



A Causal AI Framework for Longitudinal Modelling of Prostate Cancer

Abstract

This project introduces a new paradigm in oncology with an interactive, Generative AI (GenAI)-based autonomous agent for managing prostate cancer (PCa). We will build a causal "digital twin" to transparently model the entire disease trajectory by integrating complex multimodal data (imaging, labs, notes) into a unified patient view.

Our approach directly aligns with the Challenge's Technological and Clinical Areas. Technologically, we will fuse multimodal data, use GenAI for medical data augmentation via counterfactuals and synthetic data, and create a medical knowledge base by disentangling disease signals from confounders. Clinically, our agent will improve predictive diagnosis by simulating disease progression and enable personalized treatment selection by forecasting therapeutic outcomes.

The framework is designed as a trustworthy, explainable, and safe AI system compliant with the EU AI Act. This high-risk, high-gain project will establish a new European standard for medical GenAI, delivering a transformative tool to enhance clinical decisions, improve patient outcomes, and revolutionize cancer care.

1 Excellence

1.1 Vision: A Paradigm Shift from Prediction to Causal Understanding

This project puts forward a radical new vision for clinical AI: a **paradigm shift from simple prediction to deep causal reasoning**, explicitly designed for a disease whose biology, histopathology, and phenotype evolve under therapy. The proposed solution focuses on prostate cancer (PCa) and proposes an end-to-end suite of interactive GenAI agents—orchestrated as a clinician-in-the-loop super-agent—that accompanies each patient across the full pathway, from predictive diagnosis to personalized treatment selection and longitudinal monitoring. We aim to build not just another predictive tool, but a dynamic, interactive **digital twin** of a patient's disease trajectory. This twin will (i) encode histopathologic and imaging heterogeneity, (ii) learn cause-and-effect mechanisms that link interventions to downstream biological and radiographic states, and (iii) simulate future disease trajectories, dynamically updating its predictions to incorporate new clinical data and interventions.

The Clinical Challenge. The management of PCa is fraught with complexity, stemming from the dual risks of overdiagnosis and undertreatment. The natural history of the disease is a complex, multi-stage process influenced by the tumor's biological state, the patient's individual characteristics, and the impact of clinical interventions. Age at presentation varies widely, complicating risk stratification. Widespread screening with prostate-specific antigen (PSA) has led to the frequent detection of indolent, clinically insignificant cancers, resulting in overtreatment and significant morbidity [1, 2, 3]. Conversely, many PCa deaths are due to the late diagnosis of aggressive or metastatic disease [1]. The stakes of misstaging are incredibly high; the distinction

between localized and metastatic disease carries profound prognostic implications, with the 5-year survival rate dropping from nearly 100% to as low as 30–40% [4, 5].

A Disease in Motion. Compounding matters, histopathology meaningfully diverges and therapy must adapt to that biology. Under treatment pressure, tumors can (trans)differentiate along alternate evolutionary paths (e.g., AR-independent, low-PSA neuroendocrine disease), often at unknown times. This yields PSA–imaging discordance, shifting lesion phenotypes, and non-trivial uncertainty about when to switch or optimize therapy. Meanwhile, the option set is expanding—from targeted agents to radioligand therapies (e.g., Lu-PSMA), advanced imaging (PSMA-PET), and liquid-biopsy assays .

Horizon Alignment and What is Radically New. Our program delivers two clinical objectives as agents: (1) a **Predictive-Diagnosis Agent** for individualized risk stratification and (2) a **Treatment-Selection Agent** that forecasts trajectories and policy value with explicit uncertainty and timing alerts. Technologically, we integrate three pillars: (A) multimodal integration across imaging, pathology, and EHR; (B) medical data augmentation to generate high-fidelity synthetic data (e.g., MRI→synthetic CT for dosimetry planning); and (C) medical knowledge representation via a causal knowledge graph. Across all components, we embed interpretability, calibrated uncertainty, and counterfactual explanations to transform the model from a black box into a trusted collaborator.

From Tool to Teammate, and from Lab to Ecosystem. The super-agent orchestrates the agents to provide a coherent, explainable plan. We will conduct controlled proof-of-concept studies against standard-of-care with external validation and maintain continuous uncertainty calibration and bias audits. To maximize European added value, we will leverage and contribute to shared research assets (e.g., imaging archives) and will generate and share synthetic datasets to accelerate portfolio-wide benchmarking.

Forging the Technology-to-Come. This project is a high-risk, high-gain endeavor to create a **foundational causal framework** that: (1) integrates multi-scale evidence into robust, updatable decision processes; (2) monitors phenotype shifts and triggers timely therapy pivots with explicit uncertainty; and (3) establishes a new European reference architecture for trustworthy, explainable, and regulation-ready medical AI in prostate cancer—while providing a template for other evolving, treatment-pressured malignancies.

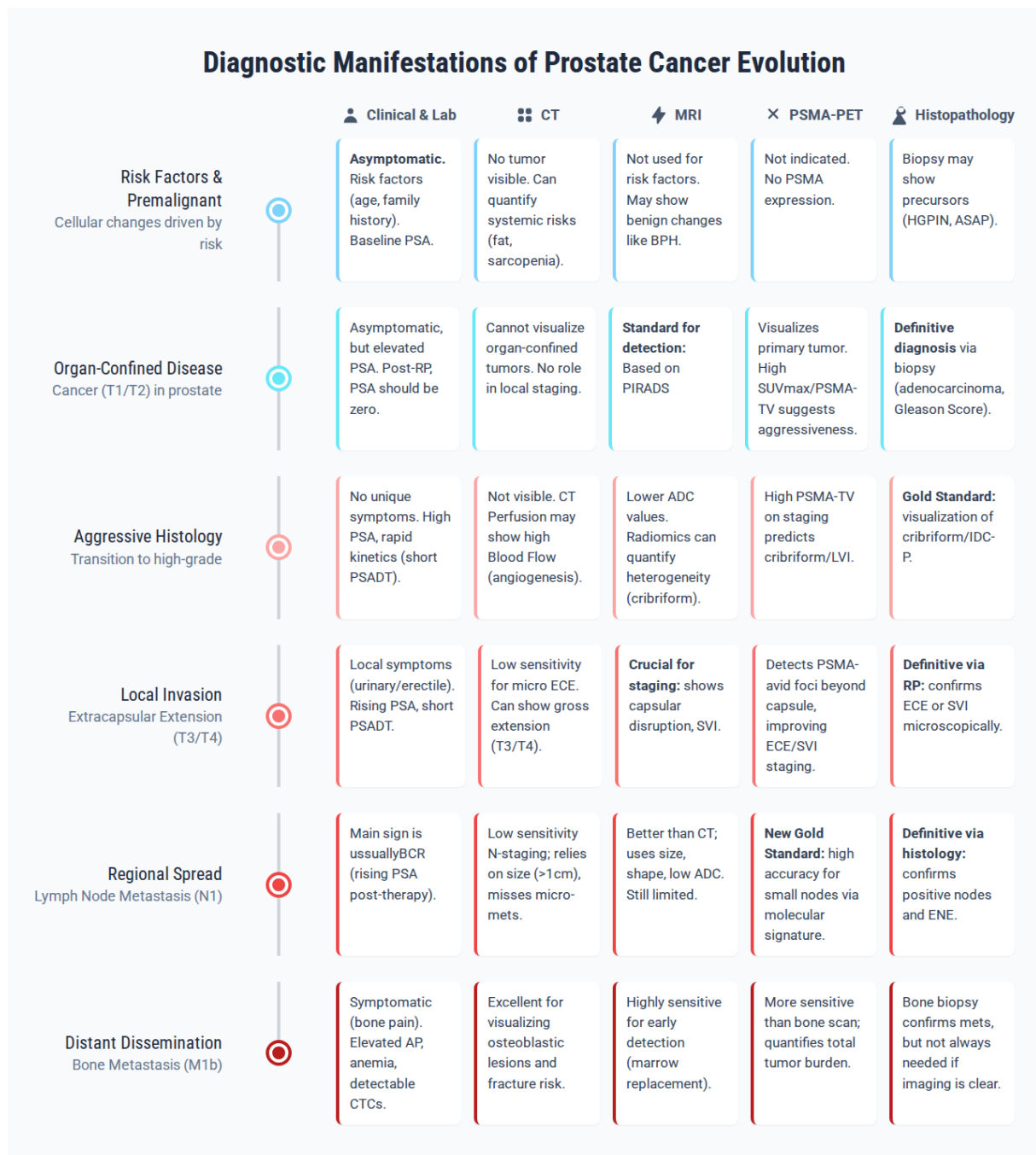


Figure 1: The natural history of prostate cancer, illustrating the stages our model will learn. The disease evolves based on tumor stage, patient state, and interventions, with information inferred from multiple clinical and imaging modalities. Our hypothesis is moreover that stage of a disease at initial diagnosis is the key to the best further therapy

Abbreviations: ADT (Androgen Deprivation Therapy), AP (Alkaline Phosphatase), ASAP (Atypical Small Acinar Proliferation), BCR (Biochemical Recurrence), BPH (Benign Prostatic Hyperplasia), CRPC (Castration-Resistant Prostate Cancer), CTCs (Circulating Tumor Cells), ECE (Extracapsular Extension), ENE (Extranodal Extension), HGPIN (High-Grade Prostatic Intraepithelial Neoplasia), LVI (Lymphovascular Invasion), PSADT (PSA Doubling Time)

1.2 Applicant’s Team Profile and Clinic Expertise

The applicant institution is uniquely positioned to succeed in this high-risk, high-gain endeavor. The applicant, the Universitätsklinik für Radiologie und Nuklearmedizin, possesses not only the requisite multimodal data and a proven track record in developing advanced AI models, but also the critical clinical expertise to ensure the solution’s reliability. Our approach is grounded in the principle that diagnostic data (PET/MR, CT, etc.) contains the key patient information, and our advantage as nuclear medicine specialists and radiologists is the deep understanding of its interpretation and central importance in the diagnostic chain. The project team is working on projects involving:

- **Synthetic Image Generation:** We have hands-on experience in generating synthetic CT scans from PET data, a foundational skill for the generative components of this proposal.
- **Predictive Modeling:** Our team has developed and validated models for predicting disease progression by fusing clinical and imaging data in Prostate cancer [6], directly relevant to the core of this project.
- **Clinical Application Development:** We have created a user-friendly GUI for structured reporting and developed an LLM-based support application for thyroid cancer guidelines, demonstrating our ability to translate research into practical clinical tools.

