NIPH
Norwegian Institute of Public Health

REPORT

2021

Implementation of machine learning in evidence syntheses in the Cluster for Reviews and Health Technology Assessments: Final report 2020-2021

Utgitt av	Folkehelseinstituttet Område for helsetjenester
Tittel	Implementering av maskinlæring i kunnskapsoppsummeringer i klynge for vurdering av tiltak: Sluttrapport 2020-2021
English title	Implementation of machine learning in evidence syntheses in the Cluster for Reviews and Health Technology Assessments: Final report 2020-2021
Ansvarlig	Camilla Stoltenberg, direktør
Forfattere	Ashley Elizabeth Muller, prosjektleder, Folkehelseinstituttet Heather Ames, Folkehelseinstituttet Jan Himmels, Folkehelseinstituttet Patricia Jacobsen Jardim, Folkehelseinstituttet Lien Nguyen, Folkehelseinstituttet Christopher Rose, Folkehelseinstituttet Stijn van de Velde, Folkehelseinstituttet
ISBN	978-82-8406-231-0
Publikasjonstype	Report (Rapport)
Antall sider	29 (79 inklusiv vedlegg)
Oppdragsgiver	Folkehelseinstituttet
Emneord(MeSH)	biomedical; technological assessment, health; unsupervised machine learning; supervised machine learning; deep learning
Sitering	Muller AE, Ames H, Himmels J, Jardim PJ, Nguyen L, Rose C, Van de Velde S. Implementering av maskinlæring i kunnskapsoppsummeringer i klynge for vurdering av tiltak: Sluttrapport 2020-2021 [Implementation of machine learning in evidence syntheses in the Cluster for Reviews and Health Technology Assessments: Final report 2020-2021] –2021. Oslo: Folkehelseinstituttet, 2021.

Table of contents

Key Messages

Machine learning (ML) has the potential to increase the efficiency of evidence syntheses. During 2020-2021, a team in the division for Health Services at the Norwegian Institute of Public Health, tested and documented pros and cons of using ML in various phases of the conduct of various evidence syntheses, and built employees' competence in using ML. This report describes the work undertaken by the ML team, project results and lessons learned.

The ML team focused attention on ML functions and systems available within EPPI Reviewer: Priority screening, Custom and Pre-built classifiers, RobotReviewer to assess Risk of Bias, Automatic text clustering, and Microsoft Academic Graph (MAG). We implemented ML functions across 19 project teams and trained 23 employees. We found that utilizing ML in our reviews increased speed, with no identified threats to methodological quality. Screening time was reduced by 60-90% in all projects. Automated study categorization – while applicable to a smaller range of projects – reduced manual time in this phase by 60-70%.

ML can, and should, change usual project workflows. The review process can become less linear and more cyclical, and several tasks can be conducted in parallel. However, workflow changes are not insignificant for those involved, and future ML work would benefit from a structured approach to both change management and innovation diffusion.

The report concludes with lessons learned and experiences gained. They shaped our proposals for future ML strategies, covering capacity-building, innovative activities, evaluation of effect, and workflow optimization.

Title:
Implementation of machine learning in evidence syntheses in the Cluster for Reviews and Health Technology Assessments: Final report 2020-2021

Publisher:
The Norwegian Institute of Public Health conducted the project based on an initiative by the Cluster of Reviews and Health Technology Assessments, Division for Health Services at the NIPH

Type of publication:
Report

Activity timeline:
Dec 2020 - June 2021

Machine learning functions evaluated and implemented:

Priority screening
Classifiers (3 types)
RobotReviewer to assess Risk of Bias
Automatic text clustering
Microsoft Academic Graph (MAG)

Hovedbudskap

Maskinlæring kan bidra til betydelig effektivisering av kunnskapsoppsummeringsprosesser. Et lag i Området for helsetjenester ved Folkehelseinstituttet evaluerte og dokumenterte i 2020-2021 fordeler og ulemper ved maskinlæring i flere faser av kunnskapsoppsummeringer, og bygde medarbeidernes kompetanse i å bruke ulike funksjoner. Denne rapporten beskriver lagets arbeid, resultater og erfaringer.

Maskinlæringslaget fokuserte på funksjoner som er tilgjengelig i EPPI-Reviewer verktøyet: «priority screening», flere typer classifiers, RobotReviewer for å vurdere risiko av skjevheter, «automatic text clustering», og Microsoft Academic Graph. Vi implementerte funksjonene i 19 prosjekter og opplærte 23 medarbeidere. Et hovedfunn er at maskinlæringsfunksjoner reduserte manuell tidsbruk, uten reduksjon i metodisk kvalitet. Tidsbruk på vurdering av studier gikk ned med 60-90 % i alle prosjekter. Automatisk studiekategorisering reduserte tidsbruk i denne fasen med 60-70 %.

Maskinlæring kan og bør endre dagens arbeidsflyt. Kunnskapsoppsummeringsprosessen kan bli mindre lineær og mer syklisk, og flere oppgaver kan gjøres samtidig. Slike endringer kan være vesentlige for alle involverte, og i framtidig maskinlæringsarbeid vil det være nyttig med en strukturert tilnærming til både endringsledelse og innovasjonsspredning.

Rapporten avslutter med erfaringer og lærdommer. Disse formet vårt forslag til framtidige strategier relatert til kompetansebygging, innovasjonsaktiviteter, evalueringer og arbeidsflytoptimalisering.

Tittel:
Implementering av maskinlæring i kunnskapsoppsummeringer i klynge for vurdering av tiltak: Sluttrapport 2020-2021

Hvem står bak denne publikasjonen?
Folkehelseinstituttet utførte studien basert på et initiativ fra klynge for vurdering av tiltak, område for helsetjenster i FHI

Type publikasjon:
Rapport

Tidsperiode for prosjektet:
Des 2020 - Juni 2021

Maskinlæringsfunksjoner som vi evaluerte og implementerte:

Priority screening
Classifiers (3 types)
RobotReviewer to assess Risk of Bias
Automatic text clustering
Microsoft Academic Graph (MAG)

Preface

The Cluster for Reviews and Health Technology Assessments, Division for Health Services at the Norwegian Institute of Public Health (NIPH) decided in the fall of 2020 to conduct a project on machine learning related to the conduct of evidence syntheses. The goals were to test and document pros and cons of using machine learning in various phases of the conduct of evidence syntheses, as well as build employees' competence in using machine learning. A team of seven worked toward these goals from December 2020 until June 2021. This report describes their work.

The report is relevant for researchers and managers interested in implementing machine learning in their evidence syntheses. It is particularly relevant for evidence synthesis environments that do not have machine learning specialists.

Financing

The work was self-initiated and financed by the Cluster for Reviews and Health Technology Assessments, Division for Health Services at the NIPH.

Team members

Project leader: Ashley Elizabeth Muller
Team members: Heather Ames, Jan Himmels, Patricia Jacobsen Jardim, Lien Nguyen, Christopher Rose, Stijn Van de Velde

Conflicts of interest

All authors declare they have no conflicts of interest.

Kåre Birger Hagen	Rigmor C Berg	Ashley E. Muller
Research director	*Department director*	*Project leader*

Background

In early 2020, the Cluster for Reviews and Health Technology Assessments, Division for Health Services at the Norwegian Institute of Public Health (NIPH), became increasingly aware of the potential benefits of using machine learning (ML) in the conduct of evidence syntheses. Thus, the leader team in the cluster decided to initiate a project on ML. The project had two overarching goals: To test and document pros and cons of using ML in various phases of the conduct of evidence syntheses, and to build employees' competence in using ML. There were four objectives:

- Develop and implement a capacity-building ML strategy for the Cluster of Reviews and Health Technology Assessments
- Conduct a retrospective evaluation of ML performance in completed projects, and potentially evaluations in new projects, including recruiting and teaching project leaders
- Report results of capacity-building and evaluations to leadership and others in the Division for Health Services
- Stay abreast of methods and ongoing studies of ML in other health technology assessment organizations, and assess possibilities for collaboration

A team of seven employees (all but one) from the Cluster for Reviews and Health Technology Assessments, dedicated much of their time from December 2020 until June 2021 to the project.

The ML team's work was anchored in the preliminary NIPH strategies for the 2019-2024 period concerning automation, increasing speed of evidence syntheses, and workflow and methods innovation. One of the goals of the division-specific strategies was for the Division for Health Services to become a leader in automation and digitalization of work processes, and to use these practices to summarize evidence more efficiently.

On a related note, we mention that during this report's preparation, the preliminary NIPH strategy was being revised. The machine learning team analyzed the preliminary strategic priorities and identified a need to integrate the ongoing, siloed ML activities at NIPH into a more cohesive, cross-division approach. Accordingly, the team began contacting, mapping and discussing with other actors and research teams in NIPH involved with ML. The strategy changes we proposed are included in the new NIPH strategy: "NIPH shall be a leader in big data, machine learning, and automation within public health", under strategic priority 7. We refer readers to a separate document which details our machine learning strategy.

Project results

The following text details ML team activities undertaken January 2020 - May 2021.

Time and resources

The team of seven, including two advisors, was allocated a maximum of twelve months' working time. The resources allocated to the team were adequate, although not fully exhausted by all team members. Some team members found it difficult to prioritize this team over projects with strict deliverables and timelines. The medium size of the team allowed us to work cooperatively and divide tasks among ourselves.

Internal team capacity building and team-building

To bring team members unfamiliar with the field of ML up to date, and as a team-building exercise, we spent the first four weeks presenting new research and concepts to each other in weekly three-hour meetings, followed by discussions. Presentations are available for future use as a ML syllabus. We also used the first part of the year familiarizing ourselves with EPPI Reviewer and its functions.

Implementation and training

The ML team supported the implementation of machine learning functions in 19 projects (including the original pilot project in August 2020). Twenty-three employees were trained, of which 18 were not members of the ML team. A list of projects and employees can be provided.

Table 1 gives an overview of the team's implementation and training activities.

Table 1: *Overview of implementation and training activities*

Machine learning function	Project teams	Employees[a] trained[b]	Training materials created
Priority screening	13	13	How-to guides in Norwegian and English, educational material
Custom classifiers for screening	10	6	How-to guide, educational material
Pre-built study design classifiers	1	2	Educational material
Custom classifiers for study categorization	1	3	Educational material
RobotReviewer to assess Risk of Bias	3	8	How-to guide for project leaders, how-to guide for project members, educational material
Automatic text clustering	2	4	Educational material
Microsoft Academic Graph (MAG)	4	6	-

[a] Including ML team members. [b] Not all trained users can implement a function independently.

To support project leaders with the implementation of new ML functions, we provided one-on-one training and technical assistance. Each project received a dedicated ML team member who trained the project leader first, and then the rest of the team, and was available for immediate assistance when needed. This intensive technical assistance ensured we were able to gather the data required for evaluation and validation activities, e.g. training time required. We used a training hand-off procedure to build capacity within the team: 1) a ML team trainee sat in on an experienced ML team member's training of a project; 2) both co-led the next training; 3) finally, the ML team trainee led a subsequent training, with the experienced member sitting in for assistance.

Intensive, often one-on-one technical assistance was necessary for project leaders to understand and implement particular functions, however, providing this level of intense assistance was not sustainable or scalable. In most cases, technical assistance was not sufficient for project leaders to become confident enough to train others, although it did build their confidence in choosing to use a particular technique in future projects.

Acknowledging that one-on-one technical assistance to all project leaders was not sustainable, we developed stand-alone training materials for project leaders and/or members. These materials encourage users to begin implementation independently of the ML team. At the time of report writing (June 2021), these materials are in the final phase of piloting and feedback collection. So far, the training materials have been successful in supporting project leaders to more independently implement ML functions, and reduce technical assistance needs from the machine learning team.

There remains uncertainty in responsibility for tasks among overlapping actors providing digital support: the digital tools team (and EPPI superuser within that team), the

ML team, and EPPI software support. In response and in agreement with the digital tools team and leadership, responsibility was delegated for basic EPPI functions to the digital tools team and ML functions to the ML team. We also encouraged project leads to contact EPPI support for questions, but the threshold appeared higher for this than asking questions in-house. The new EPPI superuser's involvement in an early ML project has proven valuable as software skills were expanded with technical understanding of basic ML techniques – this overlap may be a prerequisite for optimal coordination between the two teams.

Testing and validation

While all ML functions available in EPPI-Reviewer are fully developed and have extensive documentation of validity, the majority lacked published validation studies specifically conducted within the field of evidence synthesis. We decided that internal/institutional evaluations of all functions were a necessary first step to increase trust and buy-in among colleagues. Additionally, these evaluations provided a stronger foundation to evaluate particular functions' usefulness to our workflows. Almost all evaluations were integrated into ongoing projects, with exception of the retrospective evaluation of ML within screening (NICE is leading a simulation study of retrospective studies to identify "stopping criteria" for screening, while this team built and evaluated custom classifiers using previously completed projects) and a parallel initiative of our librarians to test Microsoft Academic Graph.

We created user-friendly introductions to each ML function; please see User-friendly summaries of machine learning functions. These 1-page, introductory infographics were developed to help project leaders understand the different functions, when to use them, and how to combine them.

In the following subsections we present how we tested and validated each of the functions as well as recommendations for next steps and/or implementation. Table 2 provides is a summary. Characteristics of each function is found in the description of each function further below.

Table 2: *Overview of evaluated techniques, benefits, and recommendations*

Function	**Relevant review types**	**Workflow changes to optimize benefits**	**Benefits**	**Next steps**
Priority screening	All	Single- or auto-screening. Screening de-prioritization.	60% less time used to screen. Rapid team understanding of inclusion criteria. Rapid communication of potential review size (or other issues) to commissioner.	Scale up implementation
Custom classifiers for screening	Reviews with clear inclusion criteria and research questions	Single- or auto-screening. Screening de-prioritization.	60-90% less time used to screen, when preceded by priority screening	Scale up implementation
Pre-built study design classifiers	Reviews of RCTs. Overviews of SRs.	Single- or auto-screening. Screening de-prioritization.	Accurately identify prioritized designs to reduce screening burden	Scale up implementation
Custom classifiers for study categorization	Review updates. Rolling reviews. Literature searches with sorting. Large reviews that have already begun categorization.	Single- or auto-categorization (data extraction)	32-77% less time used to categorize. Equally as accurate as any one reviewer, blinded or non-blinded.	Evaluate further. Explore additional applications
RobotReviewer to assess Risk of Bias	Reviews of RCTs	Use as pedagogic tool, particularly for newer researchers	Equally as accurate as one researcher. No reliable time estimates.	Scale up implementation
Automatic text clustering	All	Single- or auto-screening. Screening de-prioritization. Single- or auto-categorization (data extraction).	In screening: 74% less time to screen when applied to the least relevant studies. In study categorization: Equally as accurate as one researcher. 34% less time to categorize when semi-automated; 71% less time when fully automated.	Explore additional applications. Scale up implementation within screening
Microsoft Academic Graph (MAG)	Review updates	Supplement or replace some database searches	Retrieve fewer and more relevant studies than traditional database searches. Potentially replace one or more database searches.	Librarians proceed

Explanation: RCT=randomized controlled trial, SR=systematic review.

Priority screening

Priority screening learns from researcher screening decisions and pushes relevant studies forward in the screening queue (table 3). This technique does not make screening decisions, but helps researchers identify and handle included studies first.

Table 3: *Brief description of characteristics of priority screening*

Type of machine learning	Supervised, human-in-the-loop, active learning
Combination with other ML functions	Optimizes the subsequent use of custom classifiers
Review stage	Title and abstract screening
Degree of difficulty	Easy
Support needs	Low - Can be implemented independently with email support from EPPI or ML team

Five projects contributed to this evaluation:

- Secure institutions for youth
- Understanding and helping children who resist or refuse postseparation parental contact
- Systematic review of RCTs of treatment for perpetrators of sexual violence
- The relationship of travel distance to delivery institutions and accompaniment
- The effects of covid-19 on children and youth's wellbeing

How did we test the function?

- In the pilot project, we randomized 14,000 studies to be screened as usual (randomly) or using priority. Researchers tracked time spent, and we calculated inclusion rates after regular amounts of studies had been screened.
- Subsequent projects used priority screening exclusively (with no comparison to random screening) and we tracked inclusion rates at regular intervals.

What have we found so far?

- Time savings in the screening phase: 60% less time compared to screening as usual, if used until the inclusion rate flattens and then moving to single-screening (pilot study). 90% less time when used in combination with custom classifiers and switching to single- or auto-screening for studies under or over various cut-offs (see Classifiers).
- Efficiency: 95% of all included studies are found after screening 7.5-35% of retrieved studies. The more precise the PICO (and the more precise human screening), the more efficient priority screening is, and the quicker all included studies are identified.
- Other benefits: It requires precision of inclusion criteria immediately in the screening process, and therefore a clarification of misunderstandings earlier, both within the project team and between the project team and commissioner. It also allows projects to provide commissioners with estimates of project size quickly.
- Usefulness: Highly accepted by the teams that have used it.

Workflow changes that optimize benefits

- Priority screening necessarily changes existing screening workflows, and more than any other function we have evaluated. For example, the project team should sit together electronically or in person when screening the first 200 studies, and reconcile screening conflicts much more frequently and at regular intervals.
- Move to single-screening, and/or de-prioritize screening, after the inclusion rate plateaus. To maximize time savings, build a custom classifier.
- Begin full-text screening in parallel, as relevant studies are identified immediately.

Next steps

- We are confident that priority screening can be implemented across all projects.

Classifiers

Classifiers use natural language processing to predict membership of a piece of data (e.g. text in the title and abstract of a study) into one of two binary categories: "*A*" vs "*not A*" (table 4). For example, *include* vs *exclude*, or *population of interest* vs *not the population of interest*. "Pre-built" classifiers are those that have been trained and validated. "Custom" classifiers refer to any classifiers built by a user. Within EPPI-Reviewer, several pre-built classifiers are available, and users can build their own. We conducted three separate evaluations.

