Translating image analysis methods into radiological practice – a challenge for implementation science

Current team:

* Rolf Heckemann
* Emilia Gryska
* Katka Cerna
* Birgit Heckemann
* Isabella Björkman-Burtscher
* Michael Schöll
* Meera Srikrishna

Considered input:

* Jonas Ivarsson
* Justin Schneiderman
* Isabella Björkman-Burtscher
* Ulf Nilsson (FIK)

Not in this document:

* Basic information
* Ethics
* Relevance
* Budget
* Administrating organisation
* Review panels
* Participating researchers
* CV
* Publications

[Instructions on Forte web site (PDF)](https://forte.se/app/uploads/sites/2/2019/12/implementation-research-2020-instructions-for-the-application-form-ta.pdf)

Level 1 headings: provided by Forte/PRISMA. Level 2 headings are ours

TOC for our overview (not for inclusion in proposal)

**Table of Contents**

[Purpose, research questions, theories, background and the originality of the project (max 20 000 characters) 2](#__RefHeading___Toc18441_3084900236)

[Synopsis 2](#__RefHeading___Toc18443_3084900236)

[For discussion 4](#__RefHeading___Toc18445_3084900236)

[Background 5](#__RefHeading___Toc18447_3084900236)

[Theories we will build upon 7](#__RefHeading___Toc18449_3084900236)

[Consolidated Framework for Implementation Research 7](#__RefHeading___Toc18451_3084900236)

[Participatory design 7](#__RefHeading___Toc18453_3084900236)

[Research questions 8](#__RefHeading___Toc18455_3084900236)

[New attempt 8](#__RefHeading___Toc18457_3084900236)

[Old version 8](#__RefHeading___Toc18459_3084900236)

[Theories 9](#__RefHeading___Toc18461_3084900236)

[Originality 9](#__RefHeading___Toc18463_3084900236)

[Keywords 9](#__RefHeading___Toc18465_3084900236)

[Study design, methods for data collection and analysis (max 15 000 characters) 9](#__RefHeading___Toc18467_3084900236)

[Interdisciplinary and/or multidisciplinary approach (max 2 500 characters) 11](#__RefHeading___Toc18469_3084900236)

[Gender and diversity perspectives in the content of the research (max 2 500 characters) 12](#__RefHeading___Toc18471_3084900236)

[Work plan (max 15 000 characters) 12](#__RefHeading___Toc18473_3084900236)

[References for the project description (max 7 000 characters) 12](#__RefHeading___Toc18475_3084900236)

[Gender distribution of the project group (max 500 characters) 12](#__RefHeading___Toc18477_3084900236)

[Image uploader 12](#__RefHeading___Toc18479_3084900236)

# Purpose, research questions, theories, background and the originality of the project (max 20 000 characters)

## Synopsis

The purpose of the project proposed here is to **bridge the wasteful divide between image analysis science and clinical diagnostic radiology**. The outcome will be that radiologists will better be able to **utilize research output** to achieve **the highest possible efficiency and quality**, and that patients will receive **more accurate diagnoses in a more efficient manner**.

The divide is evident through the fact that radiologists' diagnostic work consists mostly of looking at images, applying image interpretation skills, and reporting findings – essentially **working in a qualitative manner.** Automatic image analysis has next to no role in the radiological routine. Yet, scientific papers about medical image analysis methods appear at a staggering rate, mostly providing **quantitative information about the imaged subject’s health state** (2300 articles in 2019 [Source: PubMed search on *"Image Processing, Computer-Assisted/methods"[Mesh] AND 2019[pdat]*]). Methods are typically seen by the authors as having "clinical potential", and it appears natural to assume that **combining qualitative and quantitative data provides a more solid information base** for clinical decisions.

As long as the obstacles in the way of translating imaging science efforts into radiological practice persist, **substantial amounts of effort and resources are going to waste,** and **opportunities to reduce human suffering are missed.**

To **prevent this waste in the future,** and **to maximize the benefit to patients of advances in medical imaging science,** we need to bridge the divide between research and practice. Searching for causes and solutions is a **classical case for applying the tools of implementation research** and thus **matches the remit of the call**. Thanks to the expertise of our research team that covers radiology, imaging science, software engineering / artificial intelligence, participatory design, and implementation science, **we are in a unique position to systematically generate knowledge and procedures for connecting algorithm development with clinical practice.** We will formalize the procedures to maximize their usefulness in contexts outside of our exemplars, meeting the need, formulated by Eccles and Mittman (2006), to develop “methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice and, hence, to improve the quality and effectiveness of health services and care” [@ecclesWelcomeToImplementation2006].