Crucially, the project’s leadership team possesses a rare and powerful combination of interdisciplinary expertise. The team includes two medical doctors (MDs) with medical PhDs. One leader combines a specialization in nuclear medicine with a Master’s degree in IT, providing a direct bridge between clinical application and technical implementation. The second leader holds a dual specialization in both nuclear medicine and radiology, offering an invaluable holistic perspective on the entire imaging pipeline from initial staging through monitoring to qualification to radioligand therapies. The collaboration with external partners, including the Abteilung für Nuklearmedizin at Universitätsmedizin Halle, Universitätsmedizin Charite, and the Radiological Practice Rad. Sudenburg, will provide access to diverse data and clinical perspectives, further strengthening the project’s foundation.



Figure 2: Algorithm of treatment in prostate cancer based on EAU Guidelines [7, 8].

This unique blend of expertise is complemented by the applicant's clinic, which has extensive, hands-on experience in managing patients through the entire spectrum of prostate cancer therapies (as shown on Figure 2) including late-stage prostate cancer, that is typically referred to as metastatic castration-resistant prostate cancer (mCRPC), is an incurable and fatal disease, despite substantial therapeutic advances [9, 10]. Managing mCRPC is complex, focusing on extending life, delaying disease progression, and maintaining or improving health-related quality of life (HRQOL) [9]. Our clinical practice routinely administers a wide array of treatments, including second-generation Androgen Receptor Pathway Inhibitors (ARPIs) like abiraterone, enzalutamide, apalutamide, and darolutamide [9, 10, 11]; taxane-based chemotherapies such as docetaxel and cabazitaxel [9, 10]; immunotherapy with checkpoint inhibitors like Pembrolizumab

[11, 12]; precision therapies such as PARP inhibitors (Olaparib) for patients with relevant genetic alterations [10, 11]; and cutting-edge radioligand therapies (RLT) including ^{177}Lu -PSMA and bone-targeted agents like Radium-223 [9, 13, 14]. The complexity of these cases necessitates regular multidisciplinary tumor boards, where detailed treatment decisions and patient responses are meticulously documented. Our framework is explicitly designed to support and integrate data from multiple such tumour boards across our collaborating institutions, learning from the collective expertise. This rich, real-world documentation provides an unparalleled source of high-quality data, crucial for training a causal model that can accurately learn the effects of different therapeutic interventions.

A key strength of this project is its foundation in a cutting-edge clinical environment defined by the use of ^{177}Lu -PSMA theranostics, a revolutionary and modern approach for managing mCRPC. This paradigm combines therapy ("thera") and diagnostics ("nostics") by using a single molecular target—the Prostate-Specific Membrane Antigen (PSMA)—for both imaging and treatment, representing a new and very promising therapy [15]. PSMA is a protein that is highly overexpressed (up to 1,000-fold) on the surface of prostate cancer cells, making it an ideal target [15, 16].

The theranostic workflow is a form of personalized medicine. It begins with a diagnostic PET scan using a PSMA-targeting molecule labeled with a diagnostic radionuclide (e.g., Gallium-68). This scan precisely identifies PSMA-positive tumor locations, confirming that the patient is a suitable candidate for the therapy [15, 17]. If eligibility is confirmed, the patient is then treated with a nearly identical molecule, but this time labeled with a therapeutic radionuclide, Lutetium-177. The ^{177}Lu is a short-range beta-particle emitter that delivers highly targeted radiation directly to cancer cells, minimizing damage to surrounding healthy tissue [15, 18].

This approach has proven to be a clinical breakthrough. The landmark Phase 3 VISION trial demonstrated that ^{177}Lu -PSMA-617 significantly prolonged both overall survival (15.3 vs 11.3 months) and radiographic progression-free survival (8.7 vs 3.4 months) in patients with advanced mCRPC, leading to its FDA approval in 2022 [19, 13, 12, 20]. The therapy is not only effective but also has a favorable safety profile compared to traditional chemotherapy, with fewer grade 3 or 4 adverse events reported in the TheraP trial (33% vs 53%) [21, 22].

Our project will leverage this unique dataset of paired diagnostic (^{68}Ga -PSMA PET/CT) and therapeutic (^{177}Lu -PSMA SPECT/CT) scans. This data provides an unparalleled opportunity to model the direct biological effects of a highly targeted and effective therapy, a critical component for building a robust causal model of disease progression and treatment response. Furthermore, because this therapy is administered at advanced stages, patients often have a rich, longitudinal history of prior treatments (including other systemic therapies like ARPIs and chemotherapies) and imaging. This provides the ideal data for our temporal models, giving us a unique "look-back" opportunity to model the entire patient path from a near-end point. The complexity of these cases also necessitates multidisciplinary tumor boards, giving us a unique opportunity to benchmark our model's recommendations against real-world expert consensus, making the applicant institution exceptionally well-suited for this project.

1.3 Objectives

Taking into account our previous experience, clinical knowledge and the guidelines described in the call, we have defined the main objectives of the project as follows:

1. **To Pioneer a Causal, Longitudinal Model of Prostate Cancer:** The core scientific objective is to develop and validate a novel Causal AI framework, built upon a Neural Jump Ordinary Differential Equation (NJDE) architecture. This model will be capable of learning the underlying dynamics of disease progression from sparse, irregularly-sampled,

multi-modal data, while explicitly modeling the impact of clinical interventions like surgery and therapy. This foundational technology will establish a new European standard for clinical AI.

2. **To Achieve True Explainability through Counterfactual Reasoning:** We will move beyond opaque "black-box" models by building a system that can generate clinically plausible, visual counterfactuals. This will enable clinicians to ask "what-if" questions and receive intuitive, visual explanations for the model's predictions, fostering the clinical trust necessary for widespread adoption and establishing a new paradigm for trustworthy AI.
3. **To Master Heterogeneity through Principled Disentanglement:** A key technical objective is to develop a robust Variational Autoencoder (VAE) framework that can disentangle the core components of medical imaging data. The model will learn to separate true disease signals from patient-specific anatomy, age-related changes, and technical confounders (e.g., scanner type, site-specific artifacts), ensuring the model is robust, generalizable, and ready for deployment across diverse European healthcare systems.
4. **To Create a Dynamic "Digital Twin" for Personalized Prognosis:** We aim to deliver a tool that can simulate a patient's future disease trajectory under various scenarios. The model will generate predictions not only of future scans but also of key clinical endpoints and structured radiological reports, providing a comprehensive, personalized prognostic tool that has the potential to revolutionize patient management.
5. **To Quantify Uncertainty for Safer Clinical Decisions:** We will develop and integrate robust methods for uncertainty quantification directly into our framework. The model will not only provide predictions but also a reliable measure of its own confidence, a critical feature for safe clinical use. This allows the system to "know what it doesn't know," flagging high-uncertainty cases for closer human review and preventing over-reliance on automated outputs.
6. **To Validate the Framework in a Real-World Clinical Context:** The project will culminate in a rigorous evaluation of the model's performance, using both quantitative metrics (e.g., predictive accuracy, counterfactual quality) and a qualitative, clinician-in-the-loop study to assess its real-world clinical plausibility, utility, and impact on diagnostic workflows, thereby paving the way for its integration into the European healthcare ecosystem.
7. **To Develop a Treatment-Choice Helper:** A key objective is to create a clinical decision support tool that can select the optimal treatment option or combination, as well as the ideal timing for interventions. This will be achieved by designing interventions that are explicitly optimized to maximize the expected time to progression, providing clinicians with data-driven strategies for patient management.

1.4 Concept and Methodology: **A Feasible Path** to a Groundbreaking Technology

Our methodology is a direct answer to the profound challenges of building trustworthy clinical AI. We aim to design a novel, **four-stage causal framework** that deconstructs the immense complexity of prostate cancer progression into a series of well-defined, manageable tasks. This modular approach is the key to our project's feasibility: it ensures training stability, enhances model interpretability, and allows us to systematically embed clinical domain knowledge as strong inductive biases. This guides the model to learn the true underlying causal mechanisms of the disease, rather than superficial correlations. This design leverages the project team's

extensive experience in synthetic image generation (CT from PET) and predictive modeling, and is illustrated in Figure 4.

1.4.1 Stage 1: Building the Bedrock of Clinical Validity with Supervisor Models

The first stage builds a suite of specialized "supervisor" models. These models act as expert guides, providing strong, clinically-validated signals that will enforce a valid structure on the more complex generative models in subsequent stages. Our team’s prior success in developing predictive models from fused clinical and imaging data provides a strong foundation for this work package.

- **Ordinal Classifiers:** For clinical scores with an inherent order (e.g., PI-RADS, TNM stage), standard categorical classifiers are suboptimal. We will implement a **differential ordinal learning framework** that explicitly encodes the ordered structure by combining a standard categorical loss with a differential ordinal loss, ensuring the model understands that a higher grade implies a worse prognosis [23, 24].
- **Censored Survival Regressor:** To predict time-to-progression (TTP), we will train a survival model that properly handles right-censored data. This will be achieved using a censored regression loss (e.g., Logistic Hazard) combined with a ranking loss regularizer (e.g., SurvRNC) to ensure correct risk ordering in the learned feature representation [25, 26, 27].
- **Anatomical and Confounder Models:** We will fine-tune a pre-trained TotalSegmentator model [28] to provide anatomical ground truth. Furthermore, we will train dedicated models to predict technical confounders (e.g., scanner type, dosage, data source location) and biological factors like patient age and BMI. While age and BMI are related to the disease, they also independently influence imaging appearance. To improve the quality of our counterfactuals, we will partially disentangle these effects from the primary disease signal, allowing us to model their impact on imaging separately. This approach ensures the model generalizes well and produces more clinically plausible counterfactuals [29, 30]. While sex as a biological variable is fixed for prostate cancer, the project will also investigate and mitigate potential data biases related to other demographic factors (e.g., age, ethnicity) to ensure the model’s equity and generalizability.