Table 4: *Brief description of characteristics of classifiers to screen or categorize*

Type of machine learning	Supervised, human-in-the-loop
Combination with other ML functions	Ideal after priority screening
Review stage	Title and abstract screening, or data extraction
Degree of difficulty	High. Requires both understanding of the ML process behind it, and high user skills in EPPI.
Support needs	Our user guide can be followed. 60-120 min of ML team support to help project leaders the first time.

Custom classifiers for screening

This type of classifier is useful for all systematic reviews and health technology assessments (HTAs) with clearly defined research questions and inclusion criteria. It is not recommended for overviews of overviews, broad scoping reviews with multiple research questions, or for reviews with novel definitions of interventions, exposures, etc. The accuracy depends on model quality, which the ML team can help project leaders assess in order to proceed correctly.

Nine projects contributed to this evaluation: an update of a covid-19 rapid review, one EUnetHTA rolling collaborative review and two updates, three scoping reviews, three reviews of RCTs/cohort studies, and one overview of reviews.

How did we test the function?

- Review of RCTs: We built a custom classifier after having screened (using priority screening and pre-built classifiers) 13.5% of references. We auto-screened all studies <10% likely, then manually single-screened to quality control. Screeners tracked time.
- Review of cohort studies: We built a custom classifier after having screened 61% of references. We deprioritized and single-screened all studies <30% likely, while writing the report.
- EUnetHTA rolling review and covid-19 update: We built a classifier first after having screened the first 1000 studies, and at regular increments thereafter, and repeated during subsequent updates.
- The remaining studies contributed to a retrospective evaluation. In seven completed reviews, we trained classifiers using random samples of 50 and 100 studies, as well as the first 25 studies included and a random 25 excluded studies (balanced between included and excluded), applied these to the remaining studies, and compared classifications with actual screening decisions

What have we found so far?

A <30% cut-off criteria is highly accurate to predict exclusion:

- Studies below this cut-off can be auto-screened as irrelevant.
- No studies included at full-text are lost.
- 18-90% fewer studies can be screened at title and abstract level.
- Studies included first by priority screening should be used to train the classifier. These classifiers performed better than models with larger but randomly chosen training sets.
- This applies to SRs with clear research questions and well-defined interventions or exposures.

There are significant time savings even using a more conservative cut-off:

- In practice: Auto-screening <10% relevant studies saved 48 hours (36% of total screening time), with complete accuracy.
- Retrospective estimates:
  - Auto-screening <10% and >90% relevant studies, saves 90% of screening time.
  - Single-screening <50% relevant studies saves 60-70% of screening time.
- This applies to systematic reviews with clear research questions and well-defined interventions or exposures.

When custom classifiers do not work:

- In broad scoping reviews with multiple RQs or novel definitions of exposure, the data was not good enough to create a strong model. 1-2% of included studies were missed using a <30% cut-off.

What do we need to do next to find out more?

- Evaluate in a qualitative evidence synthesis.

- Improve training materials to make new users more independent and to reduce training burden on the ML team.
- Scale up teaching of necessary basic ML knowledge, to reduce user threshold to use this technique.
- Consider making guidelines regarding a cut-off threshold that could be implemented in evaluated product types.

Pre-built study design classifiers

This type of classifier is applied to identified studies to identify three specific study designs: RCTs, systematic reviews, and economic evaluations. We did not evaluate the economic evaluation classifier. These classifiers are already fully developed and validated.

The following projects contributed to this evaluation:

- Pilot and retrospective evaluation: Systematic review of RCTs of treatment for perpetrators of sexual violence (12,000 references, 1.5% included at title and abstract, 0.1% included at full-text). Prioritized study designs: systematic reviews, then RCT, then n-RCT.
- Retrospective evaluation: Overview of reviews of remote patient monitoring RCTs (3,000 references, 4.8% included at title and abstract, 0.1% included at full-text). Due to a complicated research question, this project involved assessing primary studies included within systematic reviews.

How did we test the function?

- Pilot: We applied study design classifiers consecutively, according to prioritized study design: first the systematic review classifier, then RCT classifier. We prioritized screening of those classified as >50% likely. At the end of the project, we checked all included studies' classifier score to see if they had been captured by the relevant study design classifier.
- Retrospective evaluations: We retrospectively applied the relevant pre-built classifier(s) to screened studies in two reviews. We compared classifications to actual screening and inclusion decisions.

What have we found so far?

- Highly accurate: Pre-built classifiers are excellent at identifying study designs, confirming previous research. In the pilot study, 100% of included RCTs were identified by RCT classifier (as well as two included n-RCTs).
- <30% cut-off is accurate to auto-screen and reduces screening burden: They can be trusted to auto-screen irrelevant designs using a <30% cut-off, with no relevant studies lost. In the retrospective evaluations, auto-screening would have reduced screening burden by 25-76% studies at the title and abstract level, and 2-63% at full-text level.
- >50% cut-off is accurate to prioritize relevant designs. In the pilot study, 7 of 8 included studies were identified by the SR and RCT classifiers (the remaining study was a different study design and identified by a custom classifier). These were captured after having screened only 13.5% of 12,000 references.

Next steps:

- These are well-developed and there is no need for further internal evaluation.
- Improve training materials to make new users more independent and to reduce the training burden on the ML team.
- Scale up teaching of necessary basic ML knowledge, to reduce user threshold to use this technique.

Custom classifiers for study categorization

This type of classifier is relevant for review updates, rolling/living reviews, and other large projects (3000+ studies). It categorizes studies based on titles/abstracts, which can be used as a direct form of data extraction, or as a sorting exercise in order de/prioritize or target screening or other actions.

The following projects contributed to this evaluation:

- Covid-19 living map: Studies were manually categorized according to title/abstract to at least one population and one intervention. Thousands of new studies each week required significant scaling up of activities.
- EUnetHTA rolling collaborative HTA on rare medications for covid-19: The team could not rely solely on priority screening, as rare medications were not being picked up and thus the algorithm could not learn to identify them. Neither could the team rely on manual screening, due to the amount of studies and the rolling deadlines.

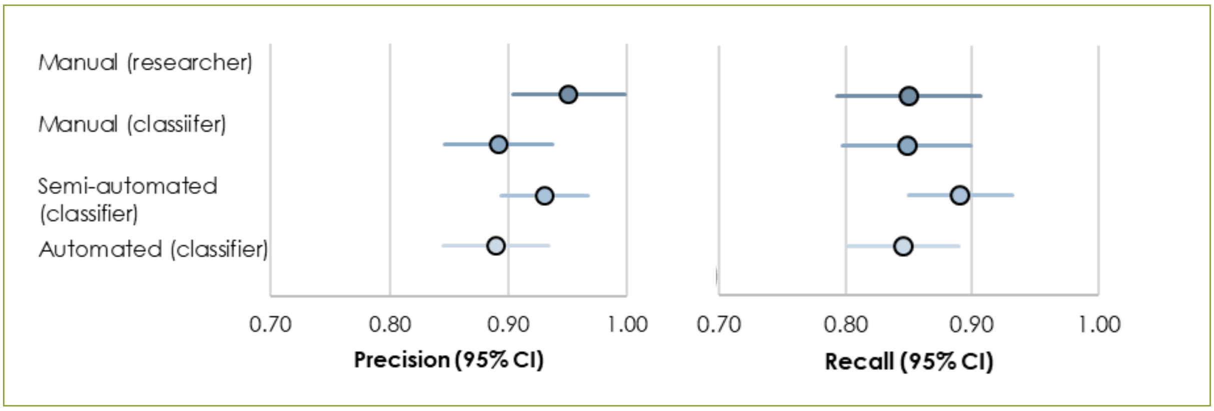
How did we test the function?

- Covid-19 living map: After categorizing 2,400 studies, we built custom classifiers to predict the 50 most common categories. 200 unscreened studies were randomized into 1 of 3 arms (2 researchers blinded to each other, fully manual; fully automated, with quality-control by 1 researcher; semi-automated, with 1 researcher non-blinded to the classifiers and 1 researcher as quality-control). Three researchers were randomly assigned studies within each arm. Precision, recall, and time were tracked.
- EUnetHTA rolling review: Classifiers were built to identify studies of prioritized rare medications that they team had not yet identified through priority screening. That is, classifiers identified studies of thematic relevance to prioritize for human screening, rather than identifying studies relevant for inclusion.

What have we found so far?

- 60-70% time savings in categorization compared to manual practice
- Successfully identified rare studies for further screening, which otherwise would not have been identified through priority screening
- Equal accuracy compared to manual practice (Figure 1)

__Figure 1:__ Accuracy of custom classifiers



What do we need to do next to find out more?

- Continue evaluation in future review updates or rolling reviews.
- Scale up implementation through teaching and training so that more project leaders can be independent.

RobotReviewer to assess Risk of Bias

RobotReviewer is fully developed ML system that assesses the first four domains of Cochrane's Risk of Bias tool and extracts relevant text to justify each assessment (table 5). It is integrated into EPPI Reviewer, as well as a standalone web-based tool.

__Table 5:__ Brief description of characteristics of RobotReviewer to assess Risk of Bias

Type of machine learning	Semi-automated, human-in the-loop: the user can accept suggestions for domain assessments and attach text snippets or amend them.
Combination with other ML functions	Not required
Review stage	Risk of Bias assessment for RCTs
Degree of difficulty	In EPPI Reviewer: intermediate skills. In the web-based version: no skills needed, but this is a slower alternative to EPPI Reviewer, and users were less positive.
Support needs	Minimal: Follow our how-to guide at your own pace. The EPPI superuser can you help you if you get stuck.

We tested RobotReviewer in two systematic reviews of RCTs involving six researchers.

- Work-related interventions for people on long-term sick leave: N=23 RCTs contributed 148 domains. Two experienced and two newer researchers. One researcher-pair used RobotReviewer within EPPI Reviewer; one pair used the RobotReviewer website.
- Systematic review of RCTs of treatment for perpetrators of sexual violence: N=3 RCTs contributed 12 domains. One experienced and one newer researcher. One researcher used EPPI Reviewer and the other used the RobotReviewer website.

How did we test the function?

- RCTs were randomly assigned into two arms for assessment: RobotReviewer within EPPI Reviewer, or the RobotReviewer website.
- All researchers were able to see RobotReviewer's domain and text suggestions while they made their own (i.e. no blinding). We measured human changes to RobotReviewer's domains (160 in total), changes from individual human assessments to final assessments, whether RobotReviewer's extracted text was deemed correct by humans, and time spent by every human on every step (administration, training, individual assessment, reconciliation, etc). Each person was also asked to report their overall impressions of the utility of RobotReviewer.

What have we found so far?

Accuracy

- RobotReviewer was as accurate as any one researcher: researchers accepted 83% of RobotReviewer's assessments (133 of 160), and 81% (129 of 160) of each other's assessments.
- In 79% of domains, there was complete agreement between RobotReviewer's assessment, a human's assessment, and the final assessment after agreement with another human. In only 4% of domains did RobotReviewer under-estimate bias. For all other domains, automated RoB was over-estimated.
- Text snippets were sufficient for 86% of domains (86 of 104). This means researchers did not have to extract text justifications for 86% of these domains.
- Human corrections to RobotReviewer did not correlate with human experience level (i.e. no sign of confirmation bias among newer researchers), or with reviewer order (i.e. no sign of confirmation bias among the first of two researchers).

Time and resource use

- Using RobotReviewer in EPPI Reviewer took 40% less time than using the web-based version. However, time use varied substantially by individual, and estimates must be taken with caution. Time use did not vary consistently according to experience level, amount of human corrections to RobotReviewer, or even amount of human corrections during reconciliation.
- We did not evaluate time use without automation.
- Administration time without needing to train a team (1 leader, 2 members, 1 support/analysis person): 2.6 hours. Administration time when training was needed, for an entirely new project team: 5 hours.

Acceptance

- Newer researchers said the extracted text helped focus their attention to the relevant parts of the study to examine, and that this saved time. Experienced researchers were, at worst, ambivalent. No one was negative to using RobotReviewer in the future, particularly the EPPI integration.

- Most researchers are not interested in replacing one reviewer with RobotReviewer, but in adding RobotReviewer to the existing process of two reviewers.

What do we need to do next to find out more?

- Recommendation: Repeat this evaluation in two new social/welfare reviews.
- Recommendation: Explore adaptation to Cochrane's Risk of Bias version 2.
- Optional: If time saved compared to fully manual RoB assessment is of interest, repeat this evaluation in a large review; ideally with the same participants.
- Optional: repeat this evaluation and measure acceptance more systematically.
- Proceed with capacity-building by highlighting accuracy over time saved.

We have an ongoing manuscript reporting these results which will be submitted in the fall.

Automatic text clustering

Clustering algorithms analyze the distribution of words, parts of words, or terms in titles and abstracts, then uses the specifications of the user to make clusters based on dis/similarity, with descriptive names (table 6). The references in a review are assigned to one or more automatically identified clusters, such that any two references within the same cluster are similar in some useful way, and any two clusters are dissimilar in some useful way. Each cluster's references, text (titles/abstracts), and search terms can be examined.

Table 6: *Brief description of characteristics of automatic text clustering*

Type of machine learning	Unsupervised
Combination with other ML functions	When used to help screen irrelevant references: useful to precede with priority screening and custom classifiers
Review stage	Title and abstract screening, data mapping, study categorization, searching
Degree of difficulty	Intermediate
Support needs	High: ML team provides an introduction and is available for troubleshooting. The user can follow EPPI's guides and contact the NIPH EPPI superuser or EPPI Centre for support.

Automatic document clustering was tested across the following projects:

- Pilot project for study categorization: Secure institutions for youth, a systematic literature search with sorting.
- Pilot project for use in screening: Systematic review of RCTs of treatment for perpetrators of sexual violence
- The relationship of travel distance to delivery institutions and accompaniment

How did we test the function?

- *Study categorization or data mapping*: We compared time use, precision and recall of manual study categorization (humans using human-designed categories), fully automated clustering (machine using machine-designed categories), and semi-automated clustering (human using machine-designed categories), in a simplified systematic review. All 128 studies in a review were categorized by two humans manually. We then ran the clustering algorithm, and randomly assigned all studies to be either coded by a human researcher blinded to cluster assignment (mimicking two independent researchers) or by a human researcher non-blinded to cluster assignment (mimicking one researcher checking another's work); the gold standard was agreement by a third researcher. Finally, we compared the original cluster assignments to this gold standard.

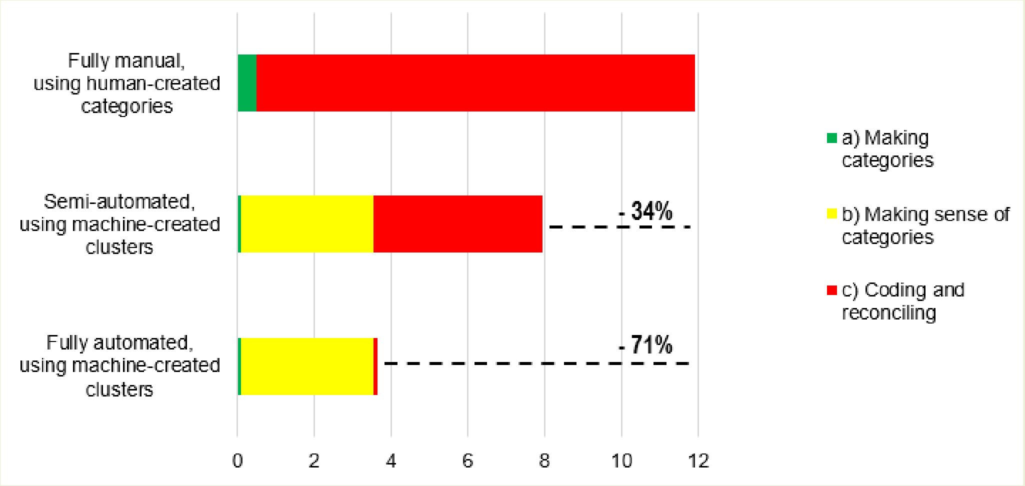
- *Screening*: We applied auto clustering to half of all unscreened studies that had already been classified as irrelevant. One researcher screened as usual, while a second used the clusters to help screen. We tracked productivity.

What have we found so far?

Data mapping:

- Most of the machine-created clusters were meaningful and useful, and some overlapped with manual categories. Machine-created clusters also uncovered one category not identified by human researchers – but it could not have been used to sort studies into the pre-determined categories.
- Equal accuracy: When humans categorized according to the auto clustering scheme, automated clustering had similar precision to both blinded and non-blinded researchers (e.g., 88% vs 89%), but higher recall (e.g., 89% vs 84%).
- No evidence of confirmation bias: Researchers blinded and non-blinded to the cluster assignments did not categorize differently.
- Time saved: Semi-automated clustering took 34% less time than fully manual categorization of 128 studies, including time spent making the categories/clusters to final agreement. Fully automated clustering took 71% less time (figure 2).

***Figure 2:** Time used for categorzation of 128 studies (hours)*



Screening:

- Time saved: 74% less time used to screen irrelevant studies (383 excluded/20 min with clusters, including the time needed to make the clusters, compared to 100 excluded /20 min).

Usefulness:

- Study categorization / data mapping: Ideal for simpler products (scoping reviews, systematic literature with sorting), to quickly become familiar with available data and uncover similarities and differences between studies.
- Screening: The more studies to screen, the more useful auto clustering is. It is particularly useful to screen or auto-screen irrelevant studies near the end of the priority screening process.
- Norwegian studies can be clustered.
- References without abstracts (often grey literature) are difficult to cluster.

What do we need to do next to find out more?

- For use in screening: test in 1-2 more projects with large amounts of studies, to confirm time saved. Randomize half of studies to be screened as usual, and half to be clustered and then screened.
- For use in search term identification: a librarian team should evaluate usefulness of automatically vs manually identified terms, in a finished search strategy.
- Clustering is a well-known ML technique. We should explore other innovative ways of applying auto clustering to systematic reviews, e.g. sampling within QES.
- Scale up implementation.

A manuscript reporting these results has been accepted upon minor revisions to *Research Synthesis Methods*.

