# Project Evaluation Criteria

### Technical Merit (30%)

- Algorithm sophistication
- Code quality and reproducibility
- Methodological rigor
- Performance metrics

### Innovation (25%)

- Novel approach or application
- Creative problem-solving
- Advancement over existing work

### Clinical Relevance (20%)

- Real-world applicability
- Clinical workflow integration
- Potential patient impact
- Stakeholder consideration

### Presentation Quality (15%)

- Clarity and organization
- Visual effectiveness
- Q&A engagement

**Documentation (10%):** README, code comments, technical report completeness, reproducibility instructions

## Technical Merit 30%

### Core Technical Components

🧠 Algorithm Design

💻 Code Quality

🔬 Methodology

📊 Performance

#### ▸ Algorithm Sophistication

Evaluation of the computational approach's complexity and appropriateness for the medical imaging problem. Considers whether advanced techniques (deep learning, ensemble methods, optimization algorithms) are properly applied.

✓ Good Example:

Implementation of a U-Net architecture with attention mechanisms for tumor segmentation, with proper justification for architectural choices based on dataset characteristics.

✗ Poor Example:

Using a simple linear classifier for complex multi-class diagnosis without exploring more sophisticated approaches or explaining the choice.

## ▸ Code Quality & Reproducibility

Assessment of code organization, documentation, version control, and ability for others to reproduce results. Includes proper dependency management, random seed setting, and clear execution instructions.

- ✓ Well-structured, modular code with clear function definitions
- ✓ Comprehensive comments explaining complex logic
- ✓ Requirements.txt or environment.yml for dependencies
- ✓ Fixed random seeds for reproducible results
- ✓ Clear README with setup and execution steps

## ▸ Methodological Rigor

Proper experimental design including appropriate train/validation/test splits, cross-validation when applicable, statistical significance testing, and handling of data imbalance or bias.

✓ Key Considerations:

Stratified k-fold cross-validation • Patient-level splitting (no data leakage) • Multiple evaluation metrics • Statistical significance testing • Ablation studies to validate design choices

### ▸ Performance Metrics

Selection and reporting of appropriate evaluation metrics for the medical imaging task. Should include sensitivity, specificity, AUC-ROC, F1-score, or other domain-specific metrics with proper interpretation.

| Sensitivity | Specificity | AUC-ROC |
| --- | --- | --- |
| **95.2%** | **92.8%** | **0.94** |

## Innovation 25%

### Innovation Dimensions

💡 Novel Approach

🎨 Creative Solutions

🚀 Advancement

## ▸ Novel Approach or Application

Evaluation of originality in applying techniques to medical imaging problems. This could involve adapting methods from other domains, combining techniques in new ways, or addressing previously unsolved problems.

✓ Innovative Examples:

• Applying self-supervised learning from natural language processing to medical image pre-training
• Using graph neural networks to model anatomical relationships in multi-organ segmentation
• Developing federated learning frameworks for privacy-preserving multi-hospital collaboration
• Creating interpretable AI models that highlight diagnostic features for radiologist review

## ▸ Creative Problem-Solving

Assessment of how well the team identifies and addresses challenges unique to medical imaging, such as limited labeled data, class imbalance, domain shift, or the need for explainability.

✓ Creative Solutions:

**Problem:** Limited annotated training data
**Solution:** Implementing synthetic data generation using GANs with medical expert validation, combined with few-shot learning techniques

**Problem:** Model interpretability for clinical adoption

**Solution:** Integrating attention maps and Grad-CAM visualizations that align with known anatomical structures

## ▸ Advancement Over Existing Work

Demonstration of improvements over baseline methods or state-of-the-art approaches. Requires thorough literature review and comparison with existing techniques, showing measurable advantages.

- ✓ Comprehensive literature review of current methods
- ✓ Clear baseline comparisons with standard approaches
- ✓ Quantitative improvement metrics (accuracy, speed, efficiency)
- ✓ Discussion of trade-offs and limitations
- ✓ Potential for broader impact beyond immediate application

⚠ Important Note:

Innovation doesn't always mean complexity. Sometimes the most innovative solutions are elegant simplifications that achieve comparable results with fewer computational resources or better interpretability.

# Clinical Relevance 20%

## Clinical Impact Framework

- Real-World Use
- Workflow Integration
- Patient Impact
- Stakeholders

### ▸ Real-World Applicability

Assessment of whether the solution can be practically deployed in clinical settings. Considers computational requirements, inference time, data availability, and regulatory compliance needs.