1.4.2 Stage 2: Mastering Heterogeneity through Principled Disentanglement

At the heart of our solution to data heterogeneity is a hierarchical Variational Autoencoder (VAE) trained for each imaging modality to learn a disentangled latent space. The key innovation is partitioning this space into four independent, semantically meaningful components: Anatomy (Z_A), Disease (Z_D), Patient State (Z_P), and Style/Confounders (Z_S). This separation is enforced through a composite loss function:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{VAE}} + \lambda_A \mathcal{L}_{\text{Anatomy}} + \lambda_D \mathcal{L}_{\text{Disease}} + \lambda_I \mathcal{L}_{\text{Disentangle}}$$

The disentanglement is achieved by moving beyond simple β -VAE approaches. Our loss function will apply a **targeted penalization** of the statistical dependence between latent subspaces. We recognize that some correlations are clinically meaningful (e.g., disease can affect anatomy), while others are spurious. Therefore, the model will be heavily penalized for mutual information between causally independent subspaces (e.g., Disease Z_D and Style Z_S , which includes scanner type and data source), while allowing for natural correlations between dependent factors like disease and anatomy. This is accomplished by selectively applying a Total Correlation (TC) penalty, ensuring the learned disease representation is invariant to technical artifacts without

sacrificing the reconstruction of clinically valid relationships [31, 32, 33]. The entire data curation and preprocessing pipeline is visualized in Figure 3.

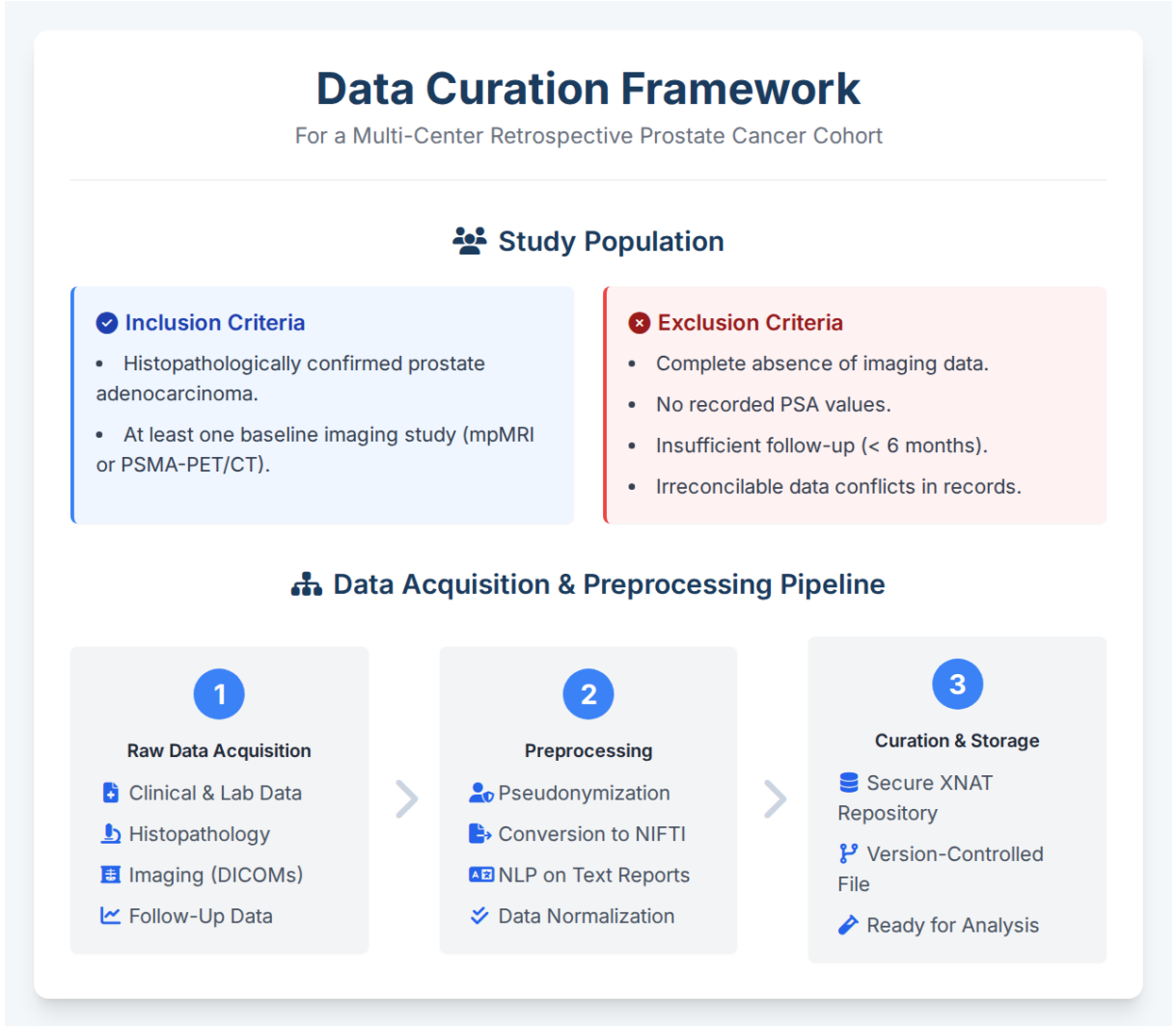


Figure 3: An infographic summarizing the data acquisition, curation, and preprocessing framework for the study cohort. It details the inclusion and exclusion criteria for patient selection and outlines the multi-step pipeline for processing clinical, histopathological, and imaging data.

1.4.3 Stage 3: Capturing Disease Dynamics with Neural Jump ODEs

This stage confronts the critical challenge of modeling disease evolution from sparse, multimodal, and irregularly-sampled clinical data. Our solution is a **Neural Jump Differential Equation (NJDE)** framework, which is uniquely suited for this task [34].

Rationale and Latent State Formulation Neural Ordinary Differential Equations (NODEs) are powerful because they model system dynamics in continuous time, making them inherently robust to irregular sampling—a defining characteristic of clinical data [34, 35, 36]. However, their primary weakness is the extreme computational cost of applying them directly to high-dimensional data like 3D images [37, 38]. Our approach strategically mitigates this by operating exclusively on the low-dimensional latent spaces learned in Stage 2. This is not just an efficiency

gain; thanks to the successful disentanglement, we do not need to pass the entire latent space to the temporal model. Instead, by training the NODE only on the most relevant parts—the disentangled **disease** (Z_D) and **patient state** (Z_P) **components**—we focus the model on learning the dynamics of disease progression itself, making the learning task more tractable and clinically relevant [39, 40, 41]. Crucially, by modeling the evolution of these purified latent vectors, we are not just fitting a curve to data points; we are learning a representation of the underlying causal trajectory of the disease, stripped of observational noise and technical artifacts.

The input for the NJDE is a carefully constructed time series of latent state vectors. For each patient, we define a unified state vector \mathbf{v} that has a fixed shape, representing a concatenation of all possible inputs. The process is as follows:

1. **Unified State Vector Definition:** A canonical tensor shape is defined for a state vector \mathbf{v} , which includes dedicated slots for the disease latent code (Z_D) and patient state code (Z_P) from each imaging modality (MRI, PET, SPECT), and for embeddings of all non-imaging data (PSA, clinical note embeddings, etc.).
2. **Time-stamped Sparse Tensor Creation:** For each time point t_i where a patient has data, a new state vector $\mathbf{v}(t_i)$ is created and initialized with zeros. The available data is then placed into its corresponding slot in the tensor. For example, at a time point with a PET scan and PSA value, only the $Z_{D_{\text{PET}}}$, $Z_{P_{\text{PET}}}$, and PSA embedding slots are filled, while all other slots remain zero.
3. **Anatomy Vector Storage:** The patient-specific anatomy vectors (Z_A) from each imaging study are not included in the dynamic state vector but are stored separately, indexed by time, for use in the final image reconstruction stage.

This procedure yields a dataset of sparse, irregularly-sampled latent state vectors, providing a computationally efficient and robust representation of each patient’s unique clinical history.

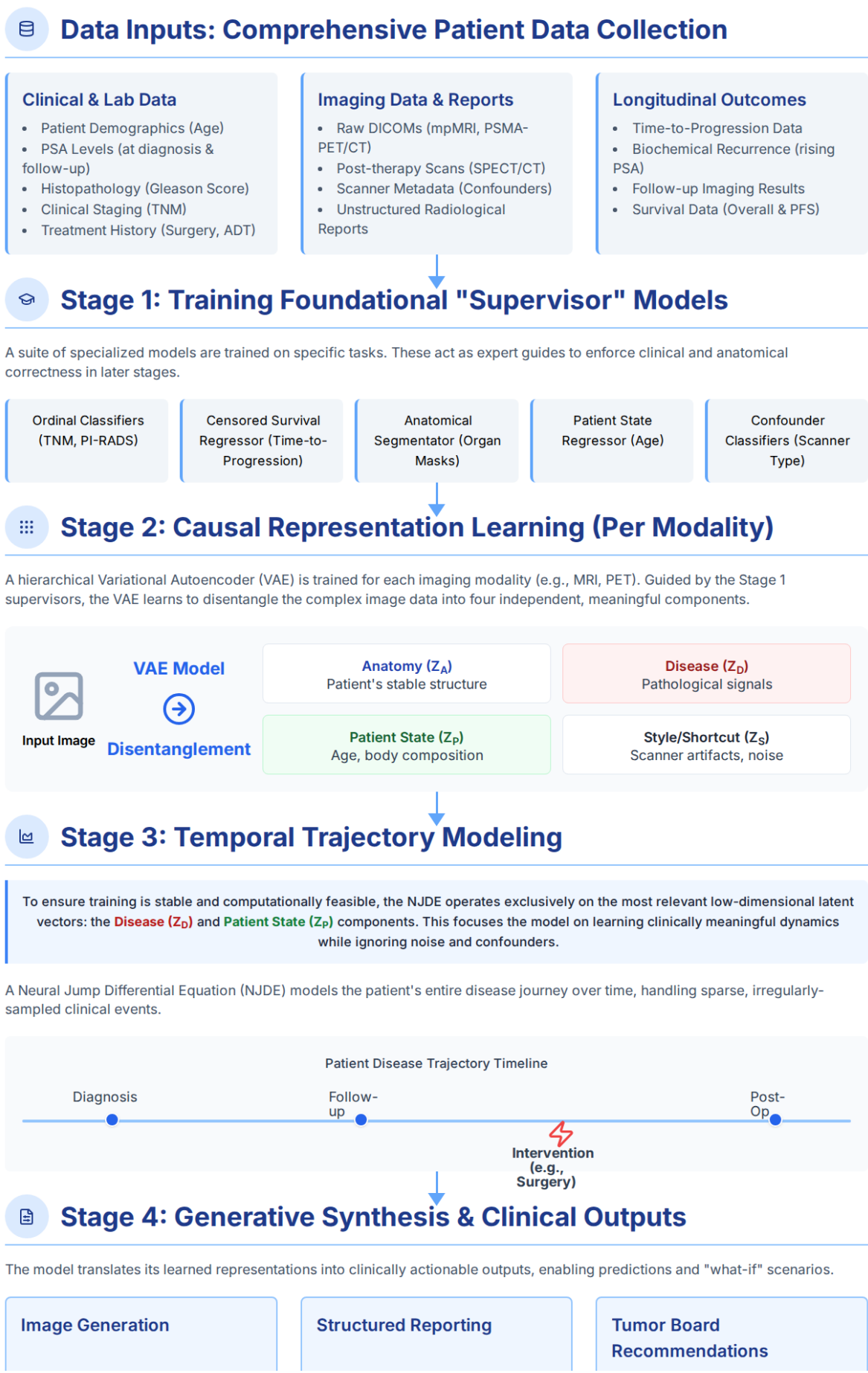
NJDE Training and Dynamics The NJDE learns the rules of disease evolution by modeling two phenomena [34]:

