BMJ Open

Automatic lesion segmentation on standard magnetic resonance images of the human brain: a scoping review protocol.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-024824
Article Type:	Protocol
Date Submitted by the Author:	15-Jun-2018
Complete List of Authors:	Gryska, Emilia; Goteborgs universitet Institutionen for kliniska vetenskaper, Avdelningen för radiofysik Schneiderman, Justin; Goteborgs universitet Institutionen for neurovetenskap och fysiologi, Sektionen för klinisk neurovetenskap Heckemann, Rolf; Goteborgs universitet Institutionen for kliniska vetenskaper, Avdelningen för radiofysik
Keywords:	Magnetic resonance imaging < RADIOLOGY & IMAGING, Scoping review, Protocol, Brain lesions, Segmentation

Automatic lesion segmentation on standard magnetic resonance images of the human brain: a scoping review protocol.

Emilia A. Gryska [1,2*], Justin F. Schneiderman[1,3], Rolf A. Heckemann[1,2]

1 MedTech West at Sahlgrenska University Hospital, University of Gothenburg, Sweden

2 Department of Radiation Physics, Institute of Clinical Sciences, University of Gothenburg, Sweden

3 Department of Clinical Neurophysiology, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Sweden

* Corresponding author: Emilia A. Gryska

Postal address: MedTech West, Röda stråket 10B, Sahlgrenska University Hospital, 413 45 Göteborg

E-mail address: emilia.gryska@medtechwest.se

Telephone number: +46 720304106

Keywords: Scoping review, Protocol, Brain lesions, Segmentation, Magnetic resonance imaging

Word count: 2,559

ABSTRACT

Introduction Automatic brain lesion segmentation from medical images has great potential to support clinical decision-making. Although numerous methods have been proposed, significant challenges must be addressed before they will become established in clinical and research practice. A rigorous and structured review of available methods will elucidate the state of the art and provide a synopsis of competing approaches to automatic brain lesion characterization.

Methods and Analysis We present the background and study design of a scoping review for automatic brain lesion segmentation methods for structural magnetic resonance imaging (MRI) according to the framework proposed by Arksey and O'Malley. We aim to identify common image processing steps as well as mathematical and computational theories implemented in these methods. We will aggregate the evidence on the efficacy and identify limitations of the approaches. Methods to be investigated work with standard MRI sequences from human patients examined for brain lesions, and are validated with quantitative measures against a trusted reference. PubMed and IEEE Xplore will be searched using carefully composed search phrases that will ensure an inclusive and unbiased overview. For matching records, abstracts will be screened to ensure eligibility. Studies will be excluded if a full paper or translation is not available, if non-standard MR sequences are used, if there is no quantitative validation, or if the method is not automatic. In the data charting phase, we will extract information about authors, publication details, and study cohort. We expect to find information about preprocessing, segmentation, classification methods, and validation procedures. We will develop an analytical framework to collate, summarize, and synthesise the data.

Ethics and Dissemination Ethical approval for this study is not required since the information will be extracted from published studies. We will submit the review report to a peer-reviewed scientific journal and explore other venues for presenting the work.

ARTICLE SUMMARY

Strengths and limitations of this study:

- This study will be the first scoping review mapping the approaches to automatic brain lesion segmentation and classification on MR images.
- We will present the state of the art and a synopsis of competing automatic brain lesion segmentation methods.
- Our study design ensures an inclusive and unbiased review while maintaining good quality of the gathered sources by proposing the requirement of a quantitative validation of the presented methods.
- We will validate our search strategy by comparing the bibliographies and citations of the most recent and most cited records with the gathered sources.
- We may have to impose a limit to the number of selected papers to keep the study feasible and exclude the earliest papers.

INTRODUCTION

In clinical practice, diagnosis of brain lesions is based on the patient's history, clinical presentation, visual assessment of appointed scans, and other laboratory examinations. Magnetic resonance imaging has become an important tool in brain lesion identification and classification due to its ability to produce images with high contrast resolution and sensitivity for abnormalities. Various conditions can give rise to such lesions. The most common causes include trauma, inflammation and autoimmune diseases, stroke, malignant or benign tumours, and infections[1]. Although brain lesions tend to appear significantly different from healthy tissues on MR scans, differentiating between brain lesion causes based on visual examination can be difficult or impossible. Still, visual interpretation is still the most common and trusted mode of image analysis in clinical practice. Accurate identification and delineation of lesion boundaries is particularly important in treatment planning for surgery or radiation therapy in tumor patients as well as determination of disease burden, prognosis, and therapy response in nearly all types of brain lesions. The process is currently commonly performed manually by an expert rater. The procedure is tedious, time-consuming, and subject to inter-rater as well as test-retest variability.