✓ Practical Considerations:

**Deployment Feasibility:** Model runs on standard hospital hardware within acceptable time constraints (e.g., inference < 30 seconds per image)

**Data Requirements:** Works with standard imaging protocols without requiring special acquisition parameters

**Robustness:** Performs consistently across different scanner manufacturers and imaging protocols

**Scalability:** Can handle typical hospital patient volumes (hundreds of scans per day)

### Clinical Workflow Integration

Evaluation of how seamlessly the solution fits into existing clinical processes. Considers integration with PACS systems, reporting workflows, and clinician decision-making processes.

- ✓ Compatible with DICOM standards and hospital IT infrastructure
- ✓ Provides results in clinically meaningful format
- ✓ Minimizes additional workload for medical staff
- ✓ Includes safeguards and uncertainty quantification
- ✓ Supports radiologist review and override capabilities

⚠ Critical Factor:

The best technical solution fails if it adds 15 minutes to each radiologist's workflow. Always consider time burden on clinical staff and design for efficiency.

### Potential Patient Impact

Analysis of how the solution could improve patient outcomes, reduce diagnostic delays, enable earlier intervention, or enhance treatment planning. Should identify specific clinical scenarios where impact is greatest.

✓ Impact Examples:

**Early Detection:** Automated screening for diabetic retinopathy enabling earlier intervention before vision loss

**Diagnostic Accuracy:** Reducing false negatives in mammography screening, catching cancers that might be missed

**Treatment Planning:** Precise tumor volume segmentation for radiation therapy planning, minimizing damage to healthy tissue

**Resource Optimization:** Triaging urgent cases for immediate radiologist review, reducing time to critical diagnoses

### ▸ Stakeholder Consideration

Recognition of multiple stakeholders including radiologists, clinicians, patients, hospital administrators, and regulatory bodies. Addresses concerns and needs of different groups.

| Radiologists | Patients | Hospitals |
|---|---|---|
| **Decision Support** | **Better Outcomes** | **Cost Efficiency** |

## Presentation Quality 15%

## Effective Communication Elements

📝 Clarity

🎯 Organization

📊 Visuals

💬 Engagement

### ▸ Clarity and Organization

Evaluation of how well the presentation conveys complex technical information in an understandable manner. Includes logical flow, clear explanations, and appropriate level of detail for the audience.

✓ Recommended Structure:

1. **Introduction (2 min):** Problem statement, clinical motivation, significance
2. **Background (2 min):** Related work, key challenges, your approach overview
3. **Methods (4 min):** Dataset, architecture, training procedure, innovation highlights
4. **Results (3 min):** Performance metrics, comparisons, visualizations
5. **Discussion (2 min):** Clinical implications, limitations, future work
6. **Conclusion (1 min):** Key takeaways, impact summary
7. **Q&A (6 min):** Address questions from evaluators

✓ Clear problem statement in opening

✓ Logical progression of ideas

✓ Technical terms explained when first introduced

✓ Smooth transitions between sections

✓ Concise summary of key findings

## ▸ Visual Effectiveness

Assessment of slide design, figure quality, and use of visual aids to enhance understanding. Includes appropriate use of diagrams, charts, example images, and highlighting of key results.

✓ Effective Visuals:

**Architecture Diagrams:** Clear, simplified network diagrams showing data flow and key components

**Results Visualization:** Side-by-side comparisons of input images, ground truth, and predictions

**Performance Charts:** Bar charts or line graphs comparing your method to baselines

**Attention Maps:** Heatmaps showing what the model focuses on for interpretability

**Clinical Impact:** Infographics showing workflow improvements or patient benefit scenarios

✗ Avoid:

Overcrowded slides with tiny text • Tables with too many numbers • Screenshots of code • Overly complex diagrams • Low-resolution images • Excessive animations • Reading slides verbatim

## ▸ Q&A Engagement

Evaluation of how well the team responds to questions, demonstrates understanding of their work, acknowledges limitations, and engages in technical discussion. Shows depth of knowledge beyond the prepared presentation.

- ✓ Listen carefully to questions before responding
- ✓ Provide direct, concise answers
- ✓ Admit uncertainty when appropriate rather than speculating
- ✓ Reference specific results or experiments when relevant
- ✓ Engage thoughtfully with critique or suggestions
- ✓ Show enthusiasm for the work and topic

> ✓ Strong Response Example:
>
> **Question:** "How would your model perform on data from different scanner manufacturers?"
>
> **Good Answer:** "That's an important consideration. Our training data came from Siemens scanners, and we haven't validated cross-manufacturer performance yet. We did implement test-time augmentation to improve robustness, but domain adaptation would be a key next step for clinical deployment. We'd propose collecting validation data from other manufacturers to quantify any performance drop."