- **Continuous Evolution (The NODE):** Between clinical events, the disease state evolves smoothly. This is modeled by a classic NODE that learns the underlying dynamics from the sparse state vectors [42].

$$\frac{d\mathbf{v}(t)}{dt} = f(\mathbf{v}(t), t; \psi) \quad \text{for } t \neq t_k$$

- **Discrete Jumps (The Interventions):** At the specific time t_k of a clinical intervention (e.g., radical prostatectomy, initiation of systemic therapies such as ADT or chemotherapy, or targeted radioligand therapy with ^{177}Lu -PSMA), the continuous evolution is interrupted by a discrete "jump." A separate neural network, g , models this instantaneous change to the state vector based on the type of intervention [43, 44]. Crucially, the model is designed to handle multiple, sequential jumps, allowing it to accurately represent a patient’s entire treatment history, which may involve several different lines of therapy over time.

$$\mathbf{v}^+ = g(\mathbf{v}(t_k), \text{intervention}_k; \phi) \quad \text{for } t = t_k$$



This hybrid approach, which explicitly separates the continuous disease course from the rapid, external impacts of clinical interventions, is critical for accurately modeling a patient’s journey [34]. The model is trained using a **leave-one-out strategy** for each patient’s time series: given all but one time point, its goal is to predict the complete state vector for the held-out time point. To handle the pervasive missing data, we employ a **masked loss function**. The loss is calculated only by comparing the predicted values to the ground truth for those elements of the state vector that were actually available (non-zero) at the target time point. This forces the model to learn a probable evolution for every data type, even from a highly incomplete dataset. By distinguishing between smooth progression and intervention-driven shocks, the NJDE learns a more biologically plausible model of the patient’s journey.

1.4.4 Stage 4: Translating Insight into Action: Generative Synthesis, Decision Support, and Structured Reporting

The final stage translates the model’s learned causal dynamics into clinically actionable outputs, evolving the framework from a descriptive tool into an active partner in clinical decision-making. This includes generating high-fidelity images for any time point or counterfactual scenario, with a lightweight Diffusion Model used as a final post-processing step to ensure realism. These visual outputs serve as the intuitive foundation for the system’s core decision support capabilities.

Crucially, the framework will function as a **Treatment-Choice Helper**. Leveraging the NJDE from Stage 3, the system will simulate multiple future patient trajectories, each conditioned on a different potential treatment intervention and timing. By comparing these simulated outcomes, the model can identify the therapeutic strategy that is explicitly optimized to **maximize the expected time to progression**. This transforms the "what-if" capability from a simple explanatory tool into a powerful engine for data-driven treatment planning.

All insights are then consolidated into automatically generated structured reports using a **Transformer-based decoder**. These reports will integrate prognostic simulations, present the comparative effectiveness of different treatment options, and culminate in a data-driven recommendation. This directly addresses the clinical need for clear, consistent, and efficient documentation [45, 46], while providing clinicians with novel, data-driven strategies for patient management. The project team’s experience in creating LLM-based support apps and GUIs from structured reports is key to ensuring this complex information is presented in an intuitive and actionable format.

1.5 Ambition

The ambition of this project extends far beyond advancing the state-of-the-art in predictive modeling. We aim to establish a new paradigm for trustworthy AI in clinical medicine, shifting the focus from correlation to causation. By developing a model that can reason, simulate, and explain, we are creating a foundational technology with the potential for profound impact. This project will not only deliver a powerful tool for prostate cancer management but will also provide a blueprint for developing causal AI models in other complex disease areas. Our work will pioneer new methods for disentanglement, longitudinal modeling, and counterfactual reasoning that will be of immense value to the wider AI research community. The successful completion of this project will represent a significant step towards the realization of truly intelligent clinical decision support systems, paving the way for a future of more personalized, more effective, and more explainable medicine.

1.6 Compliance with the EU Artificial Intelligence Act

The proposed framework is designed from the ground up to align with the principles of trustworthy AI and to comply with the requirements for high-risk AI systems as laid down in the Regulation (EU) 2024/1689 (the "AI Act"). As a system intended for use in medical diagnosis and to guide treatment, it is classified as a **high-risk AI system** under Annex III of the Act. Our methodology directly addresses the core obligations for such systems:

- **Risk Management System (Article 9):** Our project management (WP7) will implement a continuous, iterative risk management process as mandated by Article 9, which will be maintained throughout the AI system's lifecycle. Specifically, this will be operationalized through a dedicated **Risk Register**, established at the project's outset (M1). This living document will log all known and reasonably foreseeable risks to health, safety, fundamental rights, and project execution. For each identified risk, we will assess its severity and likelihood, and define concrete mitigation and contingency plans. The register will be reviewed quarterly by the project management team and discussed in detail during biannual project meetings with the external collaborators. The applicant (Universitätsklinik für Nuklearmedizin, Magdeburg) will take the lead in identifying clinical risks (e.g., misdiagnosis, incorrect treatment suggestions), while the technical team will focus on system-level risks (e.g., model bias, cybersecurity vulnerabilities, performance degradation). This structured and collaborative process ensures that risks are not only identified but actively managed and mitigated throughout the project.
- **Data and Data Governance (Article 10):** WP1 is entirely dedicated to establishing data governance practices that meet the quality criteria of Article 10. We will use relevant, representative datasets. We will proactively implement measures to detect and mitigate potential biases through the disentanglement methods in Stage 2 and the confounder models in Stage 1. For the purpose of bias detection and correction, we will process special categories of personal data only where strictly necessary and with the appropriate safeguards as permitted under Article 10(5).
- **Technical Documentation and Record-Keeping (Articles 11 & 12):** We commit to creating and maintaining comprehensive technical documentation as specified in Annex IV of the AI Act. Our system's design includes automatic logging capabilities to ensure traceability of operations, in compliance with the record-keeping requirements of Article 12.
- **Transparency and Provision of Information (Article 13):** A cornerstone of our project is to overcome the "black box" problem. The framework's ability to generate counterfactual explanations (Stage 4) is a direct implementation of the transparency requirements of Article 13. The system is designed so that its operations are sufficiently transparent to enable deployers to interpret its output and use it appropriately. We will provide detailed instructions for use that explain the system's capabilities, limitations, and the role of human oversight.
- **Human Oversight (Article 14):** The system is designed to augment, not replace, clinical experts. It functions as a decision support tool, ensuring that a natural person can oversee its functioning at all times. The design ensures that the user can understand the system's capabilities, monitor its operation, and decide to disregard, override, or reverse its output, fulfilling the requirements of Article 14.
- **Accuracy, Robustness, and Cybersecurity (Article 15):** WP5 is dedicated to rigorous validation of the model's accuracy and robustness. The system will be designed to

be resilient against errors and inconsistencies through the disentanglement of confounders (Stage 2). To comply with Article 15, we will implement a multi-layered, comprehensive security strategy that addresses AI-specific threats, ensures patient data privacy, and enforces strict regulatory compliance [47, 48]. Our approach is built on three pillars:

1. Enhancing AI Model Robustness and Integrity To defend against malicious inputs and training manipulations, we will:

- **Stress Testing and OOD Detection:** We will conduct rigorous stress tests to validate robustness and implement Out-of-Distribution (OOD) detection to flag inputs that fall outside the model’s expected data range, thereby preserving reliability [49].
- **Data Quality Assessment:** We will perform thorough checks on all data for completeness, consistency, and correctness to avoid performance degradation [50].

2. Operational, Transparency, and Regulatory Frameworks To ensure accountability and long-term reliability, we will establish:

- **Continuous Monitoring and Anomaly Detection:** We will implement continuous monitoring to identify anomalous model behaviors in real-time [47, 50].
- **eXplainable AI (XAI):** Our counterfactual explanation system is a core part of our XAI strategy, providing transparency that is crucial for clinician trust and accountability [50, 51].
- **A Comprehensive Audit Trail:** We will maintain a complete audit trail logging all user actions, system access, and predictions to ensure traceability [50, 49].

This principled approach ensures that our innovative framework is not only technologically advanced but also safe, trustworthy, and fully compliant with the Union’s legal framework for artificial intelligence.

2 Impact

The successful execution of this project will generate significant and lasting impact across multiple domains, from advancing the frontiers of science and technology to delivering tangible benefits for patients, clinicians, and the European innovation ecosystem. Our focus is not merely on creating a tool, but on establishing a new technological paradigm for clinical AI that is trustworthy, explainable, and directly aligned with the goals of the EU Cancer Mission and the strategic autonomy of the Union.