The research we propose is centered on the following four specific aims.

1. To recognize and define opportunities to enhance radiologists’ work practice with machine learning / AI-based quantitative imaging capabilities
2. To develop algorithms for quantitative image analysis that are clinically useful in terms of accuracy and robustness
3. To generate interface and workflow designs that optimize the usability of the algorithms developed in Aim 1
4. To determine and match the requirements for making the tool acceptable to practitioners.

We will achieve these aims by developing exemplars: 1) for machine-learning based brain lesion segmentation on MR images, 2) for artificial-intelligence based detection of microscopic bleeding in brain MR images, and 3) for characterization of dementia disease using processing of CT images with artificial neural networks.

We will conduct this project following a principled approach, meaning that it will a) have a solid theoretical foundation in participatory design theory and the Consolidated Framework for Implementation Research; b) be circumspective – considering changes to the course of the project based on insights gained; c) be sustainable – aiming for long-term impact on clinical practice based on economically, ecologically, and socially responsible assessment and planning; and d) be participatory – including researchers, engineers, and radiologists (Tier 1 stakeholders). We will also consider the potential benefits of consulting further stakeholders (Tier 2 stakeholders: patients, neurologists, neurosurgeons, other healthcare staff, administrative staff).

## For discussion

Hypothesis: using participatory design, we can build software tools that aid the translation of evidence-based results from imaging science into radiological practice.

Call: “[Implementation] research takes its starting point in the complexity of operations, where many different and not infrequently contradictory values and goals must be considered” – we do this; just need to say so.

Call: “It is possible to remove some of the barriers to implementation by designing and conducting research projects in collaboration between researchers and different stakeholders, such as decision makers, professionals and users.” This is what we’ll do. Should we differentiate between core stakeholders (developers, designers, radiologists) and more peripheral ones (patients, neurologists, neurosurgeons)?

Call: “[Improvement research] takes its point of departure in a problem and its present state and adopts a multidisciplinary approach. The aim is to achieve quality improvements, often at the system level.” The synopsis is pretty clear about this. Let’s take this message forward into the rest.

Call: “[Policy implementation] includes questions raised about whether organisations are appropriately structured in terms of governance and control, whether there are relevant resources and sufficient knowledge in the organisation, and how collaboration, for example between different professions, affects the conditions for successful implementation.” Could be difficult for us, but needs mentioning

From the call:

* How can end-users and different stakeholders best be involved in the implementation process and how does this affect implementation outcomes? Which local and contextual factors impact implementation processes and their outcomes?

We will design an implementation process *ad hoc* and use it for the exemplars – that way we’ll find the answers. Perhaps work on the exemplars one after the other, updating and improving the process at each step!

* How can the management’s role in facilitating organisational change be highlighted?

Involve managers at a certain point of the design process. We will group stakeholders into two tiers. Tier 1 are radiologists, engineers, and designers. Tier 2 are managers, patients, and possibly others who have not entered our awareness yet. The project design will primarily consider Tier 1 stakeholders. Tier 2 will be involved when we gain insights on how their inclusion will advance the project.

* How should goal and value conflicts be managed in order to avoid implementation problems?

Are there answers to this in participatory design theory? According to PD, participants should be brought together during the design process, creating a space where values and conflicts are discussed and (hopefully) overcome through mutual understanding.

* What dimensions of power are there in implementation and what role do they play in successful implementation?

Is there someone who is in better position to be able to make decisions than other people? For example, we can design this wonderful tool that would perfectly fit to the needs of the clinician, but if the manager will not approve (or actually understand) why this is important, they will not give it a final okey. Part of PD is basically "educating" people about the particular needs of the participants, not only the radiologists but also the managers. for example.

* What role does the existence of different interests and professional norms play, and can they facilitate or constitute obstacles in the implementation process?

Radiology as a specialty has survived many usurpation attempts by other specialties. Does this play a role? Are radiologists worried about being replaced by software?

* How can studies be designed that integrate knowledge from implementation science, improvement science and research on policy implementation?

We’ll learn by doing, right?

* What types of outcomes are most relevant to capture during implementation and in what time frame should they be studied?