Automatic image segmentation methods promise to reduce or eliminate subjective decisions in this process, facilitating fast and accurate delineation of lesions on MR brain images. Although many automatic brain lesion segmentation methods have been proposed, substantial challenges remain, for example the variable appearance

of the lesions on MR images due to unknown, possibly biological factors; differences of image acquisition protocols between centers; and the difficulty of validating such algorithms on sufficiently large case numbers. Taken together, these challenges explain why no single tool or approach has thus far been adopted in clinical or even in research practice. On surveying the literature on automatic brain lesion segmentation methods in an ad-hoc, preliminary fashion, we recognized the need for a rigorous and comprehensive review. A formalized approach to reviewing literature in this manner is the scoping review as proposed by Arksey and O'Malley[2]. Using this framework along with refinements by Levac et al.[3] and Colquhoun et al.[4], as well as elements of the PRISMA and PRISMA- P guidelines[5,6], we will map key concepts, converging developments, challenges, and promising new research avenues. The purpose of publishing the research plan at this stage is to document our objectives openly, to invite comments and suggestions, and to enhance the rigor of our study. This open documentation will compel us to follow the plan and justify any deviation. We believe that being fastidious in this manner will enhance the value of the research once completed.

Study aims and objectives

We have identified the following aims and objectives of the scoping review on existing automatic brain tumor segmentation techniques on conventional magnetic resonance images:

1. perform a comprehensive search and gather available evidence;
2. analyze, synthesize, and summarize the identified methods;
3. share the findings with the stakeholders;
4. identify common challenges, weaknesses and controversies, as well as unaddressed issues which can signify opportunities for future work to improve segmentation methods.

METHODS AND DESIGN

General

This section describes how each of the scoping study stages identified by Arksey and O'Malley[1] will be applied to the present study. The resulting draft protocol will be refined throughout the process of conducting the study. All changes will be documented in detail in a project diary and justified in the scoping review document to ensure reproducibility of the study. In this study we will balance the breadth of the included studies with the depth of the analysis of reviewed methods. The nature of the researched topic imposes certain logical limitations which, together with the inclusive nature of a scoping review, will help to create a focused yet comprehensive overview of the topic.

Identifying research questions

We have defined the following research questions that will be addressed in this study.

1. Which common image processing steps are necessary for automatic brain lesion segmentation on MR images?
2. Which mathematical and computational theories are most commonly applied in which types of brain lesions?
3. What is the efficacy of existing implementations?
4. What are the limitations of those methods and issues that should be addressed in future studies to develop a tool that is suitable for clinical use?

Identifying relevant studies

- Eligibility criteria

We established the following initial criteria for the proposed automated image processing methods to be eligible for inclusion in the scoping review. A method must be applied to one or a number (multi-modal) of commonly acquired structural MRI sequences (T1w, T2w, PDw, FLAIR, DWI) from human subjects investigated for a condition, other than primary neurodegeneration, known to cause brain lesion(s) in order to delineate, identify,

and classify these lesion(s).

The proposed methods should be validated with a manual expert segmentation of the lesions. The efficacy of the method should be reported providing quantitative scores (such as sensitivity, specificity, overlap, surface distance, volume error, or other measures of similarity, etc.). Alternative validation approaches will be considered if they have face validity. We surmise that the presence of a validation step with well-defined quantitative performance measures is an important inclusion criterion for the study despite the risk of excluding a number of records that do not provide such information, yet contain information that could, in principle, promote our objectives. A thorough validation is a necessary step in developing medical image segmentation methods. Even though we dispense with formal quality assessment of the included studies, we believe that our principled approach will enable us to provide a valuable report on the researched topic.

- Initial search

Eligible studies will be retrieved from peer-reviewed journal articles and conference papers, including review papers. Since the nature of the review topic deals with a method description rather than an intervention outcome, we presume that any other research publication types than stated in the protocol contain duplicated information or ineligible evidence. Therefore, the publication type limitation should neither substantially increase the risk of bias of the review nor limit the number of records retrieved during the screening and selection phase. We will not impose any date of publication limitations in the initial search. The authors assume that all the records that will be found during the search phase were published after MR imaging had became commonplace in clinical practice. If any earlier records should be found, we expect them to be excluded based on the eligibility criteria.

The search will be conducted using search terms constructed in English. We expect to find some papers published in other languages whose titles and abstracts were translated to English. If any of such papers will be eligible for whole-text screening, we will explore means to obtain a translation of the paper.