Microsoft Academic Graph (MAG)

Microsoft Academic Graph (MAG) is an online database and knowledge graph of 260 million scientific publications, featuring a novel data structure that is based on advanced neural network machine learning (table 7). With MAG, researchers are able to search for research semantically, similar to searching in Google, and research is linked using an iterative, machine-learning-created hierarchy of 700,000 topics – rather than having to identify research based on keywords or database-specific terms.

Within the EPPI software it is possible to use a selection of articles as a starting point to conduct literature searches of the whole database, by requesting the retrieval of similar studies. Hence the tool provides the option to update a review or supplement a search, based a previous version's included studies or an already included batch of studies from a single database.

In May 2021, Microsoft announced that the Microsoft Academic website will be retired on December 31, 2021. Although this means that introducing MAG searches more widely is not sensible, gained experience supports the use of semantic/neural network searches, which are being developed by other players in the field (Google Scholar, Web of Science, and Scopus). Our gained experience will be of relevance when evaluating usefulness of other service provider's search functions in the future.

Table 7: *Brief description of characteristics of Microsoft Academic Graph*

Type of machine learning	Neural network
Combination with other ML functions	Priority screening, custom classifiers
Review stage	Searching, title and abstract screening, review updating
Degree of difficulty	Low
Support needs	N/A – Librarians proceed

We evaluated this function in the following projects:

- Long covid
- Risk factors of covid (4th update)
- EUnetHTA rolling collaborative review of rare medications (3rd update)
- An ongoing librarian evaluation led by Lien Nguyen

How did we test the function?

- Covid projects: We used MAG as a supplementary database for an update or to complement a simple search within a review. We used priority screening to immediately identify relevant studies following database searches, then entered the included studies into MAG, and retrieved relevant studies back.
- EUnetHTA and librarian evaluation: We compared overlap between MAG and traditional database searches, to identify if studies were identified by only one of the two sources.

What have we found so far?

- MAG's retrieved studies are 3-6 times more relevant compared to a single database's retrieved studies, both at title/abstract and full-text level. MAG provided 23-50% of the studies included at full-text.
- MAG retrieves up to 85% fewer studies compared to a single database search.
- In one project's update (EUnetHTA), MAG failed to identify one included study at full-text that the traditional search identified, due to a 4+ week lag after journal publication. In the librarian evaluation, MAG retrieved all included studies.

What should a librarian team do to find out more?

- Identify alternatives to MAG, due to MAG shutting down in December 2021.
- Measure overlap between our commonly used databases and MAG (or MAG alternatives), to reduce searching in superfluous databases/sources.
- Assess whether a traditional literature search can be replaced by searching exclusively in MAG.
- Repeat this evaluation in social/welfare reviews.

- Repeat this evaluation in different review sizes, to estimate a threshold for when it is enough to search in/with MAG only.
- Explore MAG's potentials in grey literature searching, which is known to be time consuming.
- Explore the potential implications of MAG (and its alternatives) to our conventional approach to searching. We need to be prepared for the next alternative, so that we can quickly implement and evaluate its functions.

Collaboration outside of the ML team

Part of the team's work was to assess possibilities for collaboration, nationally and internationally.

National Institute for Health Care Excellence and EPPI Centre

We initiated a study with NICE and EPPI Centre to improve the priority screening algorithms within EPPI. Each organization has contributed RIS files of completed projects, and NICE and EPPI programmers are running simulations with new algorithms. This study (k > 100 projects) is the largest simulation study of ML approaches with screening, and results will be used to suggest stopping criteria for screening, or when researchers can stop manual screening.

University of North Carolina

We exchange researcher-oriented ML user guides and feedback with the University of North Carolina's information specialists, who hold responsibility for ML activities within evidence synthesis.

NIPH

We initiated talks with: Divisions for Mental and Physical Health, Health data and digitalization, Infectious Diseases, and IT.

We have reached out to researchers across the NIPH to map ongoing ML activities and interests, and held a one-hour networking meeting on 23. June 2021. The meeting goal was to be a springboard for knowledge transfer and collaboration beginning simply by communicating, as it appears that ML activities are siloed within both divisions and projects. We identified overlapping activities and drivers, and are working on next steps.

Dissemination outputs

User-friendly summaries of machine learning functions

We created 1-page, user-friendly summaries of each ML function. They were developed to help project leaders understand the different functions, when to use them, and how to combine them.

User guides adapted to NIPH workflows

See Appendix for information on user guides.

One remaining assignment that we suggest continuing with in future projects is producing template language about ML for project leaders to use in protocols and reports. Text has already been extracted from all published protocols and reports but needs to be transformed into template suggestions as well as integrated into the NIPH handbook for systematic reviews.

Manuscripts

Muller AE, Ames HMR, Jardim PSJ, Rose CJ (revision submitted and under review). Comparing automated text clustering with Lingo3G and human research categorization in a rapid review. *Research Synthesis Methods*.

Jardim PSJ, van de Velde S, Rose CJ, Ames HMR, Meneses Echavez JF, Himmels J, Muller AE (in progress). A user-centered study of automating risk of bias in real-life systematic reviews.

Røst T, Slaughter L, Nytrø Ø, Muller AE, Vist GE (in press). “Using neural networks to support high-quality evidence mapping”. *BMC Informatics*.

Presentations

Members of the team gave a number of presentations during spring 2021 (table 8).

Table 8: *Overview of presentations delivered by the ML team*

Date	Presentation title	Context and audience
02.02.2021	Drøfting av planer og aktiviteter lag for maskinlæring	Leader team, Cluster for Reviews and Health Technology Assessments
3.03.2021	Microsoft Academic Graph	Librarian *faggruppe*
23.02.2021	Testing out Microsoft Academic Graph in covid-19 rapid reviews	Citation networks in literature search - web conference, Norwegian Scientific Community for Food and Environment
15.03.2021	Getting to know the machine learning team – who we are and what we are working on	Ukestart meeting, Division for Health Services
06.04.2021	Midtveis rapport	Leader team, Cluster for Reviews and Health Technology Assessments
26.04.2021	Results of a prospective user study of RobotReviewer	Project leaders and members who participated in the user study in the Cluster for Reviews and Health Technology Assessments
08.06.2021	Scaling up machine learning with a dedicated team	Network meeting of evidence synthesis organizations: NIPH, NICE (UK), EPPI Centre (UK), ICQIG (Germany), SBU (Sweden),

		CADTH (Canada), Cochrane, Cochrane Netherlands, MAGICapp
25.05.2021	Proposal for a ML strategy	Leadership group, Cluster for Reviews and Health Technology Assessments
21.06.2021	Hvor mange roboter trenges for å vurdere Risk of Bias?	Ukestart meeting, Divsion for Health Services
23.06.2021	Introduction to HTV's ML team	Network meeting on machine learning and big data: representatives from all divisions + IT
2.11.2021	5 oral presentation abstracts submitted; no decisions yet about acceptance	CADTH online conference: "Uncertain Times, Imperfect Evidence, and the Imperative to Act"

Strategy-related outputs

We developed a proposal for a machine learning strategy for the Cluster for Reviews and Health Technology Assessments. The full strategy is presented in a separate document.

We also proposed a text for NIPH's revised strategic priorities. The following text was submitted to the management in the Division for Health Services in May 2021:

"Context: There is an increasing demand from users for high-quality products delivered faster, with greater efficiency, and at lower cost. There is also a growing societal need for high-quality, understandable, and accessible knowledge. Furthermore, rapid developments in the types of data and advanced methods available are opening opportunities to increase efficiency and speed without compromising on quality. With the revision of the strategy document, we have the opportunity to develop a clear, cross-division commitment to ML and methods innovation that can facilitate the systematic identification and implementation of tools and strategies to benefit a wide variety of products across the institute.

The problem: We have identified machine learning (ML), big data, and advanced analyses included directly or indirectly within several different strategic priorities in the 2019-2014 institute strategy.

- Forutse helsetrusler
- Stor data og avansert analyse
- Sanntidsovervåking
- På tvers av sektorer
- Enklere navigasjon
- Helsedata skal komme til nytte

But these strategies don't appear particularly coordinated or connected – which very likely means untapped opportunities for knowledge transfer, capacity-building, innovation, and de-duplication of work. For example, Jon Bohlin (Smittevern) uses machine learning in epigenetic modelling, Christian Madsen (Psykisk og fysisk helse) to predict maternal outcomes, and Yungsung Lee (Pyskisk og fysisk helse) to predict biological age based on blood samples – similar techniques can be used in vaccine development and in epidemic modeling.

The solution:

- An institution-wide vision: FHI will be an innovative organization that uses machine learning, automation, and big data to deliver our high-quality products (kunnskap, beredskap, and infrastuktur) more effectively, while also increasing accessibility, and sustainability.
- An institutional strategy that brings together the currently disjointed and vertical activities into a more cohesive, mutually beneficial and innovation-oriented collaboration. FHI products (kunnskap, beredskap, infrastuktur) will be stronger if we can facilitate in-house knowledge transfer and coordination. Based on our networking regarding only machine learning, we see quite a lot of internal expertise that can be exploited, as well as numerous opportunities for external collaboration and capacity-building.
- A Center of Excellence for knowledge innovation for machine learning, automation and big data. This will draw together/centralize/coordinate ongoing machine learning, other advanced methods, and workflow optimization projects involving arbeidsflyt, automation, and dating sharing, currently localized in Områder for smittevern, helsetjenester, helsedata og digitalisering, psykisk og fysisk helse, and IT (See figure for an example of the ongoing machine learning activities).

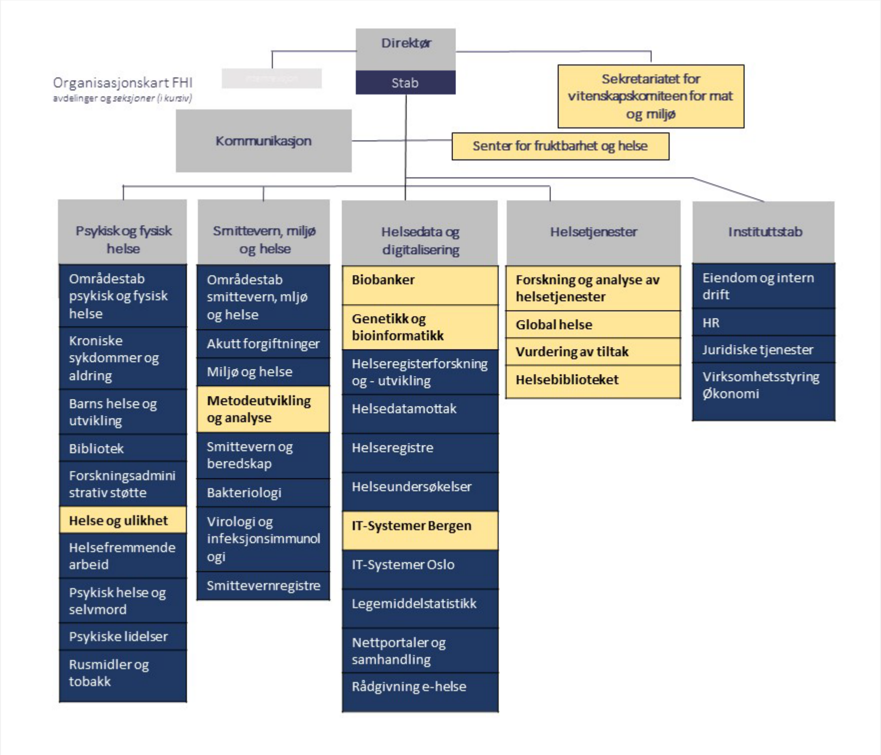

Organisasjonskart FHI
avdelinger og seksjoner (i kursiv)
Direktør
Stab
Sekretariatet for vitenskapskomiteen for mat og miljø
Kommunikasjon
Senter for fruktbarhet og helse
Psykisk og fysisk helse
Områdestab psykisk og fysisk helse
Kroniske sykdommer og aldring
Barns helse og utvikling
Bibliotek
Forskningsadministrativ støtte
Helse og ulikhet
Helsefremmende arbeid
Psykisk helse og selvmord
Psykiske lidelser
Rusmidler og tobakk
Smittevern, miljø og helse
Områdestab smittevern, miljø og helse
Akutt forgiftninger
Miljø og helse
Metodeutvikling og analyse
Smittevern og beredskap
Bakteriologi
Virologi og infeksjonsimmunologi
Smittevernregistre
Helsedata og digitalisering
Biobanker
Genetikk og bioinformatikk
Helseregisterforskning og - utvikling
Helsedatamottak
Helseregistre
Helseundersøkelser
IT-Systemer Bergen
IT-Systemer Oslo
Legemiddelstatistikk
Nettportaler og samhandling
Rådgivning e-helse
Helsetjenester
Forskning og analyse av helsetjenester
Global helse
Vurdering av tiltak
Helsebiblioteket
Instituttstab
Eiendom og intern drift
HR
Juridiske tjenester
Virksomhetsstyring Økonomi

Figure: A rapid mapping of current Machine learning activities (The yellow color represents ongoing activity)

The potential: Synergies that directly benefit existing strategies (see above).

- Through coordinating område-specific activities, internal expertise will be identified and strengthened, and thereby made available for future development.
- Increased efficiency and speed of production, while maintaining/improving quality, in the involved projects and knowledge products. Some examples: faster evidence synthesis in Område for helsetjenester, advanced epidemiological studies in Område for psykisk helse, rapid covid-19 modelling in Område for smittevern.
- Resources and time saved can be 'banked' back into development/innovation efforts.
- This center, and FHI in general, could become a model for other public health institutions (strategic priority: 'Norge i verden'). Through prioritizing ML innovation, we can demonstrate the implementation and success of cross-sectoral, horizontal programs rather than vertical, siloed initiatives."

Lessons learned

We managed to spark interest in ML, and successfully recruited and trained several project leaders and members to apply newly learned methods. Sole one-on-one trainings were, however, not sufficient for immediate method independence. To address this, educational and how-to guides were developed, and in the future, a new constellation of the ML team with more employees involved in distinct short-term roles will support scalability.

This team – initially mostly ML-novices – matured to internal training and implementation experts, through 4-5 weeks of internal capacity-building and peer-teaching. This was a sunk cost and delayed the start of other activities, although served the additional purpose of team-building. For future iterations of the team, recruiting employees with existing skills in ML and software within evidence synthesis would minimize large upfront costs.

Blocking out team members' time allowed them to prioritize ML tasks, which were often naturally de-prioritized in the face of other commissions. Related to this, team members also needed to feel confident that risk-taking was allowed and encouraged; for example, testing out a ML function in a new software for several hours and concluding that it had limited utility was still a valuable use of time.

It is crucial that the ML team continues to recruit "early adopters": employees interested in ML and innovative methods, and willing to adopt and spread new skills and knowledge. It is equally important that the team be critical and aware of ML's limitations, but such constructive criticism should be provided by team members or advisors with ML experience, not by ML-naïve/skeptic team members.

To support ML adoption and acceptability, in-house evaluations can be used, including well-developed and already validated techniques. Involving interested project leaders in the design of these evaluations may also increase subsequent acceptability. These evaluations can also be used to experiment with workflow modifications. The more workflows are changed, the more important it is that project teams feel ownership of or inclusion in those change decisions.

Home-grown, Norwegian-language training materials were popular.

ML can be a disruptive technology within evidence syntheses, although it does not have to be. The time savings we have seen in various phases of our reviews can be received as positive, as well as threatening to one’s usual role and responsibility, or both. We hope that our suggested format of the future team, with rotating short-term members will build trust in ML, but this is not a given: a goal should be to expose as many employees as possible to ML, while ensuring that concerns are heard and addressed.