2.1 Scientific and Technological Impact

This project is poised to make fundamental contributions to the field of Artificial Intelligence and its application in medicine, directly addressing the EIC Pathfinder’s goal to develop the scientific basis for breakthrough technologies.

- **A New European Standard for Trustworthy Clinical AI:** We will pioneer a shift away from correlational "black-box" models towards truly causal and explainable AI. The development of a robust, generalizable framework for learning causal models from observational clinical data will represent a landmark achievement. This provides a blueprint for future research in a wide range of diseases and establishes a new European standard for

the development and deployment of trustworthy AI systems in high-stakes environments, strengthening Europe’s technological leadership.

- **Advancing the Frontier of Generative AI:** Our work on principled disentanglement, counterfactual generation, and longitudinal modeling with Neural Jump ODEs will push the boundaries of generative AI. By creating a model that can reason about cause and effect, we are developing a foundational technology with applications far beyond medicine, including in climate science, economics, and engineering. The novel techniques developed will be of immense value to the broader European AI community.
- **Fostering an Open and Sovereign European AI Ecosystem:** We are deeply committed to open science and Europe’s digital sovereignty. We will make our code publicly available under a permissive open-source license. To ensure compliance with all data protection laws, we will **attempt to contribute** our curated, anonymized, and harmonized datasets to the **European Cancer Imaging (EUCAIM)** platform. Alternatively, we will make data available to interested institutions via secure, bilateral data sharing agreements. This strategy ensures our work can foster further research and innovation across the EU while upholding the strictest data privacy standards.

2.2 Societal and Clinical Impact


The primary impact of this project will be the profound improvement in the management of prostate cancer, leading to better patient outcomes and more efficient, resilient healthcare systems across Europe.

- **Enhanced Diagnostic Accuracy and Personalised Treatment:** By providing clinicians with a "digital twin" that can simulate disease trajectories and treatment responses, our framework will enable more accurate staging, better risk stratification, and truly personalized treatment planning. This will help to reduce both over- and under-treatment—a critical issue in prostate cancer—thereby minimizing treatment-related side effects and significantly improving patient quality of life.
- **Empowering Clinicians and Reducing Workload:** The model’s ability to generate intuitive, counterfactual explanations ("what if this lesion were benign?") and automated structured reports will empower clinicians, augmenting their decision-making process and freeing up valuable time from tedious documentation. This will improve workflow efficiency, reduce burnout, and allow clinicians to focus on what matters most: patient care.
- **A Foundation for a New Era in European Oncology:** While focused on prostate cancer, the foundational technology developed in this project is modality-agnostic and disease-agnostic. It has the potential to be adapted to other cancers (e.g., breast, lung) and complex chronic diseases, paving the way for a new era of data-driven, causal medicine that is more personalized, more effective, and more explainable, with Europe at the forefront.

2.3 Dissemination, Communication, and Exploitation

We are committed to maximizing the impact of this project through a comprehensive dissemination, communication, and exploitation strategy designed to engage stakeholders at all levels and create new market opportunities.

- **High-Impact Scientific Publications:** We will publish our methodological advancements and clinical validation results in **leading peer-reviewed journals and present at premier international conferences**.

- **Open Source and Open Data:** As stated, all code will be open-sourced via European platforms. In line with our data governance strategy, we will attempt to contribute our curated, anonymized datasets to the **European Cancer Imaging (EUCAIM)** platform or share them via secure bilateral agreements to ensure our data legacy strengthens the European research community while respecting data protection regulations.
- **Engagement with the Clinical and Patient Community:** We will actively engage with  European clinical societies. This ensures our work is aligned with real-world clinical needs and facilitates its translation into clinical practice across the EU.
- **Commercial Exploitation and Standardization:** We will explore pathways for commercial exploitation through partnerships with European medical imaging companies and health-tech startups, creating new market opportunities within the Union. Furthermore, we will actively participate in standardization bodies to promote the adoption of our structured reporting and causal AI frameworks as a new European standard for trustworthy clinical AI.

3 Implementation

3.1 Work Plan, Timeline, and Deliverables

The project is structured into seven interconnected Work Packages (WPs). The overall project timeline is visualized in the Gantt chart in Figure 5, providing a clear roadmap for the project's execution over its 36-month duration.

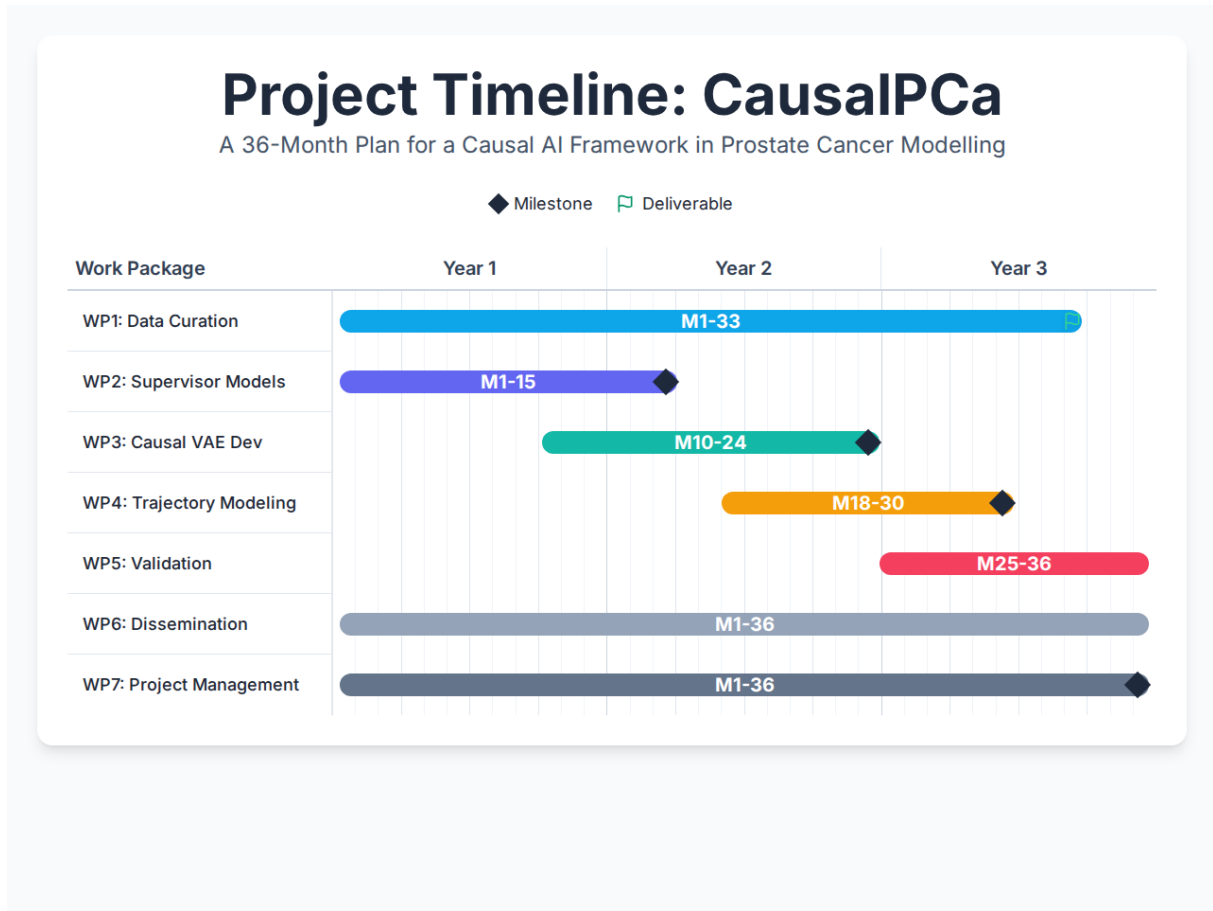


Figure 5: Project Gantt Chart illustrating the timeline for Work Packages, Deliverables (D), and Milestones (M).

- **WP1: Data Curation and Harmonization** – 30 Person-Months
- **WP2: Foundational Supervisor Model Training** – 50 Person-Months
- **WP3: Causal VAE Development** – 60 Person-Months
- **WP4: Temporal Trajectory Modeling** – 60 Person-Months
- **WP5: Validation and Clinical Integration** – 30 Person-Months
- **WP6: Dissemination, Communication, and Exploitation** – 10 Person-Months
- **WP7: Project Management** – 10 Person-Months

To monitor progress and ensure the timely completion of project objectives, a set of key deliverables is defined, as outlined in Table 1. These outputs represent the tangible results of each work package.

Table 1: List of Project Deliverables.

Del. No.	Deliverable Name	WP	Type	Dissem.	Due
D1.1	Curated, harmonized dataset (initial)	WP1	Data	CO	M12
D2.1	Validated supervisor models	WP2	Code	PU	M15
D3.1	Disentangled VAE models	WP3	Code	PU	M24
D4.1	Trained Neural Jump ODE model	WP4	Code	PU	M30
D5.1	Final validation report	WP5	R	PU	M36
D6.1	Project website and open-source repo	WP6	Web	PU	M3
D6.2	Mid-term dissemination report	WP6	R	PU	M18
D6.3	Final dissemination & exploitation plan	WP6	R	PU	M36
D7.1	Project management handbook	WP7	R	CO	M2

Type: R=Report, Data=Dataset, Code=Software, Web=Website. **Dissem:** PU=Public, CO=Confidential.

Complementing the deliverables, the project’s progress will be assessed against the verifiable milestones listed in Table 2. These milestones serve as critical checkpoints to validate the achievement of the project’s scientific and technical goals.

Table 2: List of Verifiable Milestones.

MS No.	Milestone Name	WP	Due	Means of Verification
M1	Project Kick-off and Risk Register established	WP7	M1	Kick-off meeting minutes; Initial Risk Register document available.
M2	Supervisor models achieve target performance	WP2	M15	Validation report showing mean accuracy >0.75 on internal test set.
M3	VAE models demonstrate successful disentanglement	WP3	M24	Report with quantitative disentanglement metrics (measured by reduced mutual information between dimensions of latent space without trained disentanglement) and qualitative results.
M4	NJDE model successfully predicts patient trajectories	WP4	M30	Report on temporal model performance, with mean accuracy >0.75 for predicting TNM and PSA at different time points.
M5	Final model validated for clinical plausibility	WP5	M36	Final validation report where a majority of counterfactual images are deemed plausible and useful by testing physicians.