Two online databases will be searched: PubMed and IEEE Xplore. The following search phrases will be constructed using non-controlled vocabulary to initialize the search. An advanced search in publication metadata will be conducted in both databases using the following search terms to identify potentially relevant sources (asterisk indicates a wildcard character to account for variations in the spelling of the search terms):

1. automat(ic)*
2. AND brain
3. AND lesion OR tumor OR neoplasm
4. AND segment(ation)* OR identif(ication)* OR delin(eation)* OR classif(ication)*
5. AND mri OR mr.

The search results will be exported for both databases. From the controlled vocabulary tags (MeSH in PubMed; IEEE terms and INSPEC terms in IEEE Explore), we will build frequency tables. The most common relevant terms will be added to the original search phrase to refine the search. Combining free text and index terms ensures high sensitivity search of the relevant studies. The results will be refined by applying possible limitations defined in the eligibility criteria depending on the availability in the search engine, such as document type (Journal and conference articles) or species (Human).

A separate search will be conducted in PubMed to identify potentially eligible papers that have not been indexed with MeSH yet. MeSH terms are assigned by specialists at the National Library of Medicine after a variable delay, meaning that some recent papers lack them. We will modify the search phrase and look for the search terms as well as MeSH terms in all fields, and an additional status criterion will be added to exclude MeSH indexed papers.

We will screen bibliographies of the most recent papers as well as citations of the most cited papers and compare it with the existing sample to and evaluate inclusiveness and validate the proposed search strategy. We will identify the most cited studies by dividing the number of citations of a given paper by the number of months since the publication. If there is a substantial mismatch between the existing set of selected studies, and the bibliographies and citations, we will identify additional sources by screening the bibliographies and citations of the identified set.

- Screening

The records found in the search phase will be screened to exclude irrelevant or otherwise ineligible items. The screening will be performed by finding the key terms or their synonyms in the publication title or abstract and determining if the publication is relevant. Articles that do not name in the title or the abstract or refer to any proposed method of any form of identification of any type of brain lesion will be excluded under the assumption that those papers either do not contain enough evidence for the method to be eligible for the synthesis, or do not propose a lesion segmentation and classification method. The following key terms (and their synonyms) will be considered:

1. method
2. identification
3. brain lesion
4. magnetic resonance imaging

At the screening stage, records will be excluded if a given study has previously been reported or any of the study characteristics stated in the abstract clearly do not match the eligibility criteria for this study.

Study selection

For items selected during the screening stage, full-text articles will be retrieved. The following criteria will be considered as a reason for exclusion of a paper from the review:

- full paper not available
- translation not available
- modality other than structural MRI used
- no quantitative validation found
- semi-automatic method proposed.

The terms automatic and semi-automatic segmentation do not have a widely accepted definition. For our purposes, a semi-automatic method shall be one that relies on expert's decisions during the segmentation process, while an automatic method is one that requires a user to provide (possibly preprocessed) images and launch the program, after which all decisions regarding lesion segmentation and characterization are made without human interaction.

An artificial limit of the number of papers included in this study may be imposed after completing the study selection phrase. If the study becomes unfeasible due to the number of selected papers, we will decide to include only a portion of the original sample and exclude earliest publications.

Data charting

In the data charting phase, the following study information will be extracted from every eligible record: author(s), year of publication, country of origin, and funding information (if available). We will chart the demographic information of the patients and MRI sequence(s) used in a given study as well as the type of brain lesion(s) studied. Based on our knowledge in the field of medical image segmentation we expect to find information on the following main categories in the method description[7]:

1. Image preprocessing (e.g. registration, skull stripping, intensity inhomogeneity correction, noise reduction, intensity normalization)
2. Segmentation (e.g. supervised and unsupervised)
3. Classification (e.g. types of features used; additional features, such as an ascribed tumor grade; differentiation between tumor and peripheral reactive changes (edema); differentiation between internal tumor components (tumor proper, tissue invasion, necrosis etc.)
4. Validation (e.g. amount and types of reference data; accuracy and reliability measures).

We will identify the enumerated as well as additional categories or lack thereof in each study and, if possible, subdivide each category into appropriate subgroups. We are aware that the proposed classification may turn out to be impractical and that modifications may be necessary for a well-structured and thorough analysis. To test the proposed approach we will conduct a pilot charting on a subset of studies (selected by citation numbers) to

evaluate and refine the charted variables.