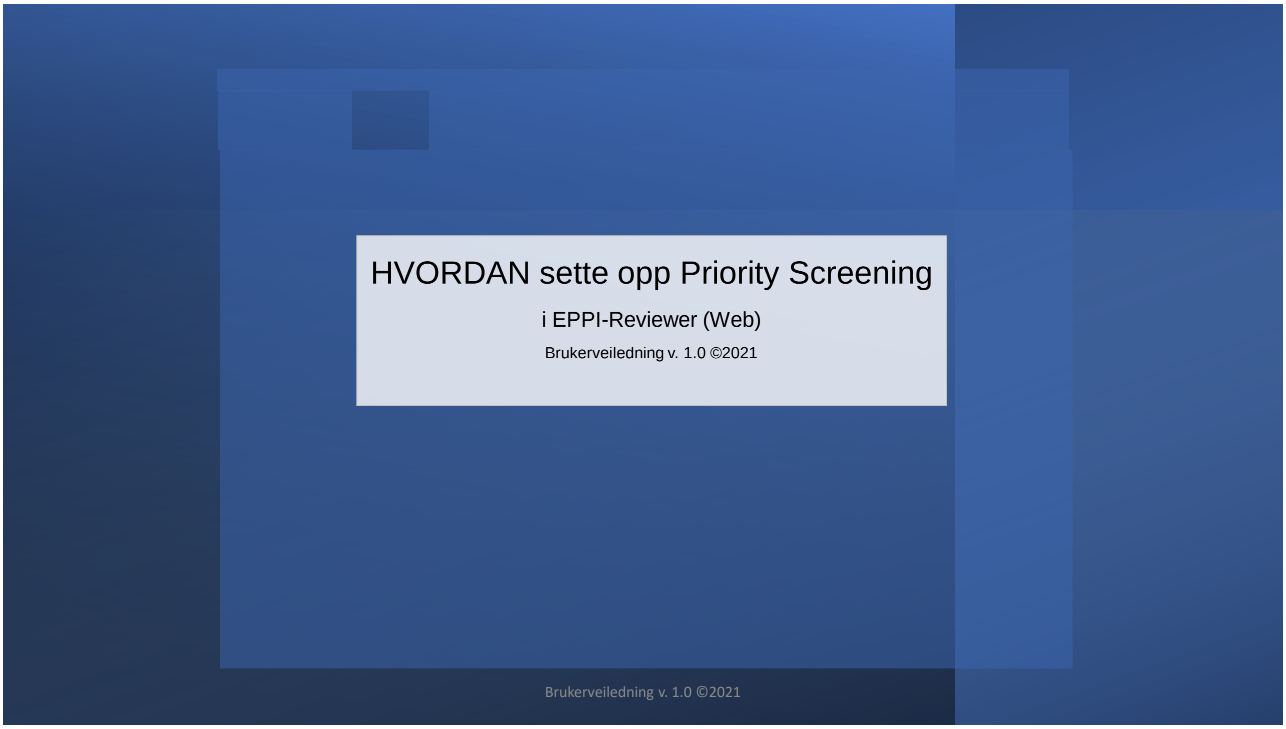
Appendices

Appendix 1: How to put up a priority screening in EPPI-reviewer

Appendix 2: Machine learning classifiers – how to build your own in EPPI 4

Appendix 3: Risk of Bias assessments with machine learning – Team leaders

Appendix 4: Risk of Bias assessments with machine learning – Team members

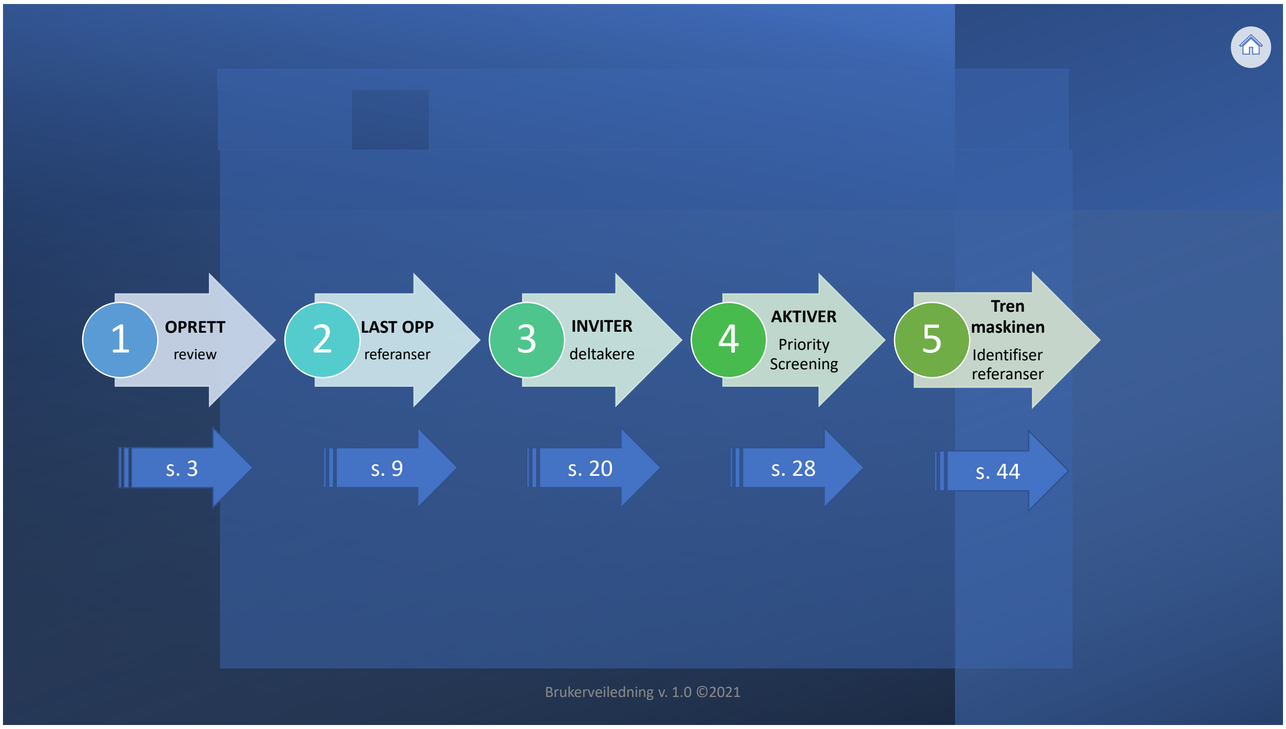


HVORDAN sette opp Priority Screening

i EPPI-Reviewer (Web)

Brukerveiledning v. 1.0 ©2021

1



1 OPPRETT review

2 LAST OPP referanser

3 INVITER deltakere

4 AKTIVER Priority Screening

5 Tren maskinen Identifiser referanser

s. 3

s. 9

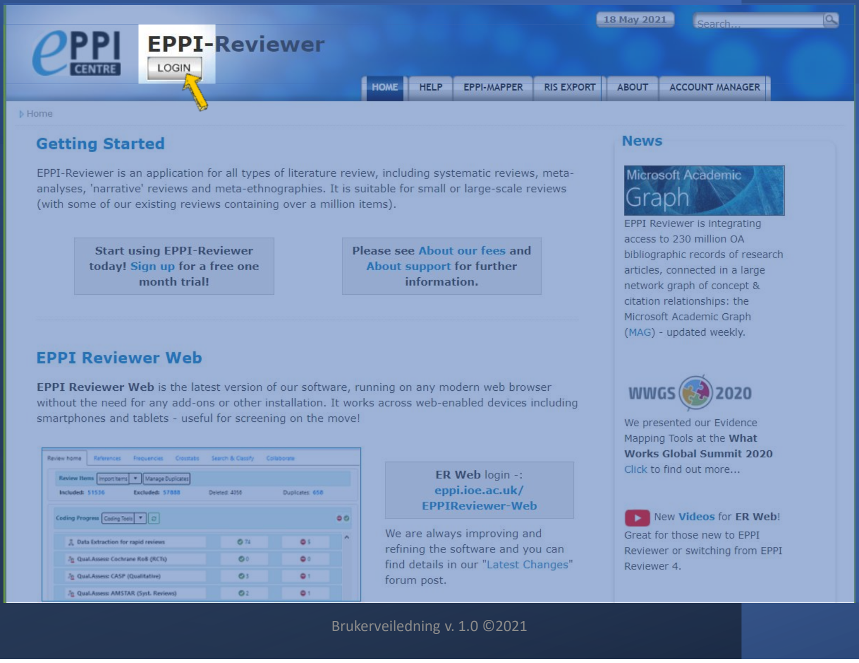
s. 20

s. 28

s. 44

2

https://eppi.ioe.ac.uk/eppireviewer-web/home



Getting Started

EPPI-Reviewer is an application for all types of literature review, including systematic reviews, meta-analyses, 'narrative' reviews and meta-ethnographies. It is suitable for small or large-scale reviews (with some of our existing reviews containing over a million items).

Start using EPPI-Reviewer today! Sign up for a free one month trial!

Please see About our fees and About support for further information.

EPPI Reviewer Web

EPPI Reviewer Web is the latest version of our software, running on any modern web browser without the need for any add-ons or other installation. It works across web-enabled devices including smartphones and tablets - useful for screening on the move!

Included: 51536 Excluded: 57888 Deleted: 4056 Duplicates: 658

Data Extraction for rapid reviews	74	5
Qual.Assess: Cochrane RoB (RCTs)	0	0
Qual.Assess: CASP (Qualitative)	3	1
Qual.Assess: AMSTAR (Syst. Reviews)	2	1

ER Web login -:
eppi.ioe.ac.uk/ EPPIReviewer-Web

We are always improving and refining the software and you can find details in our "Latest Changes" forum post.

News

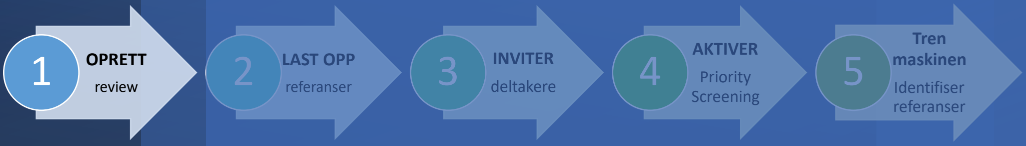
EPPI Reviewer is integrating access to 230 million OA bibliographic records of research articles, connected in a large network graph of concept & citation relationships: the Microsoft Academic Graph (MAG) - updated weekly.

We presented our Evidence Mapping Tools at the **What Works Global Summit 2020** Click to find out more...

New **Videos** for **ER Web**! Great for those new to EPPI Reviewer or switching from EPPI Reviewer 4.

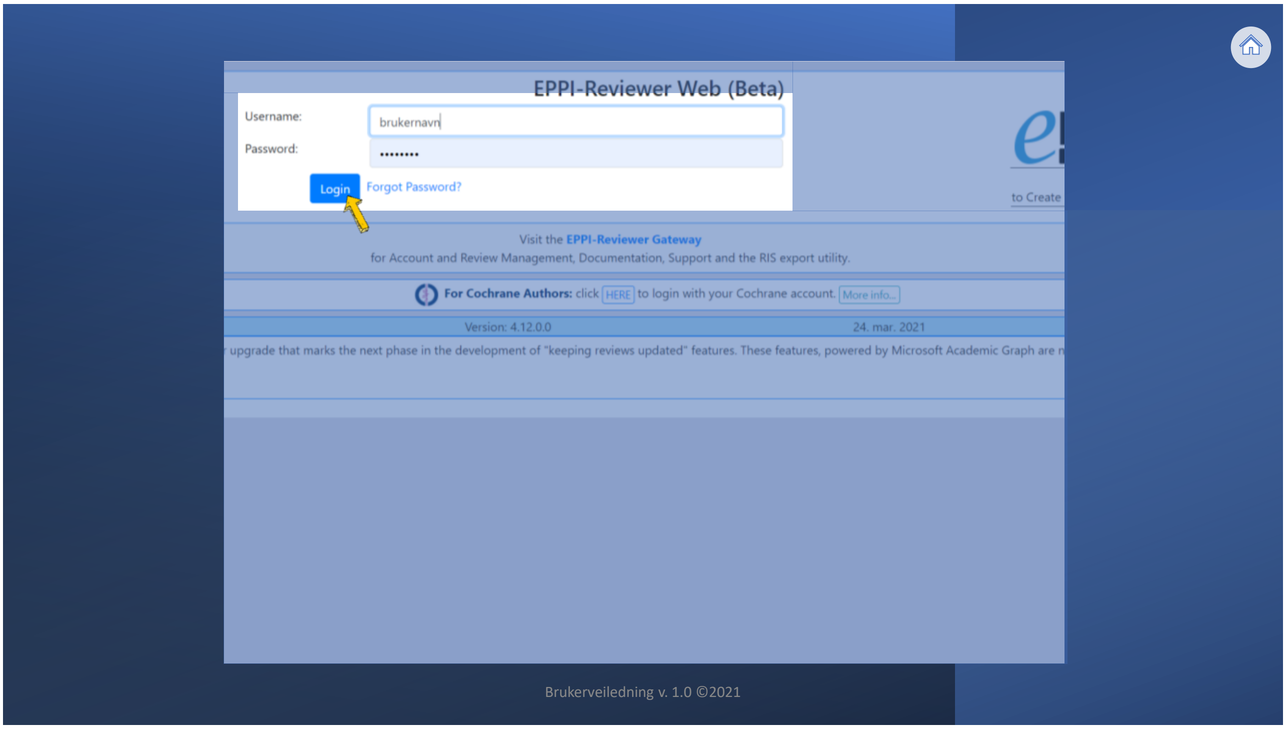
Brukerveiledning v. 1.0 ©2021

3


1 OPPRETT review
2 LAST OPP referanser
3 INVITER deltakere
4 AKTIVER Priority Screening
5 Tren maskinen Identifiser referanser

Brukerveiledning v. 1.0 ©2021

4



EPPI-Reviewer Web (Beta)

Username: brukernavn

Password: ••••••••

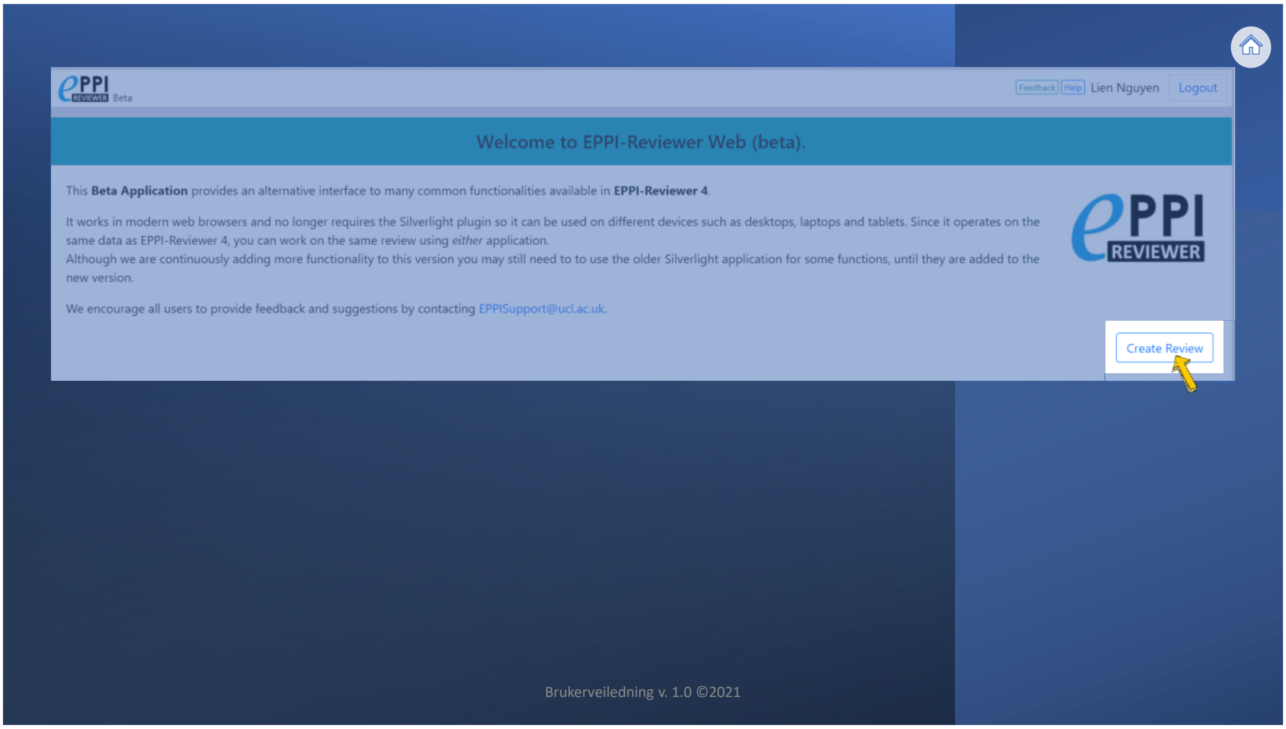
Visit the **EPPI-Reviewer Gateway**
for Account and Review Management, Documentation, Support and the RIS export utility.

For Cochrane Authors: click HERE to login with your Cochrane account. More info...

Version: 4.12.0.0 24. mar. 2021

upgrade that marks the next phase in the development of "keeping reviews updated" features. These features, powered by Microsoft Academic Graph are n

5



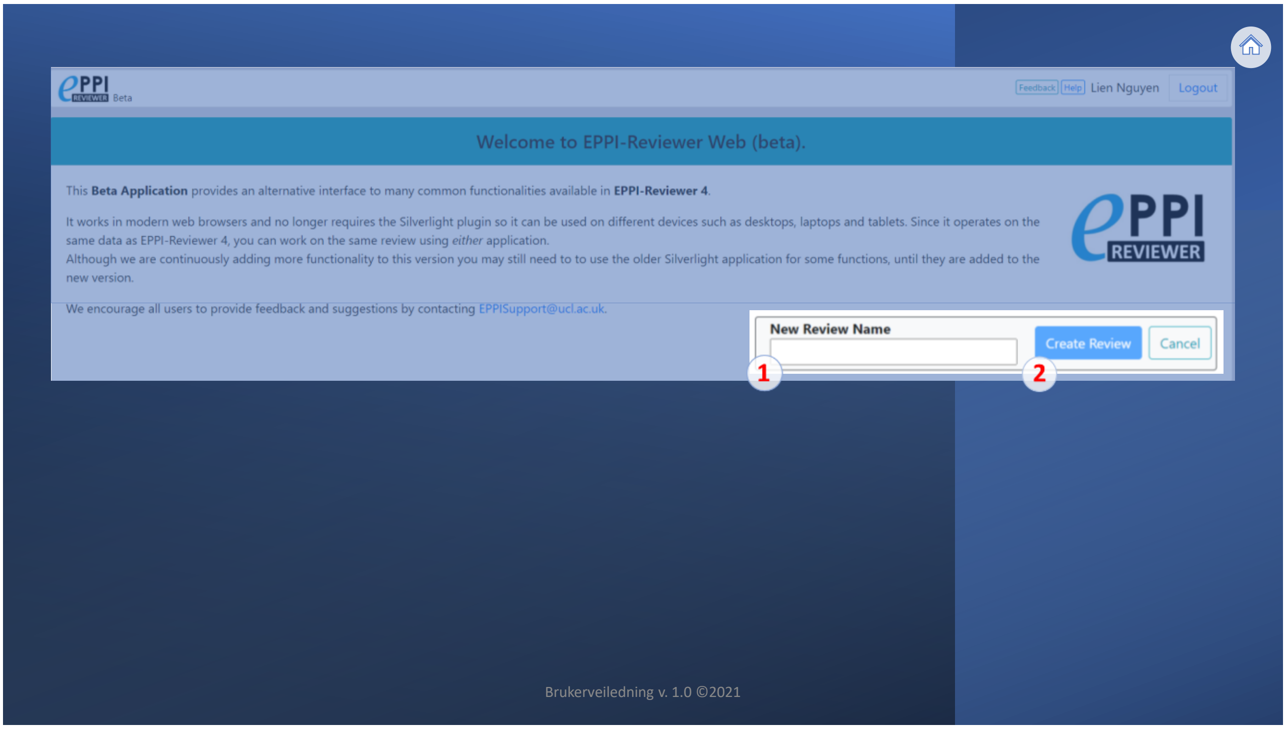
Welcome to EPPI-Reviewer Web (beta).

This **Beta Application** provides an alternative interface to many common functionalities available in **EPPI-Reviewer 4**.