3.2 Data Sources and Cohort

A key strength of this proposal is the direct access to already preprocessed, rich, longitudinal, and multimodal data from the applicant’s own clinical institutions and established collaborators as described in Table 3. This will form the core training dataset, which can be expanded with new cases during the project from our own or collaborating institutions. An ethical approval for using the preprocessed data for scientific purposes has already been obtained.

Table 3: Summary of Data Sources and Cohorts

Data Type	Contributing Institutions	Min. Patient Count
Proprietary Multi-Center Clinical Cohort		
Paired ^{68}Ga -PSMA PET/CT & mpMRI	Magdeburg, Halle, Charite	350
Paired pre-therapy ^{68}Ga -PSMA PET/CT & post-therapy ^{177}Lu -PSMA SPECT/CT	Magdeburg, Halle, Charite	550
Paired interim therapy ^{68}Ga -PSMA PET/CT & CT	Magdeburg	40
Longitudinal ^{68}Ga -PSMA PET/CT	Magdeburg	200
Diagnostic CT & mpMRI	Rad. Sudenburg (Magdeburg)	200
Total Minimum Proprietary Cohort		1340
<i>Note: The patient counts are minimum estimates and are expected to increase during the project.</i>		
Integration with Public Datasets		
Genomic, MRI, CT, PET Data	The Cancer Genome Atlas (TCGA-PRAD) [52]	14
Longitudinal MRI (297) & PSA Data	ProstateNET (EUCAIM) [53]	297

This combined data strategy provides an unparalleled foundation for this high-risk, high-gain project, mitigating the risk of data scarcity and ensuring the developed technology is robust, validated, and ready for broader clinical application.

Data Governance and AI Act Compliance (Article 10): All data will be managed within a secure, learning-ready environment based on the XNAT platform, ensuring robust data hosting and management. Our data governance framework is designed to meet the stringent quality criteria of Article 10 of the EU AI Act. The data collection is overseen by our clinical partners, and all data undergoes a rigorous curation and verification process by trained clinicians to ensure it is relevant, representative, and as free of errors as possible. To address potential biases, our disentanglement methods (Stage 2) and confounder models (Stage 1) are designed to detect and mitigate biases related to demographics and acquisition hardware. For the purpose of bias detection and correction, we will process special categories of personal data only where strictly necessary and with the appropriate safeguards as permitted under Article 10(5), ensuring our dataset is suitable for training a high-risk AI system. All data will be handled in strict compliance with GDPR and ethical guidelines.

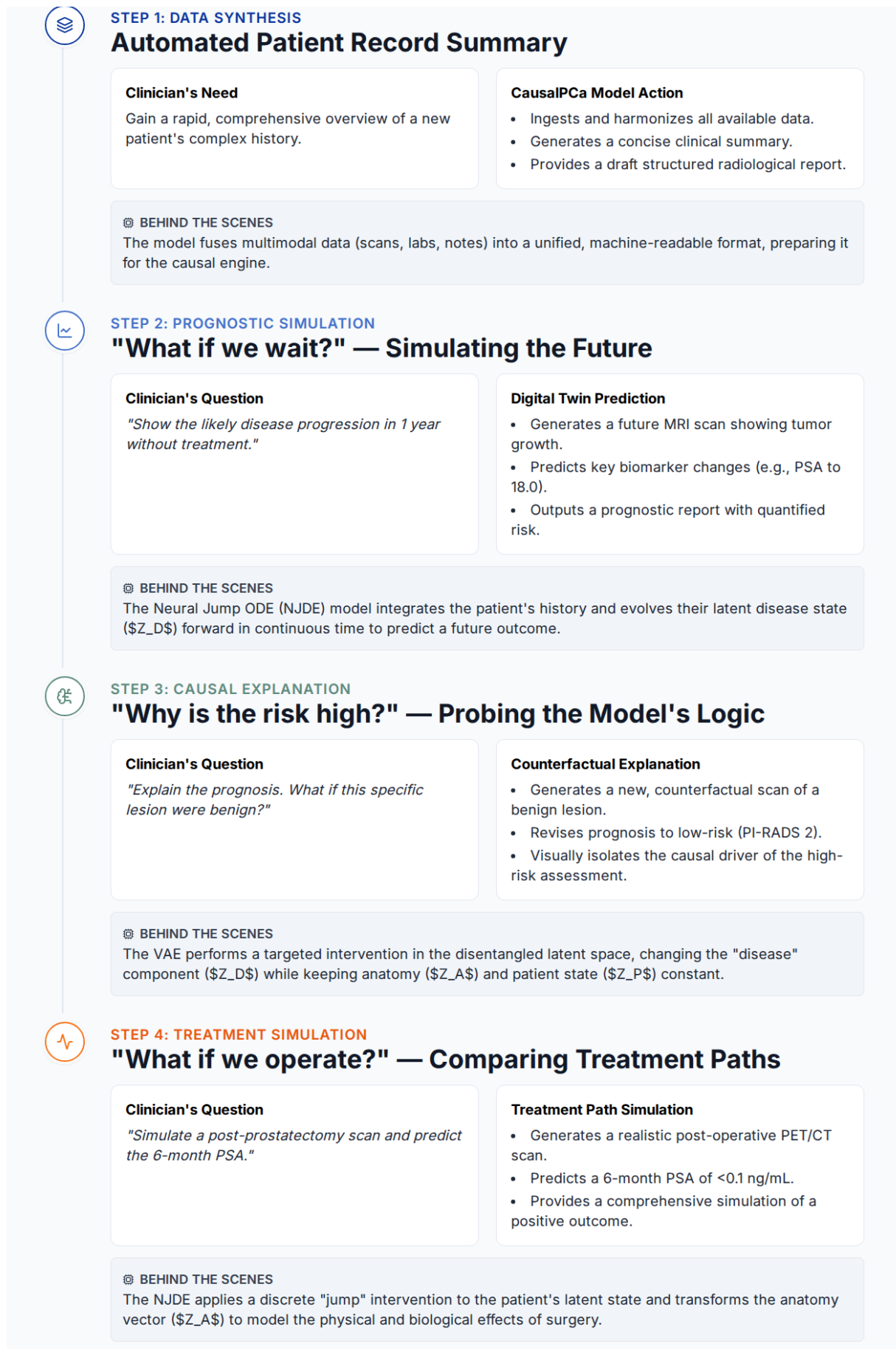


Figure 6: The CausalPCa Digital Twin in action.²⁰ This workflow demonstrates how a clinician can interact with the AI.

3.3 Illustrative Clinical Workflow

To make our vision concrete, Figure 6 illustrates the step-by-step workflow of a clinician interacting with the CausalPCa Digital Twin. This infographic demonstrates how the tool moves beyond simple prediction to become an interactive partner in clinical decision-making, seamlessly integrating data synthesis, prognostic simulation, causal explanation, and treatment planning into a single, intuitive interface.

3.4 Evaluation and Validation

The project’s success will be measured through a rigorous, multi-faceted evaluation plan that combines quantitative metrics with qualitative, clinician-in-the-loop studies to assess real-world utility and trustworthiness.

3.4.1 Quantitative Metrics

Model performance will be assessed using a comprehensive suite of metrics tailored to each sub-task:

- **Supervisor Model Performance:**
 - **Classification:** Accuracy, Area Under the Curve (AUC), F1-Score, and Quadratic Weighted Kappa for ordinal tasks.
 - **Survival Regression:** Concordance Index (C-index) and Mean Absolute Error on censored data.
- **Generative Model Performance:**
 - **Image Generation Quality:** Fréchet Inception Distance (FID), Structural Similarity Index (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS) to measure realism and fidelity [54, 55].
 - **Counterfactual Quality:** We will use a comprehensive suite of metrics to assess axiomatic soundness (effectiveness, composition, reversibility) [56, 57], validity (success rate of flipping a classifier’s decision) [58, 55].

3.4.2 Uncertainty Quantification

A key feature for clinical trust is the model’s ability to quantify its own uncertainty. Our VAE-based architecture is intrinsically probabilistic and allows for robust uncertainty quantification. We will employ two complementary methods:

- **Latent Space Sampling:** For a given input, we will draw multiple samples from its learned latent distribution. The variance in the resulting predictions will serve as a robust measure of the model’s epistemic uncertainty [59].
- **Direct Variance Utilisation:** The variance vector σ^2 produced by the VAE encoder is a direct indicator of feature-level uncertainty. We will concatenate this variance vector to the mean vector as a direct input to downstream models, allowing them to learn to depend more on high-confidence features [60].

3.4.3 Clinical Plausibility and Workflow Integration Study

Beyond quantitative metrics, a practical, clinician-in-the-loop study is essential to assess the model’s real-world utility and trustworthiness.

- **Assessing Counterfactual Plausibility and Usefulness:** To evaluate the quality of our generated explanations, we will conduct a human-grounded study with radiologists and oncologists. Clinicians will be presented with a series of cases, each including an original image and its corresponding model-generated counterfactual (e.g., an image of a tumorous prostate and its benign-looking counterfactual). They will score the counterfactuals on a Likert scale for:
 - **Clinical Plausibility:** Does the generated image appear realistic and anatomically sound?
 - **Usefulness for Explanation:** Does the visual difference between the original and counterfactual image provide a clear and understandable reason for the model’s prediction?
- **Measuring Workflow Improvement with Structured Reports:** To assess the impact of the automated reports, we will perform a comparative workflow study. A control group of clinicians will review patient cases using traditional free-text reports and standard image viewers. An experimental group will review the same cases using our system’s auto-generated structured reports and prognostic visualizations. We will measure:
 - **Efficiency:** Time taken to extract key information (e.g., TNM stage, PI-RADS score, presence of key findings) required for treatment planning.
 - **Accuracy and Concordance:** Inter-rater agreement and accuracy of the extracted information compared to an expert-defined ground truth.
 - **User Satisfaction:** Clinicians’ perceived efficiency, clarity, and confidence in their decisions will be assessed using a standardized questionnaire (e.g., the System Usability Scale).