Collating, summarizing, and reporting the data

Upon collecting of the eligible studies we will develop an analytical framework to collate, summarize, and synthesise the data. We will make use of summary counts and tables to provide quantitative information on the body of research on automatic brain tumor segmentation methods. While analysing the data, common procedures for the methods, types of lesions and their outcomes will be identified. We will also try to recognize discrepancies between the analysed methods and use that information to address Objective 4 (cf. Study aims and obejctives). The consultation stage of the scoping review, described in the following section, will contribute to fulfilling that objective. We allow for the possibility to adjust or expand the initial analytical framework after the consultation stage to present the gathered information according to the stakeholders' requests.

Consultation

Although the consultation stage is currently considered optional in scoping reviews, we see advantages in including this stage in our study. We will use this opportunity to share preliminary findings and refer to potential stakeholders to gain more insight into our data from different perspectives. The consultation will be conducted using a questionnaire or through interviews, however a detailed design of the consultation process will be created after finishing collating, summarizing, and internal reporting the data.

ETHICS AND DISSEMINATION

The study will contain information gathered from already published papers therefore it does not require ethical approval. We will distill the project diary (cf. General) into a review report for submission to a peer-reviewed scientific journal. In addition, we will seek to present the study at scientific conferences. Following on from the work done at the consultation stage, we will identify stakeholders outside of academia and seek to disseminate the results to them in appropriate formats (trade journal articles, lectures, laypersons' summaries, press releases).

Author Contributions: RAH conceived the idea of conducting the scoping review. EAG and RAH developed all the elements of the study design. EAG lead writing of the protocol. RAH was the main supervisor of the process of writing and JFS provided feedback on the methodology of the study as well as the manuscript. All authors approved the final version of the manuscript.

Funding statement: JFS is funded by the Knut and Alice Wallenberg Foundation grant KAW2014.0102, The Swedish Research Council grant 2017-00680, and the Swedish Childhood Cancer Foundation grant MT2014-0007.

Competing interests: None declared.

Data sharing : Data sharing not applicable to this manuscript as no datasets were generated or analysed.

REFERENCES

1. Brain lesions Causes [Internet]. Mayo Clinic. Available from: http://www.mayoclinic.org/symptoms/brain-lesions/basics/definition/sym-20050692 (accessed 2018 Jun)

2. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol*. 2005 Feb 1;8(1):19–32.

3. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci.* 2010 Sep 20;5:69.

4. Colquhoun HL, Levac D, O'Brien KK, et al. Scoping reviews: time for clarity in definition, methods,

and reporting. *J Clin Epidemiol*. 2014 Dec 1;67(12):1291–4.

5. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *PLoS Med*. 2009 Jul 21;6(7):e1000100.

6. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015 Jan 1;4:1.

7. García-Lorenzo D, Francis S, Narayanan S, et al. Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. *Med Image Anal*. 2013 Jan;17(1):1–18.

BMJ Open

Automatic brain lesion segmentation on standard magnetic resonance images of the human head: a scoping review protocol.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-024824.R1
Article Type:	Protocol
Date Submitted by the Author:	23-Nov-2018
Complete List of Authors:	Gryska, Emilia; Goteborgs universitet Institutionen for kliniska vetenskaper, Avdelningen för radiofysik Schneiderman, Justin; Goteborgs universitet Institutionen for neurovetenskap och fysiologi, Sektionen för klinisk neurovetenskap Heckemann, Rolf; Goteborgs universitet Institutionen for kliniska vetenskaper, Avdelningen för radiofysik
Primary Subject Heading:	Radiology and imaging
Secondary Subject Heading:	Radiology and imaging
Keywords:	Magnetic resonance imaging < RADIOLOGY & IMAGING, Scoping review, Protocol, Brain lesions, Segmentation

Automatic brain lesion segmentation on standard magnetic resonance images of the human head: a scoping review protocol.

Emilia A. Gryska [1,2*], Justin F. Schneiderman[1,3], Rolf A. Heckemann[1,2]

1 MedTech West at Sahlgrenska University Hospital, University of Gothenburg, Sweden

2 Department of Radiation Physics, Institute of Clinical Sciences, University of Gothenburg, Sweden

3 Department of Clinical Neurophysiology, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Sweden

* Corresponding author: Emilia A. Gryska

Postal address: MedTech West, Röda stråket 10B, Sahlgrenska University Hospital, 413 45 Göteborg

E-mail address: emilia.gryska@medtechwest.se

Telephone number: +46 720304106

Keywords: Scoping review, Protocol, Brain lesions, Segmentation, Magnetic resonance imaging

Word count: 3,735

ABSTRACT

Introduction Automatic brain lesion segmentation from medical images has great potential to support clinical decision-making. Although numerous methods have been proposed, significant challenges must be addressed before they will become established in clinical and research practice. We aim to elucidate the state of the art, to provide a synopsis of competing approaches and identify contrasts between them.