It works in modern web browsers and no longer requires the Silverlight plugin so it can be used on different devices such as desktops, laptops and tablets. Since it operates on the same data as EPPI-Reviewer 4, you can work on the same review using *either* application.
Although we are continuously adding more functionality to this version you may still need to to use the older Silverlight application for some functions, until they are added to the new version.

We encourage all users to provide feedback and suggestions by contacting EPPISupport@ucl.ac.uk.

6



Welcome to EPPI-Reviewer Web (beta).

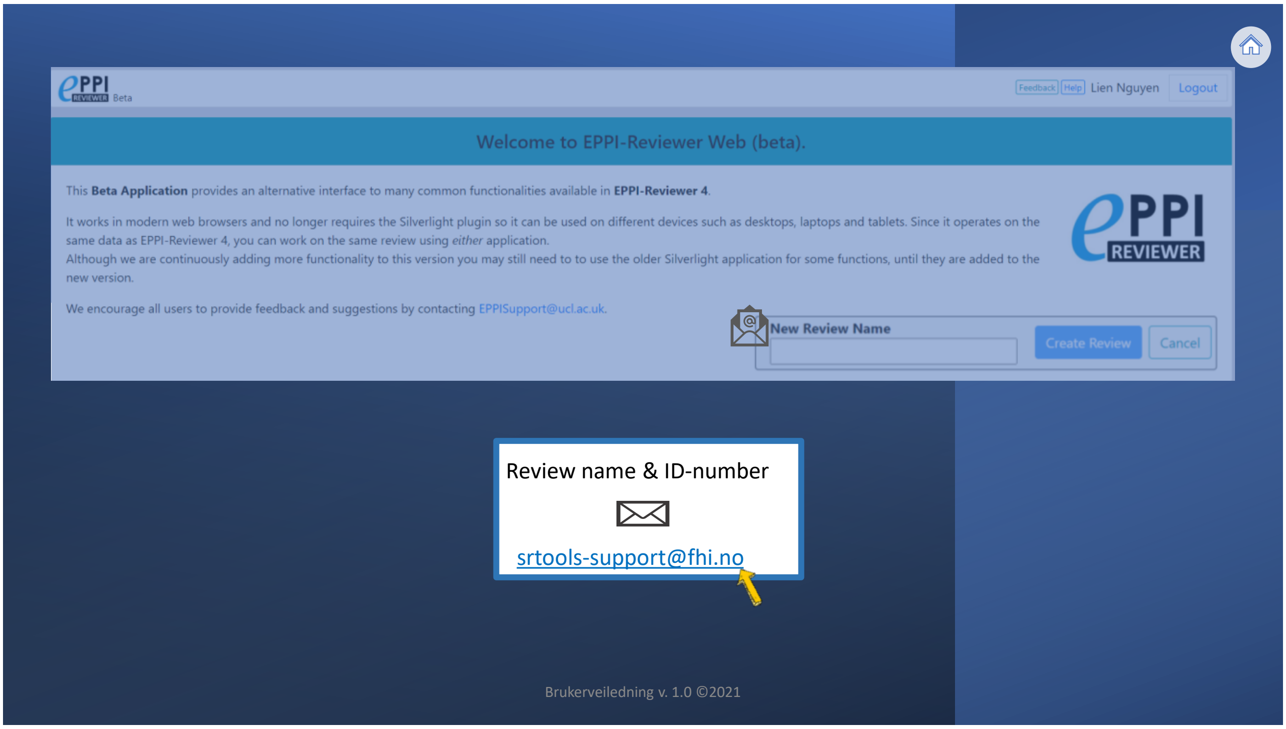
This **Beta Application** provides an alternative interface to many common functionalities available in **EPPI-Reviewer 4**.

It works in modern web browsers and no longer requires the Silverlight plugin so it can be used on different devices such as desktops, laptops and tablets. Since it operates on the same data as EPPI-Reviewer 4, you can work on the same review using *either* application.
Although we are continuously adding more functionality to this version you may still need to to use the older Silverlight application for some functions, until they are added to the new version.

We encourage all users to provide feedback and suggestions by contacting EPPISupport@ucl.ac.uk.

Brukerveiledning v. 1.0 ©2021

7



Welcome to EPPI-Reviewer Web (beta).

This **Beta Application** provides an alternative interface to many common functionalities available in **EPPI-Reviewer 4**.

It works in modern web browsers and no longer requires the Silverlight plugin so it can be used on different devices such as desktops, laptops and tablets. Since it operates on the same data as EPPI-Reviewer 4, you can work on the same review using *either* application.
Although we are continuously adding more functionality to this version you may still need to to use the older Silverlight application for some functions, until they are added to the new version.

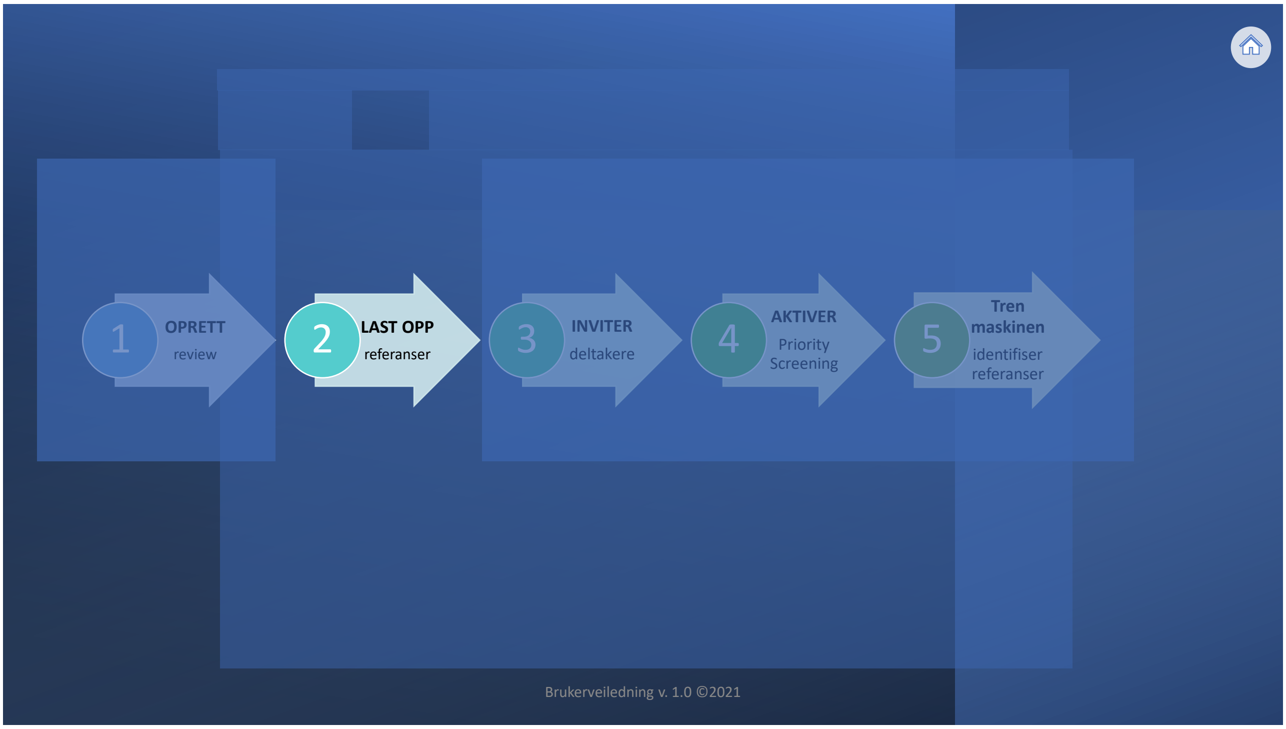
We encourage all users to provide feedback and suggestions by contacting EPPISupport@ucl.ac.uk.

Review name & ID-number

srtools-support@fhi.no

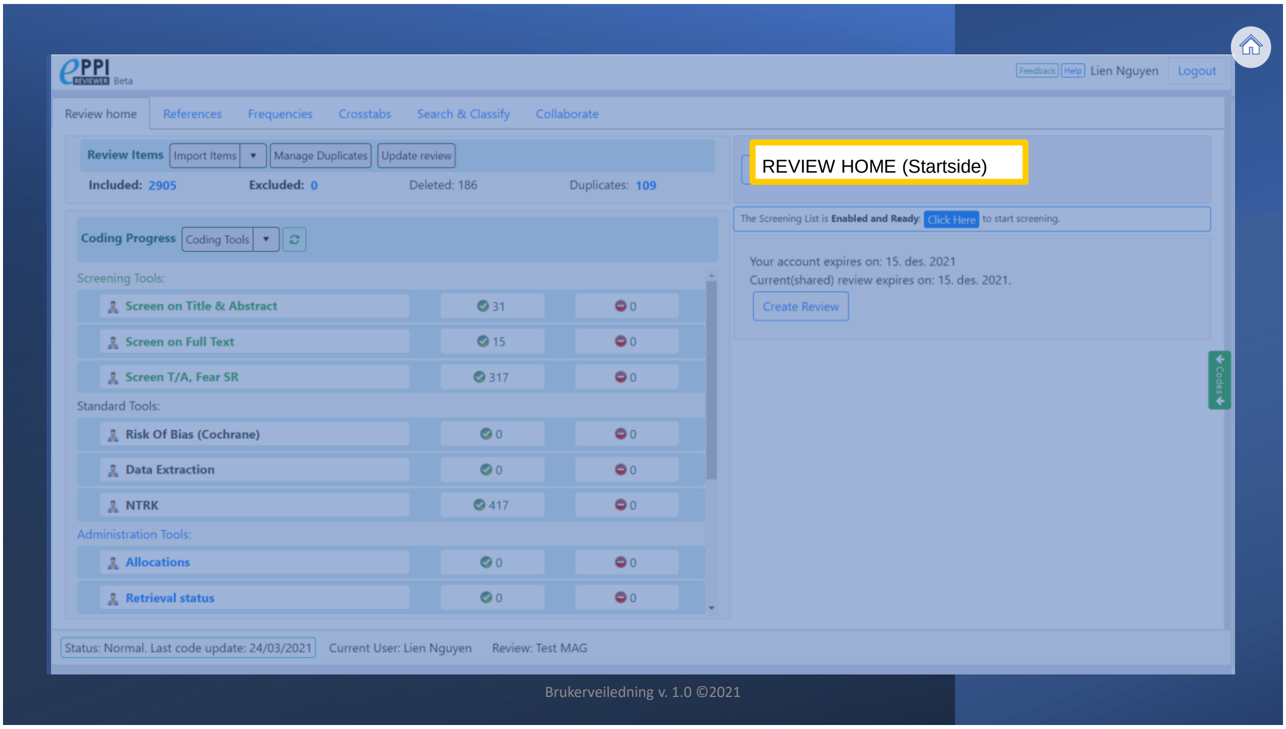
Brukerveiledning v. 1.0 ©2021

8



Brukerveiledning v. 1.0 ©2021

9



REVIEW HOME (Startside)

Lien Nguyen

Review Items

Included: 2905 Excluded: 0 Deleted: 186 Duplicates: 109

The Screening List is **Enabled and Ready**: Click Here to start screening.

Your account expires on: 15. des. 2021
Current(shared) review expires on: 15. des. 2021.

Coding Progress

Screening Tools:

Screen on Title & Abstract	✓ 31	⊖ 0
Screen on Full Text	✓ 15	⊖ 0
Screen T/A, Fear SR	✓ 317	⊖ 0

Standard Tools:

Risk Of Bias (Cochrane)	✓ 0	⊖ 0
Data Extraction	✓ 0	⊖ 0
NTRK	✓ 417	⊖ 0

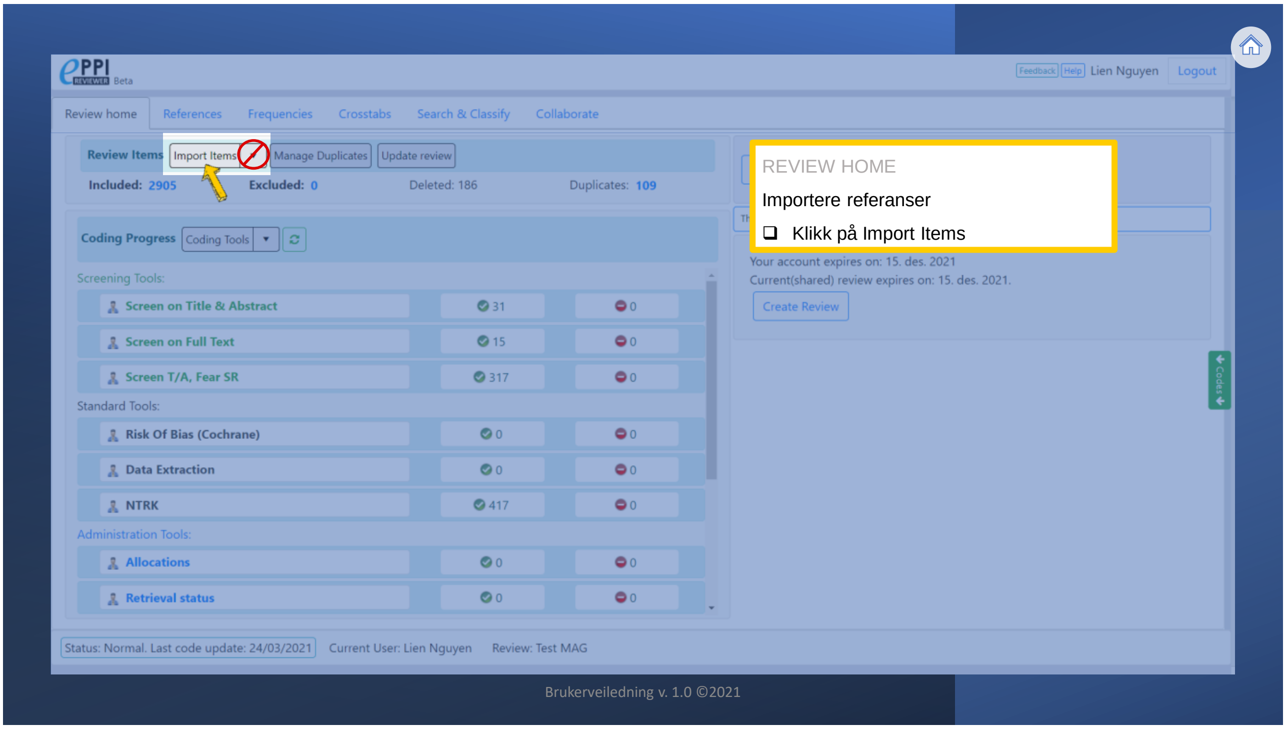
Administration Tools:

Allocations	✓ 0	⊖ 0
Retrieval status	✓ 0	⊖ 0

Status: Normal. Last code update: 24/03/2021 Current User: Lien Nguyen Review: Test MAG

Brukerveiledning v. 1.0 ©2021

10



Included: 2905 Excluded: 0 Deleted: 186 Duplicates: 109

Screening Tools:

Screen on Title & Abstract	✓ 31	⊖ 0
Screen on Full Text	✓ 15	⊖ 0
Screen T/A, Fear SR	✓ 317	⊖ 0

Standard Tools:

Risk Of Bias (Cochrane)	✓ 0	⊖ 0
Data Extraction	✓ 0	⊖ 0
NTRK	✓ 417	⊖ 0

Administration Tools:

Allocations	✓ 0	⊖ 0
Retrieval status	✓ 0	⊖ 0

Your account expires on: 15. des. 2021
Current(shared) review expires on: 15. des. 2021.