Not sure what they mean

Call: **“It is a requirement that the implementation problem is the primary object of study in the application.”**

## Background

The impetus for this proposal developed during the analysis of a scoping review we carried out on the subject of automatic brain lesion segmentation from magnetic resonance (MR) images [@gryskaAutomaticBrainLesion2019]. Brain lesions are localized abnormalities that can arise from a variety of diseases and conditions, for example cancer, stroke, or inflammation. A striking feature of the literature landscape is that there is an abundance of methods, but little discussion of implementation challenges, and a complete absence of reported practice interventions based on proposed new lesion segmentation algorithms. Most articles claim that the proposed methods have clinical potential, but such claims generally rest on validation measures (e.g. agreement between automatic and manually prepared segmentations) that are not necessarily clinically relevant. This illustrates the divide between algorithm development and clinical practice. In this divide, both resources and opportunities to reduce human suffering go to waste. Bridging the divide would enable better targeting and stronger clinical relevance of algorithm development, and it would improve radiologists’ access to research output. In consequence, the quality and efficiency of radiologists’ clinical decision-making would improve. It is unclear what the main factor or factors are that contribute to the divide; quantifying the magnitude of those contributions is even further out of reach. Accordingly there is no basis for deciding what measures would help to bridge it. This is the challenge that we propose to tackle.

We intend to build on our previous work and produce an exemplar by developing a previously described brain lesion segmentation algorithm into a practically applicable tool, and integrating it into an existing radiological workflow and evaluating its benefits in terms of effectiveness, cost savings, time savings, and improved livelihoods. Validation will be integrated as a concurrent part of the design process (rather than as a single-time point assessment), using measures that go beyond accuracy and include robustness, . By following a principled approach, we will generate findings – i.e. knowledge about barriers and potential solutions on the implementation path – that will be at least in part be generalizable and can serve as guidance for designing image analysis procedures while keeping clinical applicability in focus.

The proposed project will build on our previous work on segmentation of MR images of the human brain. The principal investigator RAH has made fundamental contributions to the field by developing and validating MAPER and Pincram software for automatic atlas-based brain image segmentation into anatomical regions. The functionality of this software could be enhanced by enabling users to additionally measure and classify focal lesions (tumour, injury, stroke, etc.). To obtain a systematic overview of the field of brain lesion segmentation, participating researcher EAG has used the rigorous and well-defined scoping review methodology [(Arksey & O’Malley, 2005)](https://doi.org/10.1080/1364557032000119616) to develop a study protocol [(Gryska et al., 2018)](https://bmjopen.bmj.com/content/9/2/e024824?utm_content=consumer&utm_medium=cpc&utm_source=trendmd&utm_campaign=bmjo&utm_term=usage-042019); the full article is currently in peer review (TODO: check if true at time of submission). The analysis highlighted a crass imbalance: while plenty of lesion segmentation methods have been proposed, and most have “clinical potential” according to authors’ claims, none of articles in our scoping review went on to describe efforts to enable clinicians to use any of these methods in practice. Method validation is often restricted in ways that are unsuitable for supporting clinical implementation. For example, a typical validation strategy is to apply the method to contest data, i.e. collections of brain images with lesions that have been manually delineated, with the manual delineation serving as the criterion standard [(Menze et al. 2014)](https://doi.org/10.1109/TMI.2014.2377694). This kind of challenge bears little resemblance to any clinical situation: to begin with, a clinical image may not contain a lesion at all. Most articles do not report on the behaviour of an algorithm in the absence of a lesion. The critical path towards sensible workflow modifications enabling the use of advanced image analysis in the routine practice of image interpretation is underexplored.

To us, this situation is a challenge and an opportunity. Implementation research offers tools and concepts that will help us identify the nature of the problem and its determinant factors. By developing an exemplar using participatory design involving radiologists as clinical end users, we will bridge the divide and map the process of implementing image analysis technology to improve clinical radiology workflows, to the ultimate benefit of patients, who will receive more accurate diagnoses in a more timely fashion than currently possible.

%%% Isabella: what kind of clinical questions does brain imaging currently answer? What is the role of lesions in brain imaging? What clinical challenges would an accurate, robust, and well-validated lesion segmentation tool address?

Present-day algorithm development lacks clinical relevance because developers often fail to consider the structure of clinical processes. For example, many brain lesion detection algorithms are built on sample data that consist exclusively of images with tumours. The software is thus developed without an awareness of the possibility that the image to be analysed does not contain a lesion at all. This lack of awareness entails the following disadvantages in our view: first, the software cannot be designed with a view to establishing absence of abnormalities, even though such a capability could be immensely helpful in clinical practice; second, the software will be prone to false positives, ie. it will delineate a lesion where there is none.