Methods and Analysis We present the background and study design of a scoping review for automatic brain lesion segmentation methods for conventional magnetic resonance imaging (MRI) according to the framework proposed by Arksey and O'Malley. We aim to identify common image processing steps as well as mathematical and computational theories implemented in these methods. We will aggregate the evidence on the efficacy and identify limitations of the approaches. Methods to be investigated work with standard MRI sequences from human patients examined for brain lesions, and are validated with quantitative measures against a trusted reference. PubMed, IEEE Xplore, and Scopus will be searched using search phrases that will ensure an inclusive and unbiased overview. For matching records, titles and abstracts will be screened to ensure eligibility. Studies will be excluded if a full paper is not available or is not written in English, if non-standard MR sequences are used, if there is no quantitative validation, or if the method is not automatic. In the data charting phase, we will extract information about authors, publication details, and study cohort. We expect to find information about preprocessing, segmentation, and validation procedures. We will develop an analytical framework to collate, summarize, and synthesise the data.

Ethics and Dissemination Ethical approval for this study is not required since the information will be extracted from published studies. We will submit the review report to a peer-reviewed scientific journal and explore other venues for presenting the work.

ARTICLE SUMMARY

Strengths and limitations of this study:

-
-
- Our study design ensures an inclusive and unbiased review while maintaining good quality of the gathered sources by proposing the requirement of a quantitative validation of the presented methods.
- We will validate our search strategy by comparing the bibliographies and citations of the most recent and most cited records with the gathered sources.
- Criteria for including studies in the scoping review may turn out to be too generous, with the number of matching papers exceeding our capacity for reviewing them.

INTRODUCTION

In clinical practice, diagnosis of brain lesions is based on the patient's history, clinical presentation, visual assessment of appointed scans, and other laboratory examinations. Magnetic resonance (MR) imaging has become an important tool in brain lesion identification and classification due to its ability to produce images with high contrast resolution and sensitivity for abnormalities. Various conditions can give rise to such lesions. The most common causes include trauma, inflammation and autoimmune diseases, stroke, malignant or benign tumours, and infections [1]. Although brain lesions tend to appear significantly different from healthy tissues on MR scans, differentiating between brain lesion causes based on visual examination can be difficult or impossible. Still, visual interpretation is the most common and trusted mode of image analysis in clinical practice. Accurate identification and delineation of lesion boundaries and classification of lesional tissue components is particularly important in treatment planning for surgery or radiation therapy in tumor patients. It is also essential for determination of disease burden, prognosis, and therapy response in nearly all types of brain lesions. The process is currently commonly performed manually by an expert rater. The procedure is tedious, time-consuming, and subject to inter-rater as well as test-retest variability [2-5].

Automatic image segmentation methods promise to reduce or eliminate subjective decisions in this process, facilitating fast and accurate delineation of lesions and classification of their components on MR brain images. Although many automatic brain lesion segmentation methods have been proposed, substantial challenges remain, for example the variable appearance of the lesions on MR images due to unknown, possibly biological factors; differences of image acquisition protocols between centers; and the difficulty of validating such algorithms on

sufficiently large case numbers [6-8]. Taken together, these challenges explain why no single tool or approach has thus far been adopted in clinical or even in research practice. On surveying the literature on automatic brain lesion segmentation methods in an ad-hoc, preliminary fashion, we recognized the need for a rigorous and comprehensive review. A formalized approach to reviewing literature in this manner is the scoping review as proposed by Arksey and O'Malley [9]. Using this framework along with refinements by Levac et al. [10] and Colquhoun et al. [11], as well as elements of the PRISMA and PRISMA-ScR guidelines[12,13], we will map key concepts, converging developments, challenges, and promising new research avenues. The purpose of publishing the research plan at this stage is to document our objectives openly, to invite comments and suggestions, and to enhance the rigor of our study. This open documentation will compel us to follow the plan and justify any deviation. We believe that being fastidious in this manner will enhance the value of the research once completed.

Study aims and objectives

We have identified the following aims and objectives of the scoping review on existing automatic brain lesion segmentation techniques on conventional magnetic resonance images:

1. Elucidate the state of the art and provide a synopsis of competing approaches to automatic brain lesion characterization
2. Identify common procedures necessary for automatic brain lesion segmentation on conventional MR images
3. Identify MR data sets with reference segmentations and/or diagnostic classifications that are publicly available for method validation;
4. Identify common challenges, weaknesses and controversies, as well as unaddressed issues which can signify opportunities for future work to improve segmentation methods.