Status: Normal. Last code update: 24/03/2021 Current User: Lien Nguyen Review: Test MAG

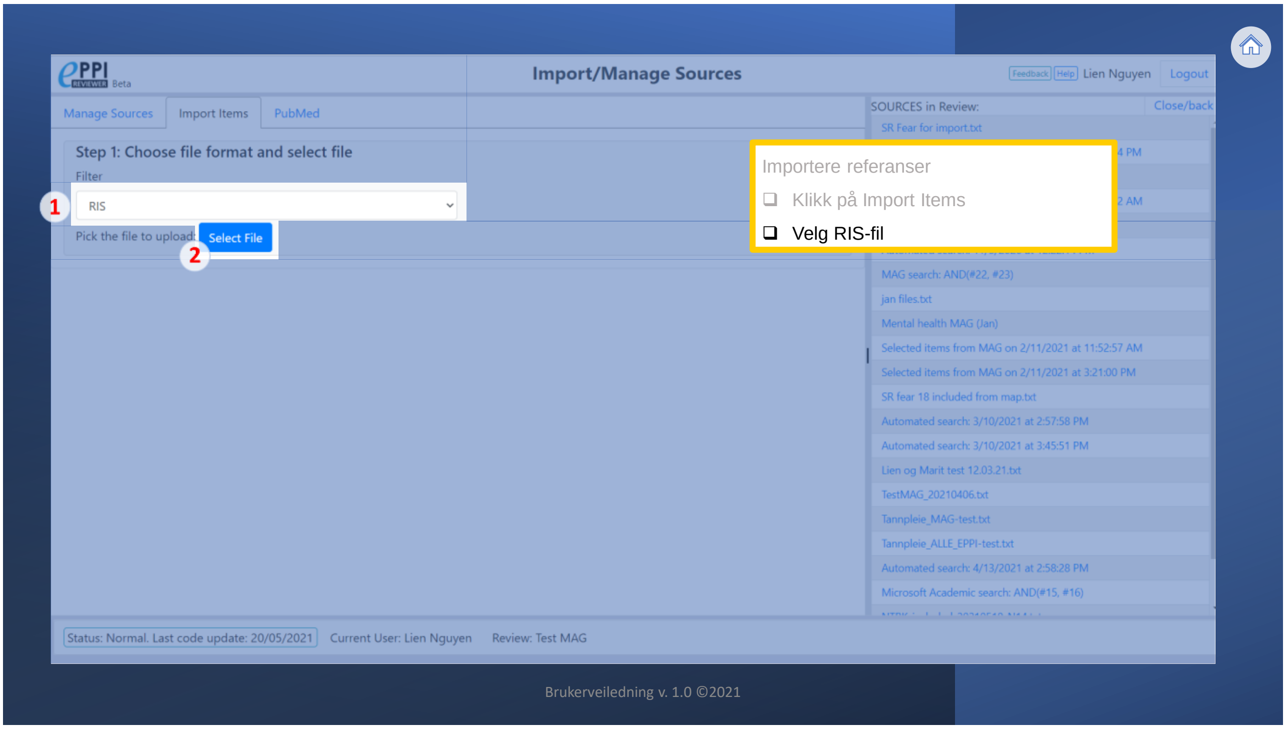
REVIEW HOME

Importere referanser

❑ Klikk på Import Items

Brukerveiledning v. 1.0 ©2021

11



Import/Manage Sources

Step 1: Choose file format and select file

Filter: RIS

Pick the file to upload:

SOURCES in Review:

- SR Fear for import.txt
- MAG search: AND(#22, #23)
- jan files.txt
- Mental health MAG (Jan)
- Selected items from MAG on 2/11/2021 at 11:52:57 AM
- Selected items from MAG on 2/11/2021 at 3:21:00 PM
- SR fear 18 included from map.txt
- Automated search: 3/10/2021 at 2:57:58 PM
- Automated search: 3/10/2021 at 3:45:51 PM
- Lien og Marit test 12.03.21.txt
- TestMAG_20210406.txt
- Tannpleie_MAG-test.txt
- Tannpleie_ALLE_EPPI-test.txt
- Automated search: 4/13/2021 at 2:58:28 PM
- Microsoft Academic search: AND(#15, #16)

Status: Normal. Last code update: 20/05/2021 Current User: Lien Nguyen Review: Test MAG

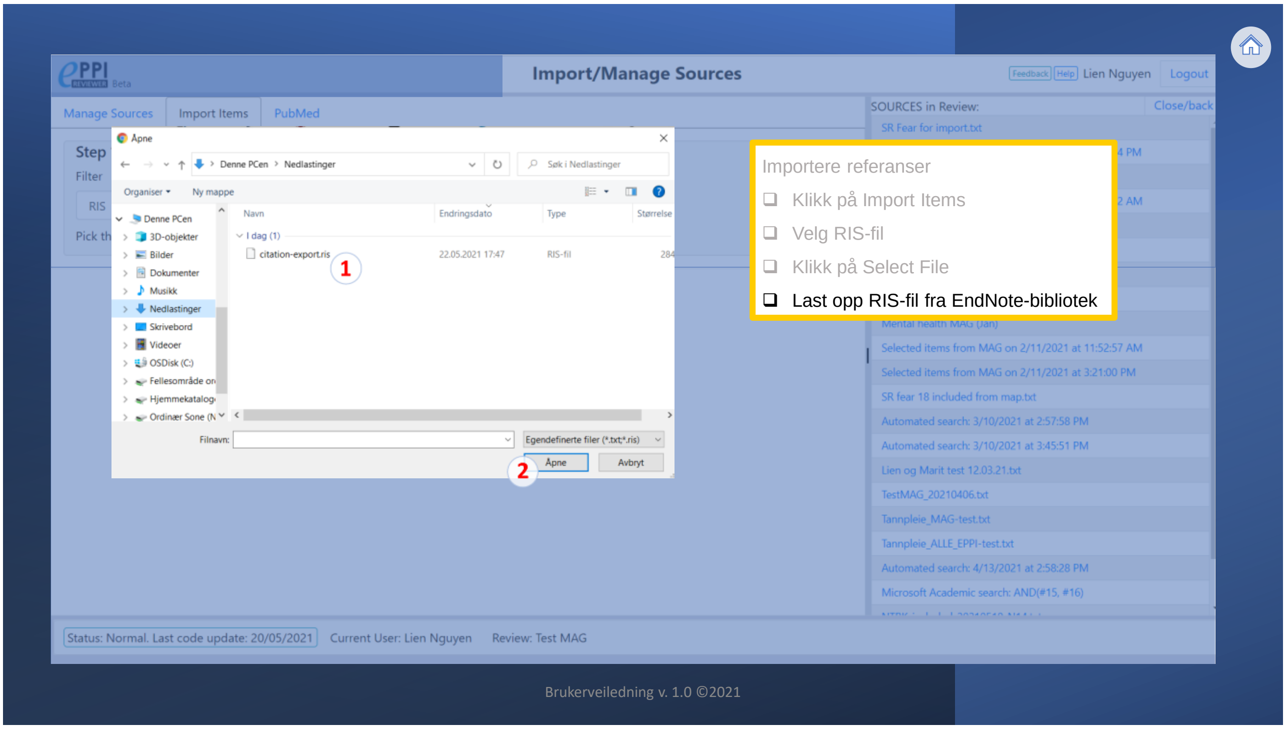
Importere referanser

❑ Klikk på Import Items

❑ Velg RIS-fil

Brukerveiledning v. 1.0 ©2021

12



citation-export.ris 22.05.2021 17:47 RIS-fil 284

Filnavn:

Egendefinerte filer (*.txt;*.ris)

SOURCES in Review:

- SR Fear for import.txt
- Mental health MAG (Jan)
- Selected items from MAG on 2/11/2021 at 11:52:57 AM
- Selected items from MAG on 2/11/2021 at 3:21:00 PM
- SR fear 18 included from map.txt
- Automated search: 3/10/2021 at 2:57:58 PM
- Automated search: 3/10/2021 at 3:45:51 PM
- Lien og Marit test 12.03.21.txt
- TestMAG_20210406.txt
- Tannpleie_MAG-test.txt
- Tannpleie_ALLE_EPPI-test.txt
- Automated search: 4/13/2021 at 2:58:28 PM
- Microsoft Academic search: AND(#15, #16)

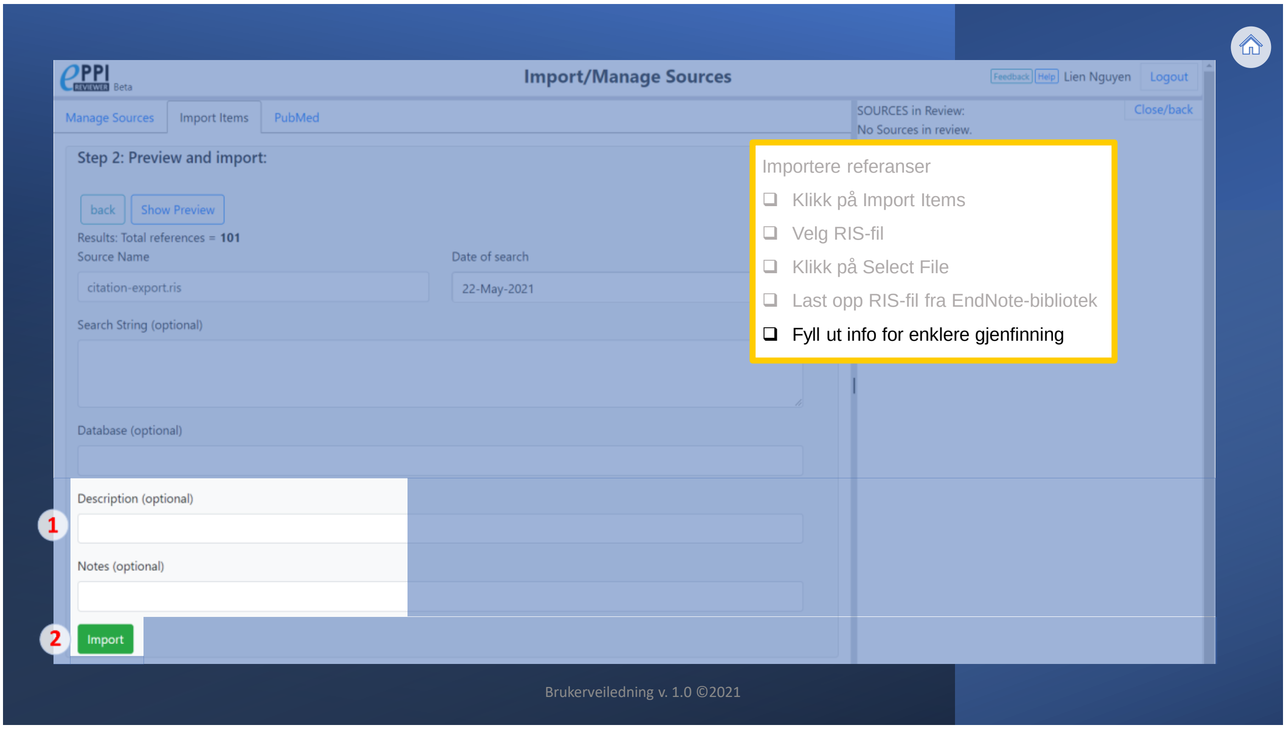
Status: Normal. Last code update: 20/05/2021 Current User: Lien Nguyen Review: Test MAG

Importere referanser

- ❑ Klikk på Import Items
- ❑ Velg RIS-fil
- ❑ Klikk på Select File
- ❑ Last opp RIS-fil fra EndNote-bibliotek

Brukerveiledning v. 1.0 ©2021

13



No Sources in review.

Step 2: Preview and import:

Results: Total references = 101

Source Name: citation-export.ris

Date of search: 22-May-2021

Search String (optional)

Database (optional)

Description (optional)

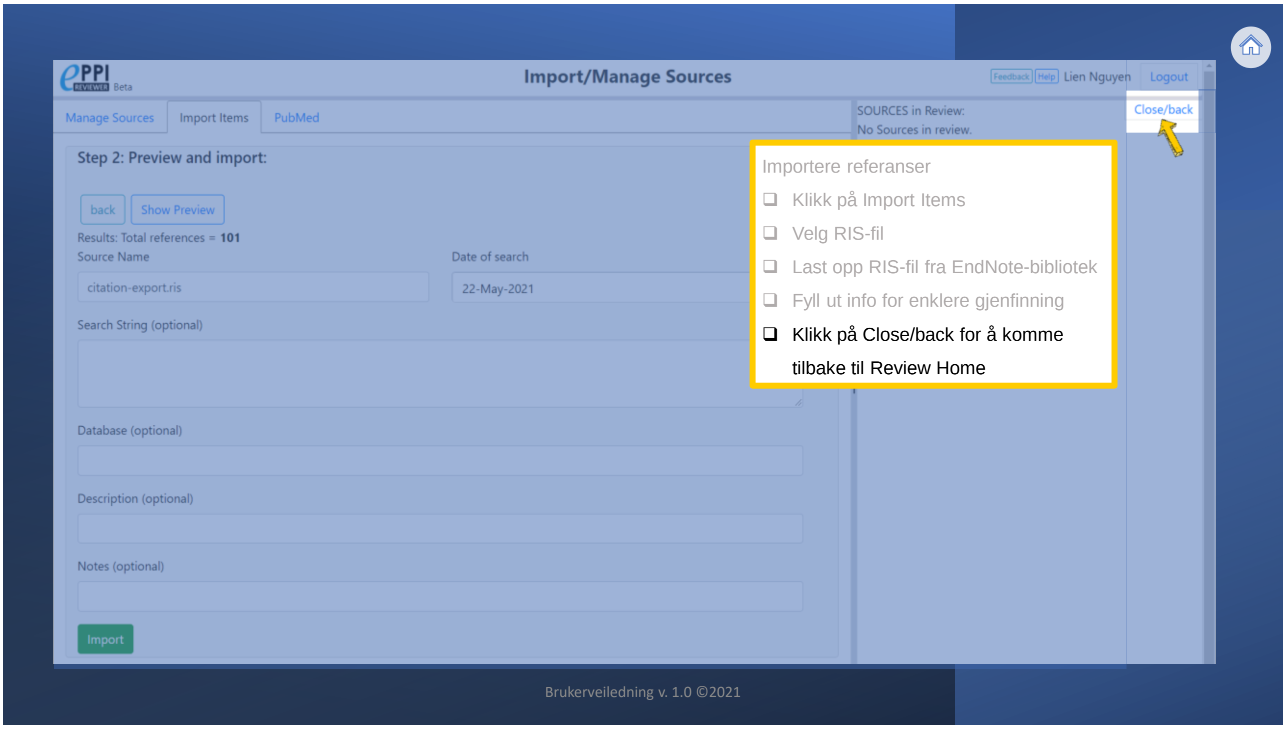
Notes (optional)

Importere referanser

- ❑ Klikk på Import Items
- ❑ Velg RIS-fil
- ❑ Klikk på Select File
- ❑ Last opp RIS-fil fra EndNote-bibliotek
- ❑ Fyll ut info for enklere gjenfinning

Brukerveiledning v. 1.0 ©2021

14



Import/Manage Sources

Manage Sources | Import Items | PubMed

Step 2: Preview and import:

back | Show Preview

Results: Total references = **101**

Source Name: citation-export.ris

Date of search: 22-May-2021

Search String (optional)

Database (optional)

Description (optional)

Notes (optional)

Import

SOURCES in Review:
No Sources in review.

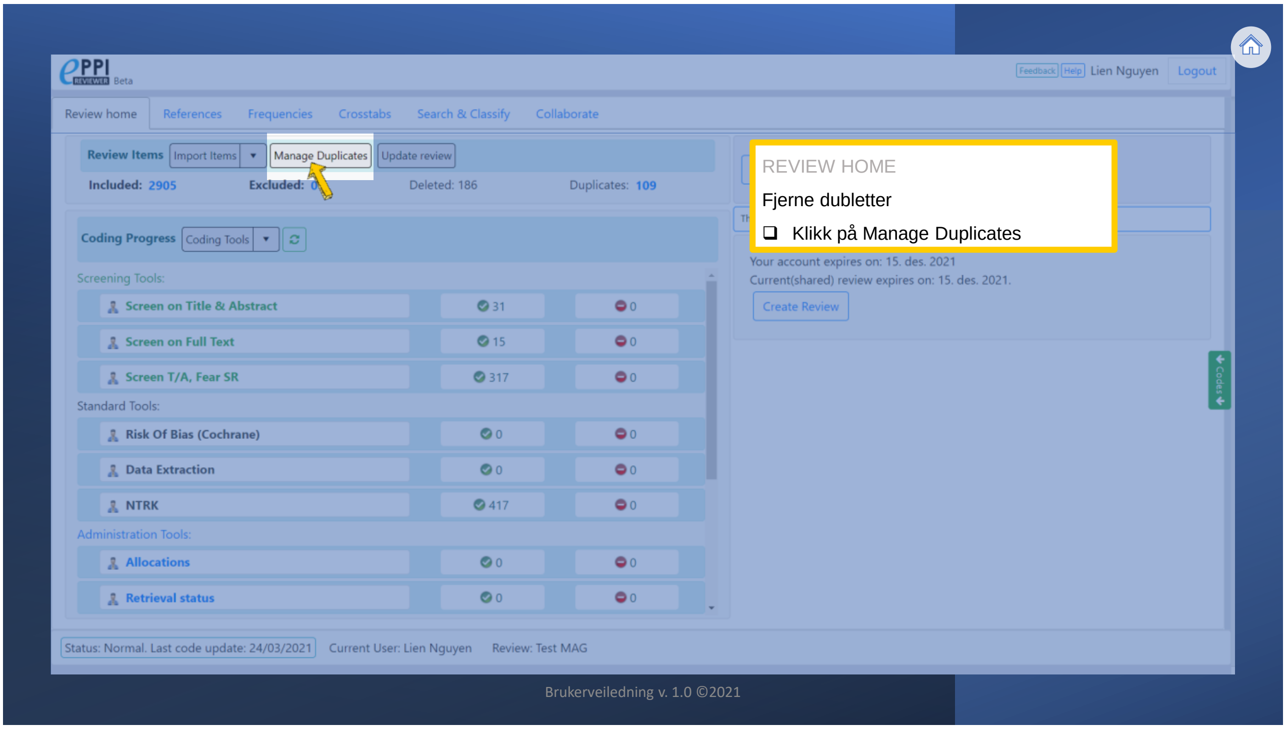
Close/back

Importere referanser

- ❑ Klikk på Import Items
- ❑ Velg RIS-fil
- ❑ Last opp RIS-fil fra EndNote-bibliotek
- ❑ Fyll ut info for enklere gjenfinning
- ❑ Klikk på Close/back for å komme tilbake til Review Home

Brukerveiledning v. 1.0 ©2021

15



Review home | References | Frequencies | Crosstabs | Search & Classify | Collaborate

Review Items | Import Items | Manage Duplicates | Update review

Included: 2905 Excluded: 0 Deleted: 186 Duplicates: 109

Coding Progress | Coding Tools

Screening Tools:

Screen on Title & Abstract	31	0
Screen on Full Text	15	0
Screen T/A, Fear SR	317	0

Standard Tools:

Risk Of Bias (Cochrane)	0	0
Data Extraction	0	0
NTRK	417	0

Administration Tools:

Allocations	0	0
Retrieval status	0	0

Your account expires on: 15. des. 2021
Current(shared) review expires on: 15. des. 2021.

