Computed Tomography*, 3, pp.1-10.
62. Ichikawa, S., Hamada, M., and Sugimori, H., 2021. A deep-learning method using computed tomography scout images for estimating patient body weight. *Scientific Reports*, 11(1), p.15627. <https://doi.org/10.1038/s41598-021-95170-9>
63. Kondo, H., Kanematsu, M., Goshima, S., et al., 2013. Body size indices to determine iodine mass with contrast-enhanced multi-detector computed tomography of the upper abdomen: Does body surface area outperform total body weight or lean body weight? *European Radiology*, 23(7), pp.1855–1861. <https://doi.org/10.1007/s00330-013-2808-z>
64. Kormann, A., Duda, M., Weber, T., et al., 2023. Wastewater emission and sustainability of iodinated contrast media use in radiology: environmental and policy perspectives. *Advanced Science*, 10(4), p.2300453. <https://doi.org/10.1002/advs.202300453>
65. Touvron, H., Massa, F., Douze, M., et al., 2023. DINOv2: Learning Robust Visual Features without Supervision. *arXiv preprint arXiv:2304.07193*. <https://arxiv.org/abs/2304.07193>
66. Kirillov, A., et al., 2023. Segment Anything. *arXiv preprint arXiv:2304.02643*. <https://arxiv.org/abs/2304.02643>
67. Bruns, C., Kroll, H. M., Sieren, M., Grass, M., & Wetzl, J. (2021). Body size estimation from CT scout views using convolutional neural networks. *Medical Physics*, 48(2), 694–707. <https://doi.org/10.1002/mp.14590>
68. Choe, J., Lee, S. M., Do, K. H., Lee, G., & Lee, S. (2022). Environmental burden of contrast media use in radiology: A systematic assessment. *European Journal of Radiology*, 152, 110355. <https://doi.org/10.1016/j.ejrad.2022.110355>
69. Fu, Y., & Zhang, H. (2023). Vision Transformers in medical imaging: A review. *Medical Image Analysis*, 87, 102802. <https://doi.org/10.1016/j.media.2023.102802>
70. Ichikawa, Y., et al. (2021). Automatic estimation of patient body weight from CT localizer images using deep learning. *Radiological Physics and Technology*, 14(1), 74–81. <https://doi.org/10.1007/s12194-020-00595-1>
71. Xu, Y., Li, Z., Zhang, Q., & Wang, Y. (2023). Evaluating Vision Transformer performance for chest radiograph classification. *Proceedings of the MICCAI 2023, Lecture Notes in Computer Science*, Springer. <https://arxiv.org/abs/2306.03944>