METHODS AND DESIGN

General

This section describes how each of the scoping study stages identified by Arksey and O'Malley [9] will be applied to the present study. In this study we will balance the breadth of the included studies with the depth of the analysis of reviewed methods. The nature of the researched topic imposes certain logical limitations which, together with the inclusive nature of a scoping review, will help to create a focused yet comprehensive overview of the topic.

Stage 1: Identifying research questions

We have defined the following research questions that will be addressed in this study.

1. Which common image processing steps are necessary for automatic brain lesion segmentation on MR images?
2. Which mathematical and computational theories are most commonly applied in which types of brain lesions?
3. What is the efficacy of existing implementations?
4. What are the limitations of those methods and issues that should be addressed in future studies to develop a tool that is suitable for clinical use?
5. What are the most commonly utilized MR data sets that provide reference lesion segmentation and/or diagnostic classification?

Stage 2: Identifying relevant studies

- Eligibility criteria

We established the following initial criteria for the proposed segmentation methods to be eligible for inclusion in the scoping review. A method must be applied to one or a number of commonly acquired MRI sequences from human subjects investigated for a condition known to cause brain lesion(s).

The proposed methods should be validated by comparison with a gold-standard reference segmentation of the lesions. The efficacy of the method should be reported providing quantitative scores such as sensitivity, specificity, overlap, surface distance, or volume error. Alternative validation approaches will be considered if they have face validity. A thorough validation is a necessary step in developing medical image segmentation methods. Even though we dispense with formal quality assessment of the included studies, we believe that our principled approach will enable us to provide a valuable report on the researched topic.

We do not define any particular study designs types as an inclusion criterion for this scoping review. This ensures inclusion of diverse approaches and designs and will potentially reveal which ones are favoured by a plurality of authors. This aligns with the generic aim of undertaking a scoping review, which is to investigate the range and type of evidence in a given field.

- Initial search

The proposed search strategy was thoroughly discussed and approved by the scoping study authors. We also took advantage of services provided by our university library and consulted with a librarian on the search strategy. Eligible studies will be retrieved from peer-reviewed journal articles and conference papers. We will not impose any limitations with respect to year of publication at this stage of the study.

The search will be conducted using search terms constructed in English.

Three online databases will be searched: PubMed, IEEE Xplore, and Scopus. The following search phrases will be constructed using non-controlled vocabulary to initialize the search. An advanced search in publication metadata will be conducted in all databases using the following search terms to identify potentially relevant sources (asterisk indicates a wildcard character to account for variations in the spelling of the search terms):

1. automat(ic)*
2. AND brain
3. AND lesion OR tumor OR neoplasm
4. AND segment(ation)* OR identif(ication)* OR delin(eation)* OR classif(ication)*
5. AND mri OR mr.

The search results will be exported from each database. From the controlled vocabulary tags (MeSH in PubMed; IEEE terms and INSPEC terms in IEEE Explore; index keywords in Scopus), we will build frequency tables. The most common relevant terms will be used to refine the original search phrase. Combining free text and index terms ensures high sensitivity of the search. The results will be refined by applying possible limitations defined in the eligibility criteria depending on the availability in the search engine, such as document type (Journal and conference articles) or species (Human).

A separate search will be conducted in PubMed to identify potentially eligible papers that have not been indexed with MeSH yet. MeSH terms are assigned by specialists at the National Library of Medicine after a variable delay, meaning that some recent papers lack them. We will modify the search phrase and look for the search terms as well as MeSH terms in all fields, and an additional status criterion will be added to exclude MeSH indexed papers.

We will screen bibliographies of the most recent papers as well as citations of the most cited papers and compare it with the existing sample to evaluate inclusiveness and validate the proposed search strategy. We will identify the most cited studies by dividing the number of citations of a given paper by the number of months since the publication. If there is a substantial mismatch between the existing set of selected studies, and the bibliographies and citations, we will identify additional sources by screening the bibliographies and citations of the identified set.

- Screening

The records found in the search phase will be exported to a reference management software (Zotero) to scan for and remove duplicated items. The screening of the identified records after duplicate removal will be conducted in two phases using web based application for systematic review – Rayyan QCRI [14].

First, we will rapidly screen the titles to exclude papers that evidently do not match the selection criteria. We will exclude any papers that deal with imaging modalities other than MRI, or where the title suggests that the study does not propose an automatic brain lesion segmentation method.