Create Review

Status: Normal. Last code update: 24/03/2021 Current User: Lien Nguyen Review: Test MAG

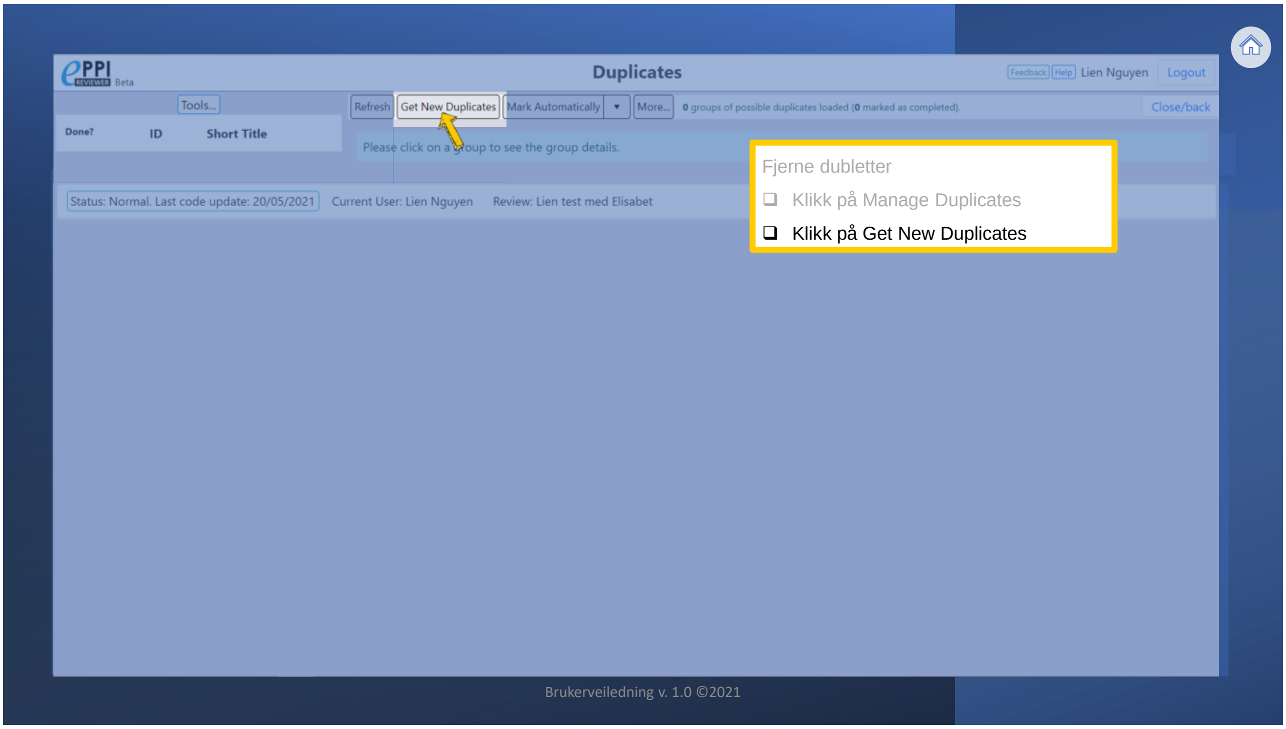
REVIEW HOME

Fjerne dubletter

- ❑ Klikk på Manage Duplicates

Brukerveiledning v. 1.0 ©2021

16



0 groups of possible duplicates loaded (**0** marked as completed).

Please click on a group to see the group details.

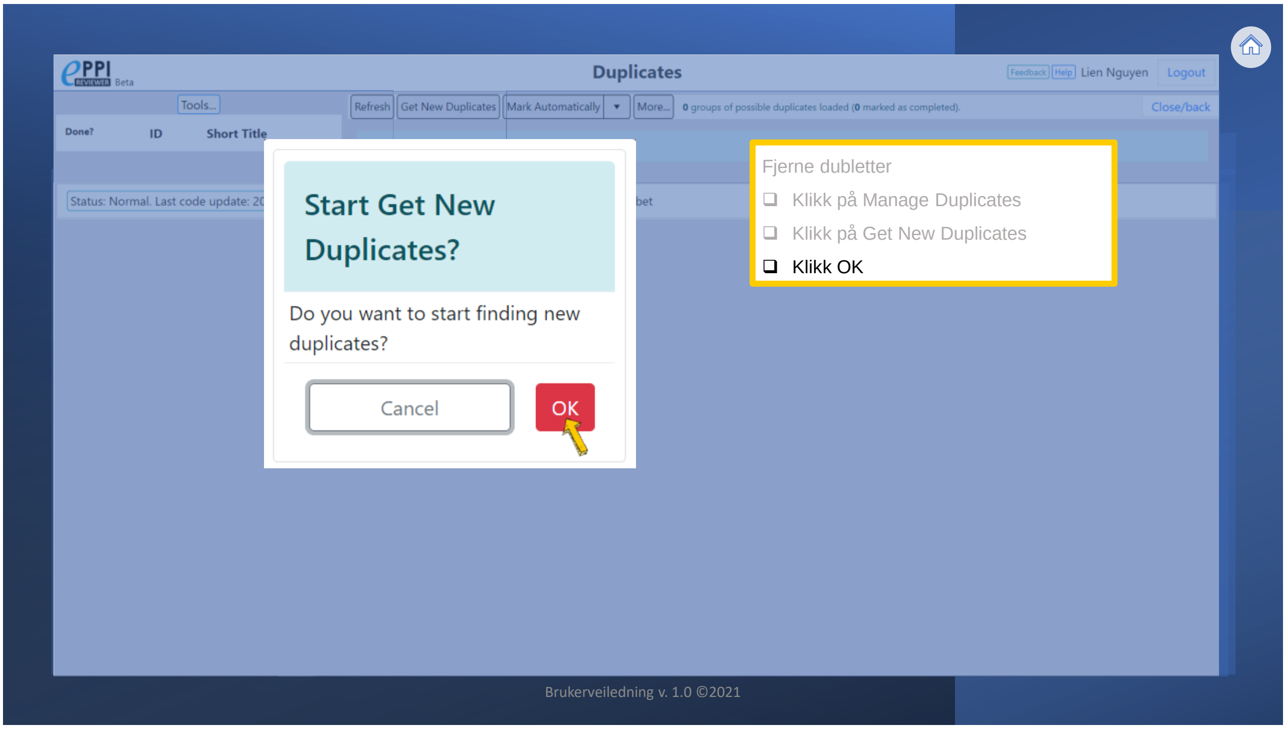
Status: Normal. Last code update: 20/05/2021 Current User: Lien Nguyen Review: Lien test med Elisabet

Fjerne dubletter

- ❑ Klikk på Manage Duplicates
- ❑ Klikk på Get New Duplicates

Brukerveiledning v. 1.0 ©2021

17



0 groups of possible duplicates loaded (**0** marked as completed).

Status: Normal. Last code update: 20

Start Get New Duplicates?

Do you want to start finding new duplicates?

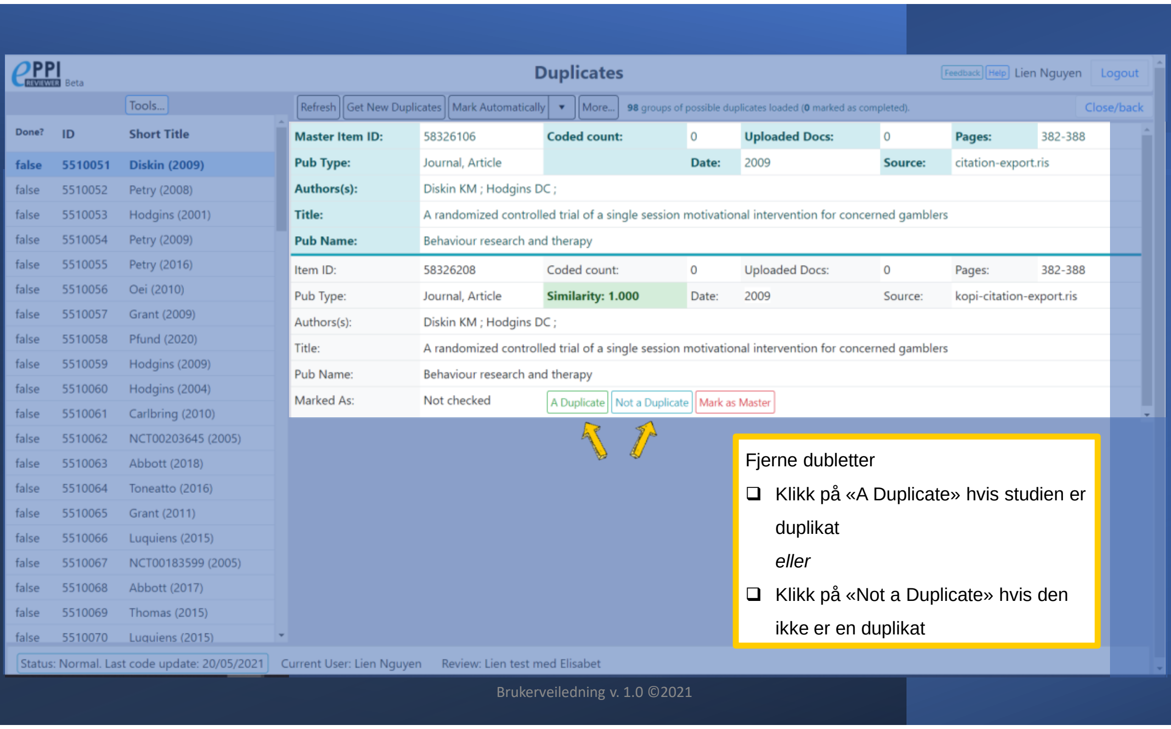
Cancel OK

Fjerne dubletter

- ❑ Klikk på Manage Duplicates
- ❑ Klikk på Get New Duplicates
- ❑ Klikk OK

Brukerveiledning v. 1.0 ©2021

18



Duplicates

98 groups of possible duplicates loaded (0 marked as completed).

Done?	ID	Short Title
false	5510051	Diskin (2009)
false	5510052	Petry (2008)
false	5510053	Hodgins (2001)
false	5510054	Petry (2009)
false	5510055	Petry (2016)
false	5510056	Oei (2010)
false	5510057	Grant (2009)
false	5510058	Pfund (2020)
false	5510059	Hodgins (2009)
false	5510060	Hodgins (2004)
false	5510061	Carlbring (2010)
false	5510062	NCT00203645 (2005)
false	5510063	Abbott (2018)
false	5510064	Toneatto (2016)
false	5510065	Grant (2011)
false	5510066	Luquiens (2015)
false	5510067	NCT00183599 (2005)
false	5510068	Abbott (2017)
false	5510069	Thomas (2015)
false	5510070	Luquiens (2015)

Master Item ID:	58326106	**Coded count:**	0	**Uploaded Docs:**	0	**Pages:**	382-388
Pub Type:	Journal, Article		**Date:**	2009	**Source:**	citation-export.ris	
Authors(s):	Diskin KM ; Hodgins DC ;						
Title:	A randomized controlled trial of a single session motivational intervention for concerned gamblers						
Pub Name:	Behaviour research and therapy						
Item ID:	58326208	Coded count:	0	Uploaded Docs:	0	Pages:	382-388
Pub Type:	Journal, Article	**Similarity: 1.000**	Date:	2009	Source:	kopi-citation-export.ris	
Authors(s):	Diskin KM ; Hodgins DC ;						
Title:	A randomized controlled trial of a single session motivational intervention for concerned gamblers						
Pub Name:	Behaviour research and therapy						
Marked As:	Not checked	A Duplicate	Not a Duplicate	Mark as Master			

Status: Normal. Last code update: 20/05/2021 Current User: Lien Nguyen Review: Lien test med Elisabet

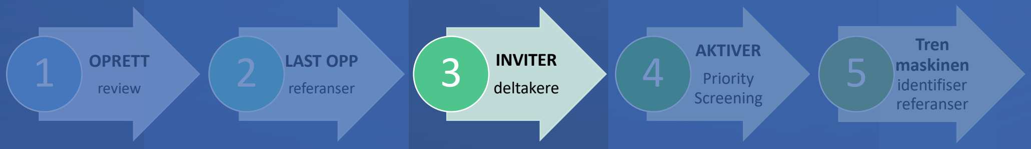
Fjerne dubletter

- ❑ Klikk på «A Duplicate» hvis studien er duplikat

  eller

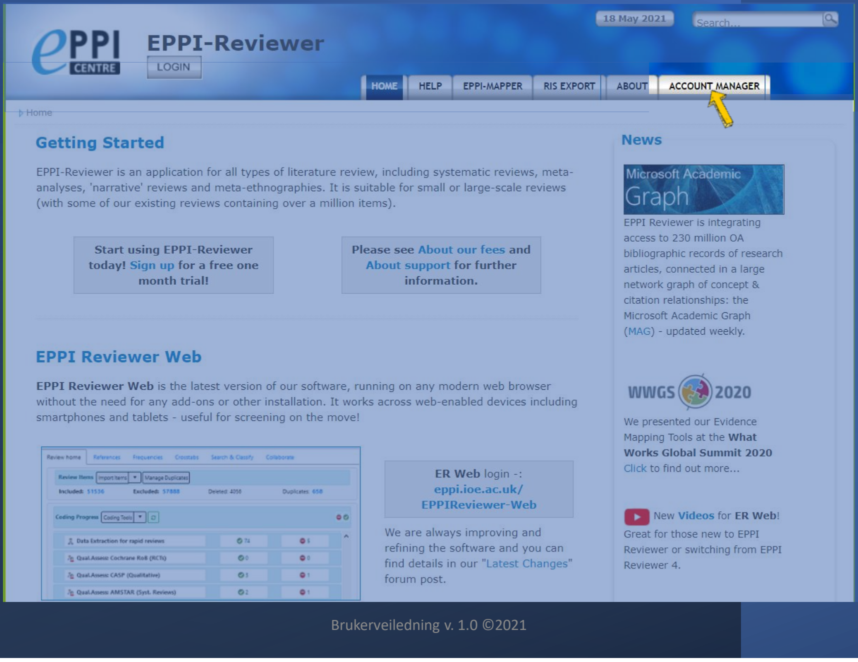
- ❑ Klikk på «Not a Duplicate» hvis den ikke er en duplikat

19



20

https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2914



Getting Started

EPPI-Reviewer is an application for all types of literature review, including systematic reviews, meta-analyses, 'narrative' reviews and meta-ethnographies. It is suitable for small or large-scale reviews (with some of our existing reviews containing over a million items).

Start using EPPI-Reviewer today! Sign up for a free one month trial!

Please see About our fees and About support for further information.

EPPI Reviewer Web

EPPI Reviewer Web is the latest version of our software, running on any modern web browser without the need for any add-ons or other installation. It works across web-enabled devices including smartphones and tablets - useful for screening on the move!

ER Web login -:
eppi.ioe.ac.uk/EPPIReviewer-Web

We are always improving and refining the software and you can find details in our "Latest Changes" forum post.

News

Microsoft Academic Graph

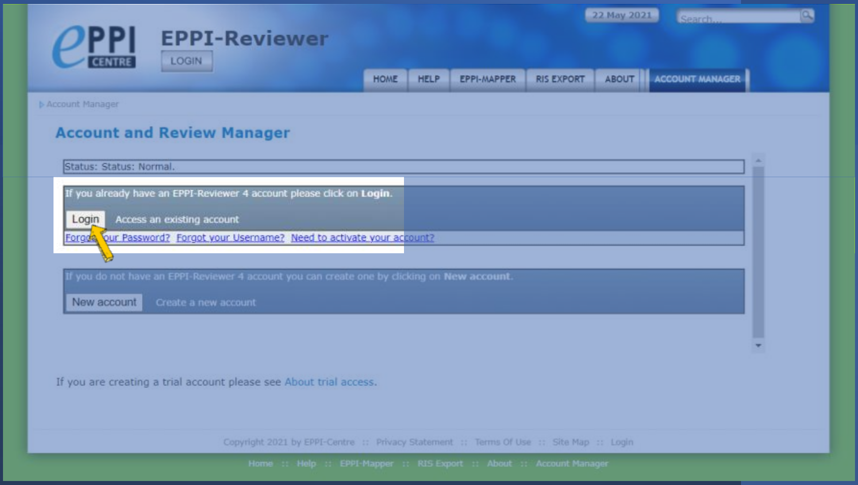
EPPI Reviewer is integrating access to 230 million OA bibliographic records of research articles, connected in a large network graph of concept & citation relationships: the Microsoft Academic Graph (MAG) - updated weekly.

WWGS 2020

We presented our Evidence Mapping Tools at the **What Works Global Summit 2020** Click to find out more...

New **Videos** for **ER Web**! Great for those new to EPPI Reviewer or switching from EPPI Reviewer 4.

21



Account and Review Manager

Status: Status: Normal.

If you already have an EPPI-Reviewer 4 account please click on **Login**.

Login Access an existing account

Forgot your Password? Forgot your Username? Need to activate your account?

If you do not have an EPPI-Reviewer 4 account you can create one by clicking on **New account**.

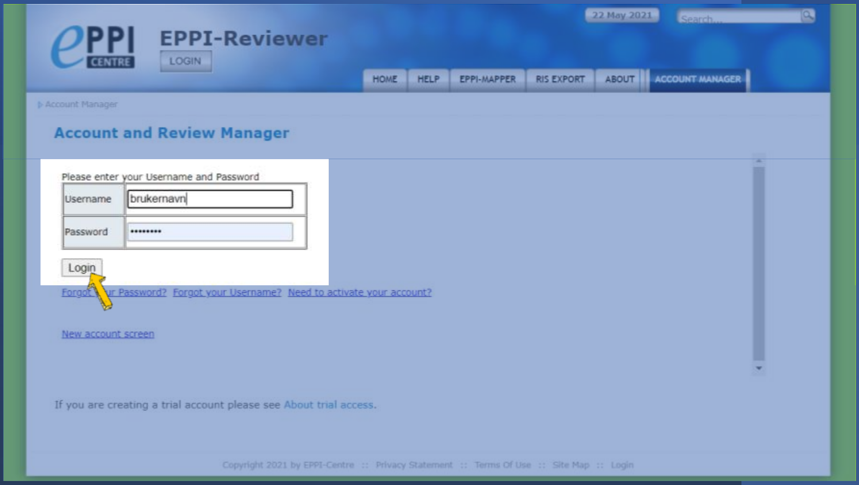
New account Create a new account

If you are creating a trial account please see About trial access.

Copyright 2021 by EPPI-Centre :: Privacy Statement :: Terms Of Use :: Site Map :: Login

Home :: Help :: EPPI-Mapper :: RIS Export :: About :: Account Manager

22



Account and Review Manager

Please enter your Username and Password

Username	brukernavn
Password	••••••••

Login

Forgot your Password? Forgot your Username? Need to activate your account?