### Theories we will build upon

To fulfill the purposes of the project, we will draw on participatory design (PD) and the Consolidated Framework for Implementation Research (CFIR) [cite] to ensure successful and sustainable uptake of the tool and to provide a structured description of factors that facilitate or hinder the implementation process.

The nature of our project fits well with the PD approach. PD builds on the idea that diverse stakeholders are involved in the design of a tool of which they will be future users. An important aspect of PD is sustainability in the sense of achieving a lasting effect on professional practice in the target setting. Based on the findings from the review, there may be huge potential for developing and implementing a relevant, suitable, and sustainable automatic brain lesion segmentation tool using the PD approach. One of the biggest challenges in PD research, however, is generalizability of the results and translating a "case study" results to a broader context (systemic uptake?). To make our findings generalizable and to comply with implementaiton research (IR) frameworks, we will complement/combine the PD with the CFIR framework.

Various frameworks and models have been proposed for studying intervention implementation. In general, these theories, frameworks, and models can be classified into three categories (Nielsen et al., 2015, Impl Sci 10, 53), depending on which aspect of the implementation a given theory provides guidelines for. The categories are:

* the process of "translating research into practice"
* the factors explaining and impacting implementation
* the evaluation of the implementation.

#### Consolidated Framework for Implementation Research

CFIR falls under the second category and provides, as the name suggests, a consolidated framework of constructs that should be addressed in implementation studies. The most relevant constructs will be chosen that fit to our problem and setup. The process and evaluation of the implementation process will be organized according to the sustainable PD approach. The approach will be complemented by the CFIR framework as a reference of the constructs we will address, processes we will apply, and impact they have on the implementation problem. This approach will help us systematize to some extent a dynamic (?) and complex problem and propose a set of guidelines for implementation of algorithm-based decision support tools for radiologists.

#### Participatory design

Participatory design builds on the idea that diverse stakeholders are involved in the design of a tool of which they will be future users.

PD builds on active participation of stakeholders such as design specialists, development specialists and domain specialists [(Bratteteig & Wagner, 2014)](https://paperpile.com/c/qafUON/svIq). This approach emphasises the active role of all the participants. However, participation in the design process alone is not enough. It is necessary that the practitioners (domain experts) can actually make informed decisions during the design process. For this, they need appropriate support [(Bratteteig & Wagner, 2016)](https://paperpile.com/c/qafUON/2S4Y).

A key feature of the PD approach is therefore mutual learning, where the design participants learn about the other domains to be able to work on a common solution [(Kensing & Blomberg, 1998)](https://paperpile.com/c/qafUON/nHj6). In addition, participating in PD requires soft skills, such as reflection: domain experts often have to reflect about their own work when participating in PD – not all the workflows are formalized, and they have to be formalized during the PD process.

An important aspect of PD is sustainability in the sense of achieving a lasting effect on professional practice in the target setting (Meurer et al. 2018). To this end, we will take the iterative nature of software development into account. Importantly, we will continually assess the demands that the new tool imposes on practitioners to change their workflow and develop skill in its use [(Heath & Luff, 2000)](https://paperpile.com/c/qafUON/j0xw).

During the design process of a tool supporting a workflow, the workflow has to be dismantled, the “black box” of the workflow has to be opened.

...

## Research questions

### New attempt

The overarching question – How can we utilize research results from imaging science to achieve improvements of quality and efficiency in clinical radiology? – breaks down into the following component questions:

1. How can radiologists best be involved in the process of implementing automatic image analysis methods into clinically useful software tools?
2. Which goal and value conflicts between radiologists and developers are exposed in the process? What are the results of applying the means that participatory design theory proposes for managing such conflicts?
3. What role does the existence of different interests and professional norms play as developers and radiologists collaborate in the implementation process?
4. How can studies be designed that integrate knowledge from implementation science, improvement science, and research on policy implementation? (unchanged from call)

### Old version

Our research questions pertain to four stages of implementation.

1. Low-level development and validation: How to develop an algorithm for automatic brain lesion segmentation and assess accuracy and robustness for clinical applicability?
2. Usability and interface: What would be a suitable interactional design to optimize usability?
3. Organization and work environment: What are the requirements for the integration of the tool into existing workflows?
4. Wider context: Are the tool and implementation process acceptable and sustainable, considering structural, political, cultural (and potentially other) context factors?

## Theories

TODO: What is our theoretical contribution?