In the second phase, abstracts of the papers that passed the previous phase will be screened to identify and exclude irrelevant or otherwise ineligible items. The screening will be performed by finding the key terms or their synonyms in the publication title or abstract and determining if the publication is relevant. Articles that do not name in the title or the abstract or refer to any proposed method of any form of identification of any type of brain lesion will be excluded under the assumption that those papers either do not contain enough evidence for the method to be eligible for the synthesis, or do not propose a lesion segmentation method. The following key terms (and their synonyms) will be considered:

1. method
2. identification
3. brain lesion
4. magnetic resonance imaging

At the screening stage, records will be excluded if a given study has previously been reported or any of the study characteristics stated in the abstract clearly do not match the eligibility criteria for this study.

Stage 3: Study selection

For items selected during the screening stage, full-text articles will be retrieved. The following criteria will be considered as a reason for exclusion of a paper from the review:

- full paper not available
- paper written in a language other than English
- modality other than standard MRI used
- no quantitative validation found
- semi-automatic method proposed.

The terms automatic and semi-automatic segmentation do not have a widely accepted definition. For our purposes, a semi-automatic method shall be one that relies on expert's decisions during the segmentation process, while an automatic method is one that requires a user to provide (possibly preprocessed) images and launch the program, after which all decisions regarding lesion segmentation and characterization are made without human interaction.

Stage 4: Data charting

In the data charting phase, the following study information will be extracted from every eligible record: author(s), year of publication, country of origin, and funding information (if available). We will note the clinical diagnosis of the patients and MRI sequence(s) used in a given study as well as the type of brain lesion(s) studied. Based on our knowledge in the field of medical image segmentation we expect to find information on the following main categories in the method description [8]:

1. Image preprocessing methods and procedures (e.g. registration, skull stripping, intensity inhomogeneity correction, noise reduction, intensity normalization)
2. Segmentation methods and particular computational and mathematical theories applied
3. MR database used
4. Validation (e.g. amount and types of reference data; accuracy and reliability measures).
5. Efficacy of the methods in terms of applicability to a predefined task
6. Limitations of the method stated by authors

Following the suggestions included in the PRISMA checklist for scoping reviews [13], we will critically appraise the information in the gathered studies. Together with the information collected in the data charting phase, it will help us identify the common patterns, efficacy, and limitations of the studies and presented methods. The elements we will appraise will include method and material description, preprocessing description, robustness of

the method, and validation procedure including reference segmentation information and segmentation evaluation measures.

To test the proposed approach we will conduct a pilot charting and appraisal on a subset of recent studies to evaluate and refine the charted variables. Implementation details will be decided during a pilot phase of data charting and critical appraisal.

Given the extent and variability in reporting the information to be extracted, we may have to update the charting form in an iterative manner even after conducting a pilot charting [10]. Levac et al. [10] note that it is nearly impossible to design an adequate charting form at the outset, and recommend iterative refinement.

Stage 5: Collating, summarizing, and reporting the data

Once the eligible studies have been collected, we will develop an analytical framework to collate, summarize, and synthesise the data. We will make use of summary counts and tables to provide quantitative information on the body of research on automatic brain tumor segmentation methods. While analysing the data, common procedures for the methods, types of lesions and their outcomes will be identified. We will also present an inclusive comparison of methods and their performance that utilize the same MR datasets and segmentaiton evaluation measures, if their study design allows for making such a comparison. We will identify discrepancies between the analysed methods and use that information to address Objective 4 (cf. Study aims and obejctives). The consultation stage of the scoping review, described in the following section, will contribute to fulfilling that objective. We allow for the possibility to adjust or expand the initial analytical framework after the consultation stage to present the gathered information according to the stakeholders' requests.

Stage 6: Consultation

Although the consultation stage is currently considered optional in scoping reviews, we see advantages in including this stage in our study. We will use this opportunity to share preliminary findings and refer to potential stakeholders to gain more insight into our data from different perspectives. The consultation will be conducted using a questionnaire or through interviews, however a detailed design of the consultation process will be created after finishing collating, summarizing, and internal reporting the data.

Study limitations

While we aim to conduct a well structured and reproducible study, we are aware that our approach has limitations. In the defined inclusion criteria, we limit the eligible sources to journal articles and conference papers. We presume that any other sources, such as posters, books, theses, etc., will contain ineligible evidence or duplicates of included papers. We are aware of that our presumption may not be true in all cases and thus we risk excluding some eligible studies.