## ▸ Presentation Tips

Professional delivery enhances the impact of your technical work. Practice timing, speak clearly, maintain appropriate pace, and show confidence in your methods and results.

| Timing | Team Dynamics | Delivery |
|---|---|---|

14 min talk
6 min Q&A

Balanced
Participation

Clear &
Confident

## Documentation 10%

### Documentation Components

README

Code Comments

Technical Report

Reproducibility

### ▸ README File

A comprehensive README is essential for others to understand and use your work. Should include project overview, setup instructions, usage examples, and key results summary.

✓ README Structure:

1. **Project Title and Description:** Clear summary of what the project does
2. **Table of Contents:** For easy navigation
3. **Installation:** Dependencies, environment setup, installation commands
4. **Usage:** How to run training, inference, and evaluation
5. **Dataset:** Data source, preprocessing steps, expected format
6. **Results:** Summary of key performance metrics
7. **Project Structure:** Overview of files and directories
8. **Citation:** How to cite your work if applicable
9. **License:** Usage rights and restrictions
10. **Contact:** Authors and contact information

## ▸ Code Comments

Well-commented code helps others understand implementation details and design decisions. Comments should explain WHY, not just WHAT the code does.

- ✓ Function/method docstrings with parameters and return values
- ✓ Explanation of complex algorithms or mathematical operations
- ✓ Rationale for hyperparameter choices
- ✓ References to papers for implemented techniques
- ✓ TODOs and known limitations clearly marked

⚠ Balance is Key:

Avoid over-commenting obvious code (e.g., "x = x + 1 # increment x"). Focus comments on non-obvious logic, design decisions, and important implementation details.

## ▸ Technical Report

A detailed technical document providing comprehensive information about the project, including methodology, experiments, results analysis, and discussions that extend beyond the presentation.

✓ Report Sections:

**Abstract:** Concise summary of problem, approach, and key findings (200-300 words)
**Introduction:** Background, motivation, clinical significance, related work
**Methods:** Detailed description of dataset, preprocessing, model architecture, training procedure
**Experiments:** Evaluation protocol, metrics, baseline comparisons, ablation studies
**Results:** Quantitative results, qualitative analysis, statistical significance
**Discussion:** Interpretation, clinical implications, limitations, ethical considerations
**Conclusion:** Summary of contributions and future directions
**References:** Properly cited literature and resources

## ▸ Reproducibility Instructions

Clear step-by-step instructions enabling someone else to reproduce your results from scratch. This includes environment setup, data preparation, training, and evaluation procedures.

- ✓ Exact Python version and library versions (requirements.txt)
- ✓ Hardware requirements (GPU memory, CPU cores)
- ✓ Random seeds for all stochastic operations

- ✓ Expected training time and computational costs
- ✓ Data download links or instructions (if publicly available)
- ✓ Pre-trained model weights or checkpoints
- ✓ Scripts for each stage (preprocessing, training, evaluation)
- ✓ Expected output format and file locations

**✓ Reproducibility Checklist:**

Can someone with similar hardware setup your environment in < 30 minutes? • Can they reproduce your main results within 10% accuracy? • Are error messages handled with clear guidance? • Is the expected runtime documented?

# Summary: Achieving Excellence

### ▸ Balanced Approach

The most successful projects demonstrate strength across all evaluation dimensions. While technical excellence is crucial, a technically sophisticated solution with poor clinical relevance or unclear communication will receive lower overall scores.

| Technical Merit | Innovation | Clinical Relevance |
| --- | --- | --- |
| **30%** | **25%** | **20%** |
| Foundation of Quality | Novel Contributions | Real-World Impact |

| Presentation | Documentation |
| --- | --- |
| **15%** | **10%** |
| Effective Communication | Reproducibility |

## ▸ Final Recommendations

- ✓ Start with a clear problem statement and clinical motivation
- ✓ Conduct thorough literature review before implementing
- ✓ Implement rigorous experimental methodology from the start
- ✓ Document as you go—don't leave it until the end
- ✓ Test reproducibility by having team members run your code
- ✓ Consider clinical workflow from the beginning, not as an afterthought
- ✓ Practice your presentation multiple times
- ✓ Prepare for critical questions about limitations and future work
- ✓ Ensure all team members understand all aspects of the project

✓ Allocate time for final polish and quality checks

**Remember:** The goal is not perfection in every criterion, but demonstrating strong competence across technical, innovative, clinical, and communication dimensions. Focus on doing a few things exceptionally well rather than everything mediocrely.