3.5 Risk Analysis, Mitigation, and AI Act Compliance

This is an ambitious, high-risk project at the frontier of AI research. We have identified the primary risks and have developed clear mitigation strategies that are intrinsically linked to our compliance with the EU AI Act. Our risk management process, as detailed in WP7, is continuous and iterative, designed to address risks to health, safety, and fundamental rights throughout the AI system’s lifecycle, in direct alignment with **Article 9** of the Act. Furthermore, our methodology is designed to comply with the ‘do no significant harm’ principle as per Article 17 of the EU Taxonomy Regulation, ensuring our research does not negatively impact environmental objectives.

The critical risks are summarized in Table 4.

Table 4: Critical Risks and Mitigation Strategies.

No.	Description	LikelihoodImpact		Mitigation Strategy
1	Training Instability and Data Scalability	High	High	Sequential, multi-stage training framework to decompose complexity. Modular design allows for isolated troubleshooting.
2	Generalizability and Overfitting to Spurious Correlations	Medium	High	Principled disentanglement to separate causal factors from confounders. Use of multi-center data and strong inductive biases (e.g., ordinal losses). Compliant with AI Act Articles 10 & 15.
3	Performance with Incomplete and Heterogeneous Data	High	High	Neural Jump ODE architecture is inherently designed for sparse, irregular data. A masked loss function allows the model to learn from all available data without being penalized for missingness.
4	Clinical Trustworthiness and the "Black Box" Problem	Medium	High	Explainability through counterfactuals allows clinicians to probe model reasoning. System is designed as a decision support tool, ensuring human oversight at all times. Compliant with AI Act Articles 13 & 14.
5	Model Drift and Performance Deterioration	Medium	High	Grounding the model in fundamental biological knowledge through disentanglement makes it more robust to superficial data shifts. Continuous monitoring will be implemented to detect performance degradation early.

Furthermore, comprehensive **technical documentation** will be maintained throughout the project as specified in Annex IV of the AI Act, and our system’s design includes automatic logging capabilities to ensure traceability of operations, in compliance with the record-keeping requirements of **Articles 11 and 12**. Furthermore, comprehensive **technical documentation** will be maintained throughout the project as specified in Annex IV of the AI Act, and our system’s design includes automatic logging capabilities to ensure traceability of operations, in compliance with the record-keeping requirements of **Articles 11 and 12**.

4 Conclusions

This project represents a bold step towards a new frontier in clinical AI. We propose not an incremental improvement, but a foundational shift from correlational pattern recognition to causal, explainable, and trustworthy artificial intelligence. Our vision is to create a dynamic "digital twin" for prostate cancer that empowers clinicians with a tool that can reason, simulate, and explain—augmenting their expertise and enabling truly personalized medicine.

The CausalPCa framework is designed to tackle the immense complexity of prostate cancer by deconstructing it into manageable, causally-grounded stages. By pioneering novel methods in disentanglement, longitudinal modeling, and counterfactual generation, we will deliver a technology that is not only powerful but also robust, generalizable, and aligned with the highest standards of the EU AI Act.

This high-risk, high-gain endeavor will deliver a transformative tool for oncology, establish a new European standard for trustworthy AI, and provide a versatile, foundational technology with the potential to revolutionize how we understand and treat complex diseases. We are confident that the applicant’s team, with its deep interdisciplinary expertise and access to unparalleled clinical data, is perfectly positioned to turn this ambitious vision into a reality, generating profound scientific, clinical, and societal impact for Europe.

References

- [1] Anwar R. Padhani and Ivo G. Schoots. Imaging-based diagnostic and therapeutic strategies for prostate cancer in the coming decades. *Radiology*, 307, 6 2023.
- [2] Amarnath Jena, Rajesh Taneja, Sangeeta Taneja, Aru Singh, Venkatesh Kumar, Anshuman Agarwal, and Narasimhan Subramanian. Improving diagnosis of primary prostate cancer with combined ⁶⁸ga-prostate-specific membrane antigen-hbed-cc simultaneous pet and multiparametric mri and clinical parameters. *American Journal of Roentgenology*, 211, 12 2018.
- [3] Matteo Caracciolo, Angelo Castello, Luca Urso, Francesca Borgia, Naima Ortolan, Licia Uccelli, Corrado Cittanti, Massimo Castellani, Mirco Bartolomei, Massimo Lazzeri, and Egesta Lopci. The role of [68ga]psma pet/ct for clinical suspicion of prostate cancer in patients with or without previous negative biopsy: A systematic review. *Cancers*, 14, 10 2022.
- [4] Jade Wang, Elisabeth O’Dwyer, Juana Martinez Zuloaga, Kritika Subramanian, Jim C. Hu, Yuliya S. Jhanwar, Himanshu Nagar, Arindam RoyChoudhury, John Babich, Sandra Huicochea Castellanos, Joseph R. Osborne, and Daniel J. A. Margolis. Reasons for discordance between 68ga-psma-pet and magnetic resonance imaging in men with metastatic prostate cancer. *Cancers*, 16, 5 2024.
- [5] Lorenzo Cereser, Laura Evangelista, Gianluca Giannarini, and Rossano Girometti. Prostate mri and psma-pet in the primary diagnosis of prostate cancer. *Diagnostics*, 13, 8 2023.
- [6] Joanna Wybranska, Lorenz Pieper, Christian Wybranski, Philipp Genseke, Jan Wuestemann, Julian Varghese, Michael Kreissl, and Jakub Mitura. Integrating ga-psma-11 pet/ct with clinical risk factors for enhanced prostate cancer progression prediction. *Cancers*, 17:2285, 07 2025.
- [7] Philip Cornford, Roderick Bergh, Erik Briers, Thomas Van den Broeck, Marcus Cumberbatch, Maria Santis, Stefano Fanti, Nicola Fossati, Giorgio Gandaglia, Silke Gillesen, Nikolaos Grivas, Jeremy Grummet, Ann Henry, Theodorus Van der Kwast, Thomas Lam, Michael Lardas, Matt Liew, Malcolm Mason, Lisa Moris, and James Catto. Eau-eanm-estro-esur-siog guidelines on prostate cancer. part ii-2020 update: Treatment of relapsing and metastatic prostate cancer. *European Urology*, 79, 10 2020.
- [8] Nicolas Mottet, Roderick Bergh, Erik Briers, Thomas Van den Broeck, Marcus Cumberbatch, Maria Santis, Stefano Fanti, Nicola Fossati, Giorgio Gandaglia, Silke Gillesen, Nikolaos

- Grivas, Jeremy Grummet, Ann Henry, Theodorus Van der Kwast, Thomas Lam, Michael Lardas, Matt Liew, Malcolm Mason, Lisa Moris, and James Catto. Eau-eanm-estro-esur-siog guidelines on prostate cancer-2020 update. part 1: Screening, diagnosis, and local treatment with curative intent. *European Urology*, 79, 11 2020.
- [9] Karim Fizazi, Ken Herrmann, Bernd J Krause, Kambiz Rahbar, Kim N Chi, Michael J Morris, Oliver Sartor, Scott T Tagawa, Ayse T Kendi, Nicholas Vogelzang, Jeremie Calais, James Nagarajah, Xiao X Wei, Vadim S Koshkin, Jean-Mathieu Beaugerard, Brian Chang, Ray Ghouse, Michelle DeSilvio, Richard A Messmann, and Johann de Bono. Health-related quality of life and pain outcomes with [177lu]lu-psma-617 plus standard of care versus standard of care in patients with metastatic castration-resistant prostate cancer (vision): a multicentre, open-label, randomised, phase 3 trial. *The Lancet Oncology*, 24, 6 2023.
 - [10] Koji Hatano and Norio Nonomura. Systemic therapies for metastatic castration-resistant prostate cancer: An updated review. *The World Journal of Men's Health*, 41, 2023.
 - [11] Jiao Ma, Lanying Li, Taiping Liao, Weidong Gong, and Chunyin Zhang. Efficacy and safety of 225ac-psma-617-targeted alpha therapy in metastatic castration-resistant prostate cancer: A systematic review and meta-analysis. *Frontiers in Oncology*, 12, 2 2022.
 - [12] Brian Ramnarain and Oliver Sartor. Psma-targeted radiopharmaceuticals in prostate cancer: Current data and new trials. *The Oncologist*, 28, 2 2023.
 - [13] Elias Chandran, William D. Figg, and Ravi Madan. Lutetium-177-psma-617: A vision of the future. *Cancer Biology and Therapy*, 23, 2 2022.
 - [14] Susan J. Keam. Lutetium lu 177 vipivotide tetraxetan: First approval. *Molecular Diagnosis and Therapy*, 26, 5 2022.
 - [15] Ute Hennrich and Matthias Eder. [177lu]lu-psma-617 (pluvictotm): The first fda-approved radiotherapeutic for treatment of prostate cancer. *Pharmaceuticals*, 15, 10 2022.
 - [16] Sui Wai Ling, Erik de Blois, Eline Hooijman, Astrid van der Veldt, and Tessa Brabander. Advances in 177lu-psma and 225ac-psma radionuclide therapy for metastatic castration-resistant prostate cancer. *Pharmaceutics*, 14, 10 2022.
 - [17] Chalermrat Kaewput and Sobhan Vinjamuri. Update of psma theranostics in prostate cancer: Current applications and future trends. *Journal of Clinical Medicine*, 11, 5 2022.
 - [18] Mohammad S. Sadaghiani, Sara Sheikhabaei, Rudolf A. Werner, Kenneth J. Pienta, Martin G. Pomper, Michael A. Gorin, Lilja B. Solnes, and Steven P. Rowe. ¹⁷⁷lu-psma radioligand therapy effectiveness in metastatic castration-resistant prostate cancer: An updated systematic review and meta-analysis. *The Prostate*, 82, 3 2022.
 - [19] Viviane J. Tschan, Francesca Borgna, Sarah D. Busslinger, Martina Stirn, Josep M. Monné Rodriguez, Peter Bernhardt, Roger Schibli, and Cristina Müller. Preclinical investigations using [177lu]lu-ibu-dab-psma toward its clinical translation for radioligand therapy of prostate cancer. *European Journal of Nuclear Medicine and Molecular Imaging*, 49, 5 2022.
 - [20] Albert Jang, Ayse T. Kendi, Geoffrey B. Johnson, Thorvardur R. Halfdanarson, and Oliver Sartor. Targeted alpha-particle therapy: A review of current trials. *International Journal of Molecular Sciences*, 24, 7 2023.