In the study identification phase, we attempt to construct a very sensitive search phrase with high precision. Our strategy may, however, turn out to be insufficiently inclusive, and a substantial number of additional, potentially eligible studies will be identified by screening the references and citations of selected papers. Since the process is time consuming, we will have to limit the number of papers we will consider for reference and citation screening. This may introduce subjectivity and we may not be able to identify every relevant study.

Since the nature of a scoping study is to identify evidence in a given field without particular assumptions about the designs of the sampled studies, creating an optimal framework for data charting, appraisal, and synthesis is a complex task. Evaluating efficacy of the gathered studies poses a particular challenge, and we may not be able to provide an objective comparative assessment of segmentation methods' efficacy.

We attempt to ensure reproducibility of our study through rigorous planning and thorough documentation of the study methodology. In addition, a single investigator (EAG) will be responsible for identifying and selecting eligible papers, as well as extracting and summarizing the relevant information for the study. We thus avoid interrater variability in the protocol implementation, accepting as a trade-off that we may have to constrain the sample to account for the amount of time available to the investigator.

Patient and Public Involvement

Neither patients nor public were involved in the development of this study design.

ETHICS AND DISSEMINATION

The study will contain information gathered from already published papers therefore it does not require ethical approval. We will distill the project diary (cf. General) into a review report for submission to a peer-reviewed scientific journal. In addition, we will seek to present the study at scientific conferences. Following on from the work done at the consultation stage, we will identify stakeholders outside of academia and seek to disseminate the results to them in appropriate formats (trade journal articles, lectures, laypersons' summaries, press releases).

Author Contributions: RAH conceived the idea of conducting the scoping review. EAG and RAH developed all the elements of the study design. EAG lead writing of the protocol. RAH was the main supervisor of the process of writing and JFS provided feedback on the methodology of the study as well as the manuscript. All authors approved the final version of the manuscript.

Funding statement: JFS is funded by the Knut and Alice Wallenberg Foundation grant KAW2014.0102, The Swedish Research Council grant 2017-00680, and the Swedish Childhood Cancer Foundation grant MT2014-0007.

Competing interests: None declared.

Data sharing : Data sharing not applicable to this manuscript as no datasets were generated or analysed.

REFERENCES

1. Brain lesions Causes [Internet]. Mayo Clinic. Available from: http://www.mayoclinic.org/symptoms/brain-lesions/basics/definition/sym-20050692 (accessed 2018 Jun)
2. Porz N, Bauer S, Pica A, Schucht P, Beck J, Verma RK, Slotboom J, Reyes M, Wiest R. Multi-modal glioblastoma segmentation: man versus machine. *PloS one*. 2014 May 7;9(5):e96873.
3. Weltens C, Menten J, Feron M, Bellon E, Demaerel P, Maes F, Van den Bogaert W, van der Schueren E. Interobserver variations in gross tumor volume delineation of brain tumors on computed tomography and impact of magnetic resonance imaging. *Radiother Oncol.* 2001 Jul 1;60(1):49-59.
4. Bø HK, Solheim O, Jakola AS, Kvistad K-A, Reinertsen I, Berntsen EM. Intra-rater variability in low-grade glioma segmentation. *J Neurooncol*. 2017;131(2):393–402.
5. Ashton EA, Takahashi C, Berg MJ, Goodman A, Totterman S, Ekholm S. Accuracy and reproducibility of manual and semiautomated quantification of MS lesions by MRI. *J Magn Reson Imaging*. 2003 Mar;17(3):300–8.
6. Gordillo N, Montseny E, Sobrevilla P. State of the art survey on MRI brain tumor segmentation. *Magn Reson Imaging*. 2013 Oct;31(8):1426–38.
7. Bauer S, Wiest R, Nolte L-P, Reyes M. A survey of MRI-based medical image analysis for brain tumor studies. *Phys Med Biol*. 2013 Jul 7;58(13):R97-129.
8. García-Lorenzo D, Francis S, Narayanan S, Arnold DL, Collins DL. Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. *Med Image Anal*. 2013 Jan;17(1):1–18.
9. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol*. 2005 Feb 1;8(1):19–32.

10. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci.* 2010 Sep 20;5:69.

11. Colquhoun HL, Levac D, O'Brien KK, et al. Scoping reviews: time for clarity in definition, methods, and reporting. *J Clin Epidemiol*. 2014 Dec 1;67(12):1291–4.

12. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *PLoS Med*. 2009 Jul 21;6(7):e1000100.

13. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, Moher D, Peters MD, Horsley T, Weeks L, Hempel S. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of internal medicine*. 2018 Oct 2;169(7):467-73.

14. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Systematic reviews*. 2016 Dec;5(1):210.