## Originality

%%% Relationship with “Professional trust and autonomous systems”

%%% Rolf, Emilia

# Keywords

%%% 1-5 keywords

%%% Emilia: frequent MeSH items from scoping review sample? two to three

%%% Plus two to three pertaining to the domain and research approach

# Study design, methods for data collection and analysis (max 15 000 characters)

Our study will use the *design case study* methodology, [@wulfPracticeBasedComputingEmpirically2015], complemented by elements of implementation research (CFIR).



Schema design case study

(from Practice-based Computing: Empirically-grounded Conceptualizations derived from Design Case Studies; in: Wulf, V.;Schmidt,K.; Randall,D.(eds): Designing Socially Embedded Technologies in theReal World, Springer, London, 2015, 111 - 150)

**Phase I: Understanding the needs and challenges of radiologists’ work**

The first phase focuses on understanding the needs and the challenges that the radiologists experience in their current state of work, that is before the deployment of the algorithm-based tools.

The methods for data collection would involve observations, interviews and short surveys. Observations and interviews will be mainly of qualitative character, and their aim is to understand the existing practice from technological, organizational and social perspective. Hence, radiologists but also other relevant stakeholders (such as managers) will be interviewed. A series of surveys based on the concepts from the implementation research (CIFR) will be developed to gain deeper understanding of the radiologist work.

The observations will be documented by writing down field notes, audio and video recording relevant moments (when relevant and possible), collecting relevant documents and taking photographs. The interviews will be audio recorded and transcribed.

The qualitative data will be analyzed through thematic analysis.

**Phase II: Algorithm-based tool development**

During the second phase, the actual tool will be developed.

A series of meetings will be organized during which technical requirements will be collected. During these meetings, possible designs will be iterated and multiple prototypes of varied fidelity developed. Relevant stakeholders will be invited in order to ensure that their input is considered but also as a way to tackle the possible conflicts or differences in their values. During the meetings, the developers together with the researchers, radiologists (other practitioners?) and other relevant stakeholders will gradually co-design a fully functional prototype.

The development process will be documented through writing down field notes and audio recordings. Audio recordings will be transcribed.

The qualitative data (field notes and audio recordings) will be analysed through thematic analysis.

At the end of this phase, another series of surveys based on the concepts from the implementation research (CIFR) will be developed to gain deeper understanding of the radiologists’ involvement in the tool development.

**Phase III: Algorithm-based tool appropriation**

During the third phase, the prototype will start being used in the clinical practice of the radiologists. This phase can be viewed still as a part of the development process, since during the use of the tool, additional information will be collected and the tool possible refined. The aim of this phase is not only to fix technical bugs but also to ensure the tool fits into the clinical practice. Hence, the radiologists will be supported in learning how to use the tool in a meaningful way for their practice and how to create and sustain new workflows connected to the tool.

This phase will be documented by another series of observations and interviews. In a similar manner as during the first phase, the observations will be documented by writing down field notes, audio and video recording relevant moments (when relevant and possible), collecting relevant documents and taking photographs. The interviews will be audio recorded and transcribed.

The qualitative data will be analyzed through thematic analysis.

The results from the pre-study and the appropriation part of the study will be compared and reflected over together with the practitioners.

Finally, another series of surveys based on the concepts from the implementation research (CIFR) will be developed to gain deeper understanding of the radiologists’ involvement in the tool development.

# Interdisciplinary and/or multidisciplinary approach (max 2 500 characters)

A team composed of experts from education, imaging scientists/engineers, designers and clinical practitioners, in a unique position to tackle the challenges of the project …

# Gender and diversity perspectives in the content of the research (max 2 500 characters)

Brain imaging is applied in a broad variety of diseases that, taken together, affect all genders, races, and social circles. Research results will be approximately equally relevant across gender, ethnic, and socioeconomic groups represented in Sweden.

…

# Work plan (max 15 000 characters)

In a participatory approach, we will determine clinicians wishes and needs for a automatic brain segmentation tool Developing a prototype of the tool Pilot testing the tool in clinical practice Assessing the acceptability and sustainability of the tool and implementation process considering structural, political, cultural context factors

%%% Package structure, Gantt chart …

At the reporting stage, we will seek to follow StaRI guidelines and checklist (27 items) to the extent that they are transferrable from the original target domain (public health) to our domain. https://www.bmj.com/content/356/bmj.i6795

# References for the project description (max 7 000 characters)

# Gender distribution of the project group (max 500 characters)

The project group has a balanced gender distribution.

# Image uploader

Gantt chart