- [21] Shahneen Sandhu Amir Iravani Michael S Hofman, Louise Emmett. Overall survival with [177lu]lu-psma-617 versus cabazitaxel in metastatic castration-resistant prostate cancer (therap): secondary outcomes of a randomised, open-label, phase 2 trial. *The Lancet Oncology*, 25, 1 2024.
- [22] Kanchi Patell, Matthew Kurian, Jorge A. Garcia, Prateek Mendiratta, Pedro C. Barata, Angela Y. Jia, Daniel E. Spratt, and Jason R. Brown. Lutetium-177 psma for the treatment of metastatic castrate resistant prostate cancer: a systematic review. *Expert Review of Anticancer Therapy*, 23, 5 2023.
- [23] Ju Cheon Lee, Keunho Byeon, Boram Song, Kyungeun Kim, and Jin Tae Kwak. Dior-vit: Differential ordinal learning vision transformer for cancer classification in pathology images. *Medical Image Analysis*, 105, 10 2025.
- [24] Clément Grisi, Kimmo Kartasalo, Martin Eklund, Lars Egevad, Jeroen van der Laak, and Geert Litjens. Hierarchical vision transformers for prostate biopsy grading: Towards bridging the generalization gap. *Medical Image Analysis*, 105, 10 2025.
- [25] Riqiang Gao, Lingfeng Li, Yucheng Tang, Sanja L. Antic, Alexis B. Paulson, Yuankai Huo, Kim L. Sandler, Pierre P. Massion, and Bennett A. Landman. Deep multi-task prediction of lung cancer and cancer-free progression from censored heterogenous clinical imaging. *arXiv*, 2019.
- [26] Antoine Rivail, Wolf-Dieter Vogl, Sophie Riedl, Christoph Grechenig, Leonard M. Coulibaly, Gregor S. Reiter, Robyn H. Guymer, Zhichao Wu, Ursula Schmidt-Erfurth, and Hrvoje Bogunović. Deep survival modeling of longitudinal retinal oct volumes for predicting the onset of atrophy in patients with intermediate amd. *Biomedical Optics Express*, 14, 5 2023.
- [27] Numan Saeed, Muhammad Ridzuan, Fadillah Adamsyah Maani, Hussain Alasmawi, Karthik Nandakumar, and Mohammad Yaqub. Survrnc: Learning ordered representations for survival prediction using rank-n-contrast. *arXiv*, 2024.
- [28] Jakob Wasserthal, Hanns-Christian Breit, Manfred Meyer, Maurice Pradella, Daniel Hinck, Alexander Sauter, Tobias Heye, Daniel Boll, Joshy Cyriac, Shan Yang, Michael Bach, and Martin Segeroth. Totalsegmentator: Robust segmentation of 104 anatomic structures in ct images. *Radiology: Artificial Intelligence*, 5, 07 2023.
- [29] Lemuel Puglisi, Daniel C. Alexander, and Daniele Ravi. Brain latent progression: Individual-based spatiotemporal disease progression on 3d brain mris via latent diffusion. *Medical Image Analysis*, 106, 12 2025.
- [30] Yundi Zhang, Paul Hager, Che Liu, Suprosanna Shit, Chen Chen, Daniel Rueckert, and Jiazhen Pan. Towards cardiac mri foundation models: Comprehensive visual-tabular representations for whole-heart assessment and beyond. *Medical Image Analysis*, 106, 12 2025.
- [31] Jana Fragemann, Lynton Ardizzzone, Jan Egger, Jens Kleesiek, and Mad Workshop. Review of disentanglement approaches for medical applications: Towards solving the gordian knot of generative models in healthcare. 3 2022.
- [32] Ashkan Abbasi and Amirhassan Monadjemi. Optical coherence tomography retinal image reconstruction via nonlocal weighted sparse representation. *Journal of Biomedical Optics*, 23, 3 2018.

- [33] Louisa Fay, Erick Cobos, Bin Yang, Sergios Gatidis, and Thomas Küstner. Avoiding shortcut-learning by mutual information minimization in deep learning-based image processing. *IEEE Access*, 11, 2023.
- [34] Daehoon Gwak, Gyuhyeon Sim, Michael Poli, Stefano Massaroli, Jaegul Choo, and Edward Choi. Neural ordinary differential equations for intervention modeling. *arXiv*, 2020.
- [35] Alistair Johnson, Lucas Bulgarelli, Tom Pollard, Steven Horng, Leo Anthony Celi, and Roger Mark. Mimic-iv. 2023.
- [36] Stav Belogolovsky, Ido Greenberg, Danny Eytan, and Shie Mannor. Individualized dosing dynamics via neural eigen decomposition. *arXiv*, 2023.
- [37] Steffen Wiewel, Moritz Becher, and Nils Thuerey. Latent-space physics: Towards learning the temporal evolution of fluid flow. *arXiv*, 2018.
- [38] Jared Quincy Davis, Krzysztof Choromanski, Jake Varley, Honglak Lee, Jean-Jacques Slotine, Valerii Likhoshesterov, Adrian Weller, Ameesh Makadia, and Vikas Sindhwani. Time dependence in non-autonomous neural odes. *arXiv*, 2020.
- [39] Matthew Ashman, Jonathan So, Will Tebbutt, Vincent Fortuin, Michael Pearce, and Richard E. Turner. Sparse gaussian process variational autoencoders. *arXiv*, 2020.
- [40] Göran Köber, Raffael Kalisch, Lara M.C. Puhlmann, Andrea Chmitorz, Anita Schick, and Harald Binder. Deep learning and differential equations for modeling changes in individual-level latent dynamics between observation periods. *Biometrical Journal*, 65, 3 2023.
- [41] Idris Bachali Losada and Nadia Terranova. Bridging pharmacology and neural networks: A deep dive into neural ordinary differential equations. *CPT: Pharmacometrics and Systems Pharmacology*, 13, 7 2024.
- [42] Rianne van den Berg, Leonard Hasenclever, Jakub M. Tomczak, and Max Welling. Sylvester normalizing flows for variational inference. *arXiv*, 2018.
- [43] Christa Cuchiero, Martin Larsson, and Josef Teichmann. Deep neural networks, generic universal interpolation, and controlled odes. *arXiv*, 2019.
- [44] Futoon M. Abushaqra, Hao Xue, Yongli Ren, and Flora D. Salim. Seqlink: A robust neural-ode architecture for modelling partially observed time series. *arXiv*, 2022.
- [45] Tobias Jorg, Moritz C. Halfmann, Gordon Arnhold, Daniel Pinto dos Santos, Roman Kloeckner, Christoph Düber, Peter Mildenerger, Florian Jungmann, and Lukas Müller. Implementation of structured reporting in clinical routine: a review of 7 years of institutional experience. *Insights into Imaging*, 14, 4 2023.
- [46] Ethan Sacoransky, Benjamin Y.M. Kwan, and Donald Soboleski. Chatgpt and assistive ai in structured radiology reporting: A systematic review. *Current Problems in Diagnostic Radiology*, 53, 11 2024.
- [47] Bourair Al-Attar. Network security in ai-based healthcare systems. *Babylonian Journal of Networking*, 2023, 11 2023.
- [48] Muhammet Alkan, Idris Zakariyya, Samuel Leighton, Kaushik Bhargav Sivangi, Christos Anagnostopoulos, and Fani Deligianni. Artificial intelligence-driven clinical decision support systems. *arXiv*, 2025.

- [49] Anita Khadka, Gregory Epiphanou, and Carsten Maple. Assuring intelligent medical devices: Ai assurance cases. *Studies in Health Technology and Informatics*, 6 2025.
- [50] Juan M. García-Gómez, Vicent Blanes-Selva, José Carlos de Bartolomé Cenzano, Jaime Cebolla-Cornejo, and Ascensión Doñate-Martínez. Functional requirements to mitigate the risk of harm to patients from artificial intelligence in healthcare. *arXiv*, 2023.
- [51] Ugoaghalam Uche James, Onuh Matthew Ijiga, and Lawrence Anebi Enyejo. Ai-powered threat intelligence for proactive risk detection in 5g-enabled smart healthcare communication networks. *International Journal of Scientific Research and Modern Technology*, 11 2024.
- [52] M. L. Zuley, R. Jarosz, B. F. Drake, D. Rancilio, A. Klim, K. Rieger-Christ, and J. Lemmerman. The Cancer Genome Atlas Prostate Adenocarcinoma Collection (TCGA-PRAD) (Version 4). The Cancer Imaging Archive, 2016. [Data set].
- [53] The ProCAncer-I project. ProstateNET - A large archive of anonymized data on prostate cancer. ProstateNET. Accessed: 2025-10-05.
- [54] Vibujithan Vigneshwaran, Erik Ohara, Matthias Wilms, and Nils Forkert. Macaw: A causal generative model for medical imaging. *arXiv*, 2024.
- [55] Sumedha Singla. Deep learning for medical imaging from diagnosis prediction to its counterfactual explanation. *arXiv*, 2022.
- [56] Aneesh Komanduri, Xintao Wu, Yongkai Wu, and Feng Chen. From identifiable causal representations to controllable counterfactual generation: A survey on causal generative modeling. *arXiv*, 2023.
- [57] Miguel Monteiro, Fabio De Sousa Ribeiro, Nick Pawlowski, Daniel C. Castro, and Ben Glocker. Measuring axiomatic soundness of counterfactual image models. *arXiv*, 2023.
- [58] Sumedha Singla, Motahhare Eslami, Brian Pollack, Stephen Wallace, and Kayhan Batmanghelich. Explaining the black-box smoothly- a counterfactual approach. *arXiv*, 2021.
- [59] Héloïse Bustin, Tom Meyer, Rolf Reiter, Jakob Jordan, Lars Walczak, Heiko Tzschätzsch, Ingolf Sack, and Anja Hennemuth. Elastonet: Neural network-based multicomponent mr elastography wave inversion with uncertainty quantification. *Medical Image Analysis*, 105, 10 2025.
- [60] Paul Friedrich, Yannik Frisch, and Philippe C. Cattin. Deep generative models for 3d medical image synthesis. 10 2024.