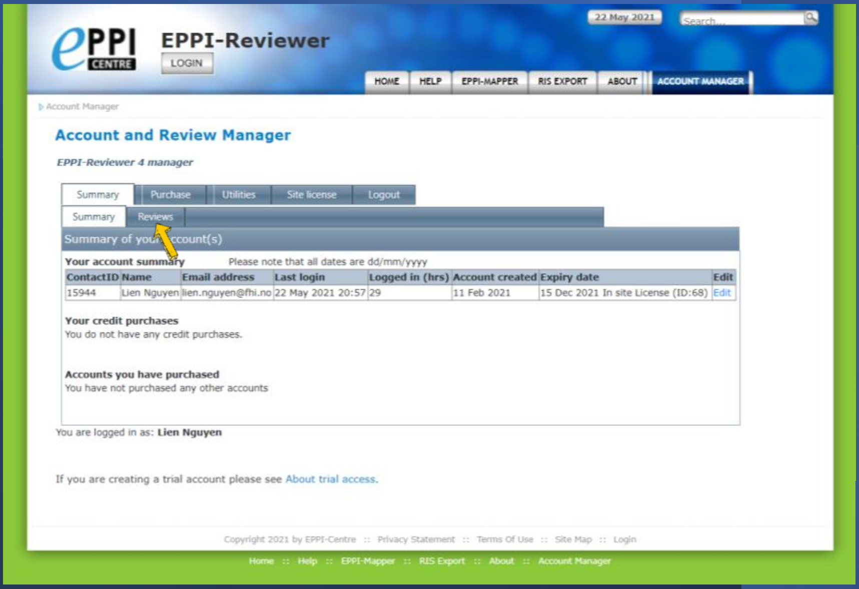
New account screen

If you are creating a trial account please see About trial access.

Copyright 2021 by EPPI-Centre :: Privacy Statement :: Terms Of Use :: Site Map :: Login

Brukerveiledning v. 1.0 ©2021

23



Account and Review Manager

EPPI-Reviewer 4 manager

Summary of your account(s)

Your account summary Please note that all dates are dd/mm/yyyy

ContactID	Name	Email address	Last login	Logged in (hrs)	Account created	Expiry date	Edit
15944	Lien Nguyen	lien.nguyen@fhi.no	22 May 2021 20:57	29	11 Feb 2021	15 Dec 2021 In site License (ID:68)	Edit

Your credit purchases
You do not have any credit purchases.

Accounts you have purchased
You have not purchased any other accounts

You are logged in as: **Lien Nguyen**

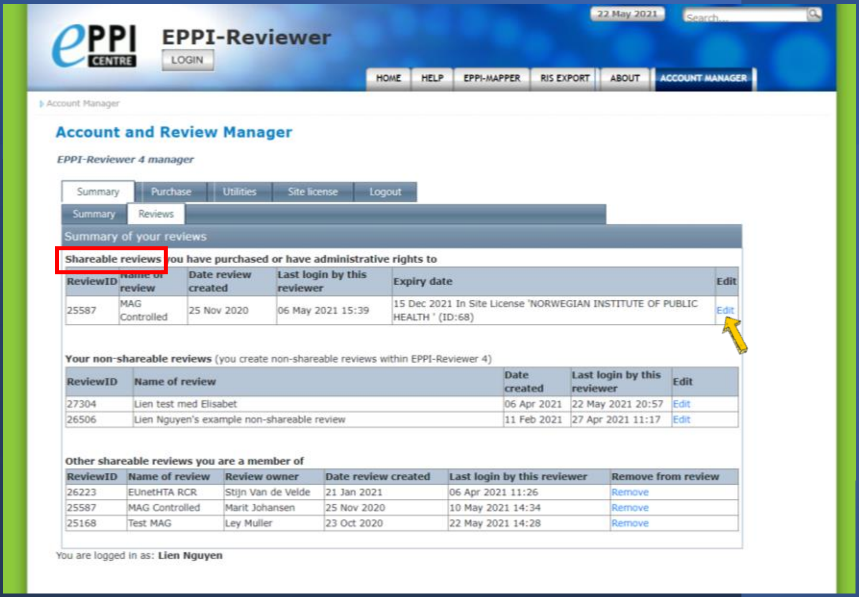
If you are creating a trial account please see About trial access.

Copyright 2021 by EPPI-Centre :: Privacy Statement :: Terms Of Use :: Site Map :: Login

Home :: Help :: EPPI-Mapper :: RIS Export :: About :: Account Manager

Brukerveiledning v. 1.0 ©2021

24



Account and Review Manager

EPPI-Reviewer 4 manager

Summary of your reviews

Shareable reviews you have purchased or have administrative rights to

ReviewID	Name of review	Date review created	Last login by this reviewer	Expiry date	Edit
25587	MAG Controlled	25 Nov 2020	06 May 2021 15:39	15 Dec 2021 In Site License 'NORWEGIAN INSTITUTE OF PUBLIC HEALTH ' (ID:68)	Edit

Your non-shareable reviews (you create non-shareable reviews within EPPI-Reviewer 4)

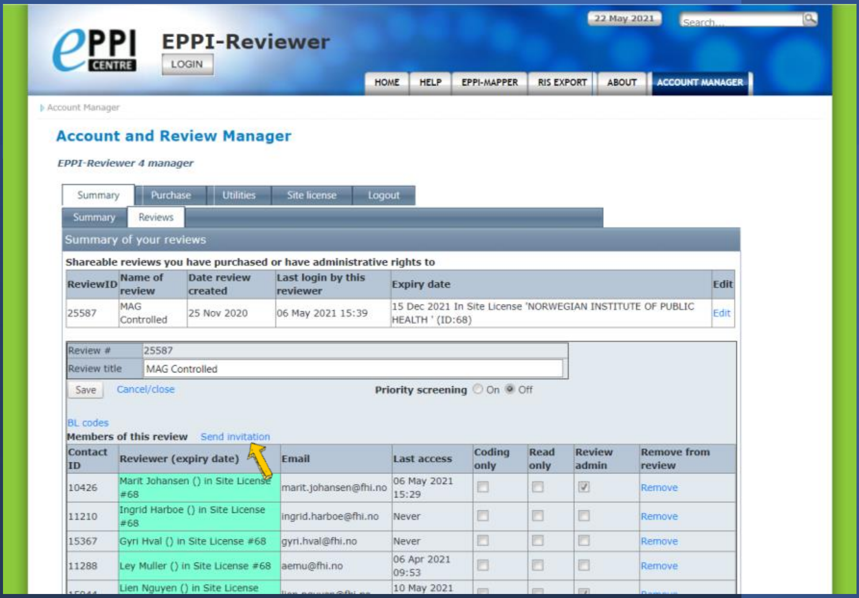
ReviewID	Name of review	Date created	Last login by this reviewer	Edit
27304	Lien test med Elisabet	06 Apr 2021	22 May 2021 20:57	Edit
26506	Lien Nguyen's example non-shareable review	11 Feb 2021	27 Apr 2021 11:17	Edit

Other shareable reviews you are a member of

ReviewID	Name of review	Review owner	Date review created	Last login by this reviewer	Remove from review
26223	EUnetHTA RCR	Stijn Van de Velde	21 Jan 2021	06 Apr 2021 11:26	Remove
25587	MAG Controlled	Marit Johansen	25 Nov 2020	10 May 2021 14:34	Remove
25168	Test MAG	Ley Muller	23 Oct 2020	22 May 2021 14:28	Remove

You are logged in as: **Lien Nguyen**

25



Account and Review Manager

EPPI-Reviewer 4 manager

Summary of your reviews

Shareable reviews you have purchased or have administrative rights to

ReviewID	Name of review	Date review created	Last login by this reviewer	Expiry date	Edit
25587	MAG Controlled	25 Nov 2020	06 May 2021 15:39	15 Dec 2021 In Site License 'NORWEGIAN INSTITUTE OF PUBLIC HEALTH ' (ID:68)	Edit

Review #	25587
Review title	MAG Controlled

Save Cancel/close

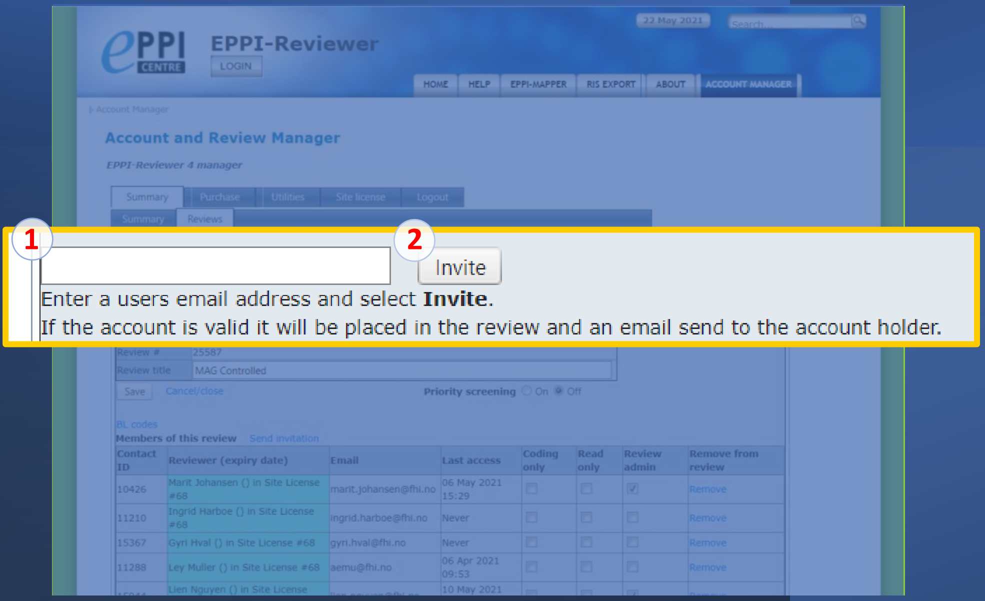
Priority screening ○ On ● Off

BL codes

Members of this review Send invitation

Contact ID	Reviewer (expiry date)	Email	Last access	Coding only	Read only	Review admin	Remove from review
10426	Marit Johansen () in Site License #68	marit.johansen@fhi.no	06 May 2021 15:29	□	□	☑	Remove
11210	Ingrid Harboe () in Site License #68	ingrid.harboe@fhi.no	Never	□	□	□	Remove
15367	Gyri Hval () in Site License #68	gyri.hval@fhi.no	Never	□	□	□	Remove
11288	Ley Muller () in Site License #68	aemu@fhi.no	06 Apr 2021 09:53	□	□	□	Remove
	Lien Nguyen () in Site License	lien.nguyen@fhi.no	10 May 2021	□	□	☑	Remove

26



EPPI-Reviewer

Account and Review Manager

EPPI-Reviewer 4 manager

1

2 Invite

Enter a users email address and select **Invite**.

If the account is valid it will be placed in the review and an email send to the account holder.

Review # 25587

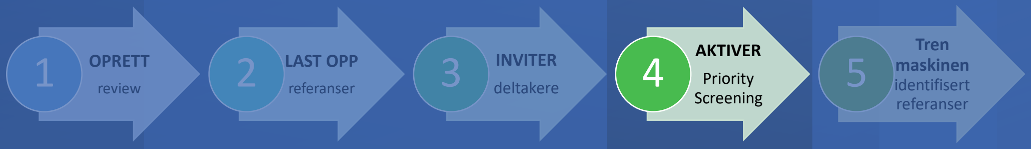
Review title MAG Controlled

Priority screening ○ On ● Off

Members of this review

Contact ID	Reviewer (expiry date)	Email	Last access	Coding only	Read only	Review admin	Remove from review
10426	Marit Johansen () in Site License #68	marit.johansen@fhi.no	06 May 2021 15:29	□	□	☑	Remove
11210	Ingrid Harboe () in Site License #68	ingrid.harboe@fhi.no	Never	□	□	□	Remove
15367	Gyri Hval () in Site License #68	gyri.hval@fhi.no	Never	□	□	□	Remove
11288	Ley Muller () in Site License #68	aemu@fhi.no	06 Apr 2021 09:53	□	□	□	Remove
[illegible]	Lien Nguyen () in Site License	[illegible]	10 May 2021				

27



1 OPPRETT review

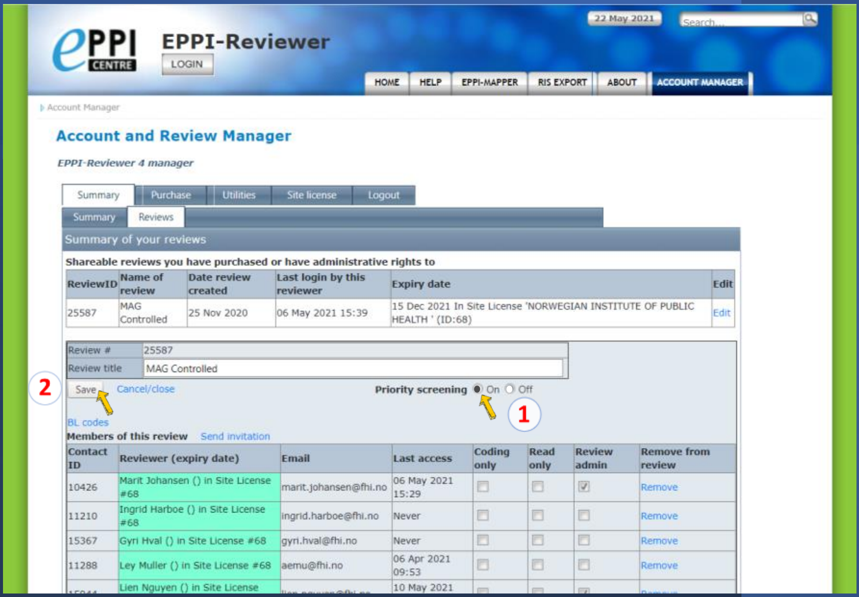
2 LAST OPP referanser

3 INVITER deltakere

4 AKTIVER Priority Screening

5 Tren maskinen identifisert referanser

28



EPPI-Reviewer

LOGIN

22 May 2021 | Search...

HOME | HELP | EPPI-MAPPER | RIS EXPORT | ABOUT | ACCOUNT MANAGER

Account Manager

Account and Review Manager

EPPI-Reviewer 4 manager

Summary | Purchase | Utilities | Site license | Logout

Summary | Reviews

Summary of your reviews

Shareable reviews you have purchased or have administrative rights to

ReviewID	Name of review	Date review created	Last login by this reviewer	Expiry date	Edit
25587	MAG Controlled	25 Nov 2020	06 May 2021 15:39	15 Dec 2021 In Site License 'NORWEGIAN INSTITUTE OF PUBLIC HEALTH ' (ID:68)	Edit

Review #	25587
Review title	MAG Controlled

2 Save | Cancel/close

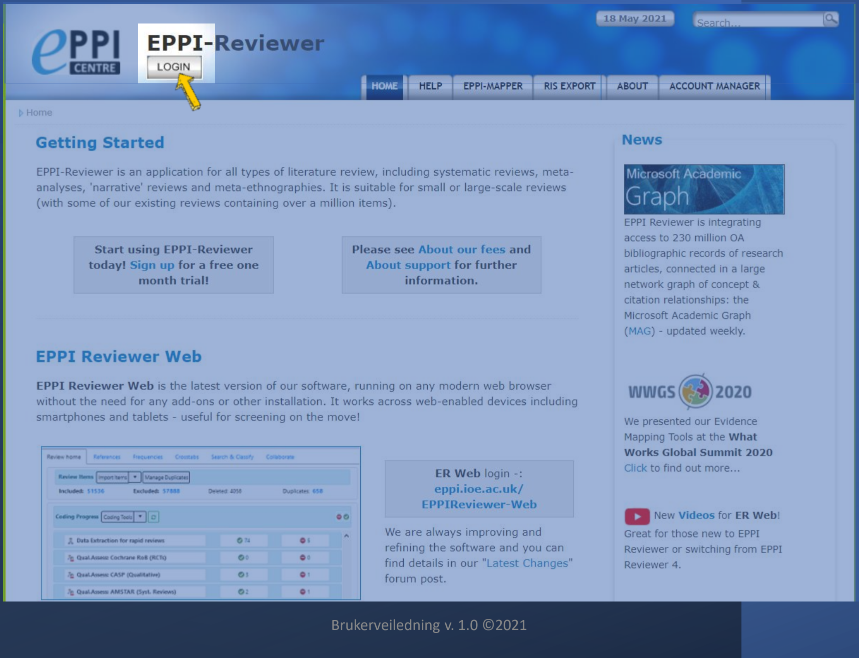
Priority screening ● On ○ Off 1

BL codes

Members of this review Send invitation

Contact ID	Reviewer (expiry date)	Email	Last access	Coding only	Read only	Review admin	Remove from review
10426	Marit Johansen () in Site License #68	marit.johansen@fhi.no	06 May 2021 15:29	□	□	☑	Remove
11210	Ingrid Harboe () in Site License #68	ingrid.harboe@fhi.no	Never	□	□	□	Remove
15367	Gyri Hval () in Site License #68	gyri.hval@fhi.no	Never	□	□	□	Remove
11288	Ley Muller () in Site License #68	aemu@fhi.no	06 Apr 2021 09:53	□	□	□	Remove
15044	Lien Nguyen () in Site License	lien.nguyen@fhi.no	10 May 2021	□	□	☑	Remove

29



EPPI-Reviewer

LOGIN

18 May 2021 | Search...

HOME | HELP | EPPI-MAPPER | RIS EXPORT | ABOUT | ACCOUNT MANAGER

Home

Getting Started

EPPI-Reviewer is an application for all types of literature review, including systematic reviews, meta-analyses, 'narrative' reviews and meta-ethnographies. It is suitable for small or large-scale reviews (with some of our existing reviews containing over a million items).

Start using EPPI-Reviewer today! Sign up for a free one month trial!

Please see About our fees and About support for further information.

EPPI Reviewer Web

EPPI Reviewer Web is the latest version of our software, running on any modern web browser without the need for any add-ons or other installation. It works across web-enabled devices including smartphones and tablets - useful for screening on the move!

ER Web login -: **eppi.ioe.ac.uk/EPPIReviewer-Web**

We are always improving and refining the software and you can find details in our "Latest Changes" forum post.

News

Microsoft Academic Graph

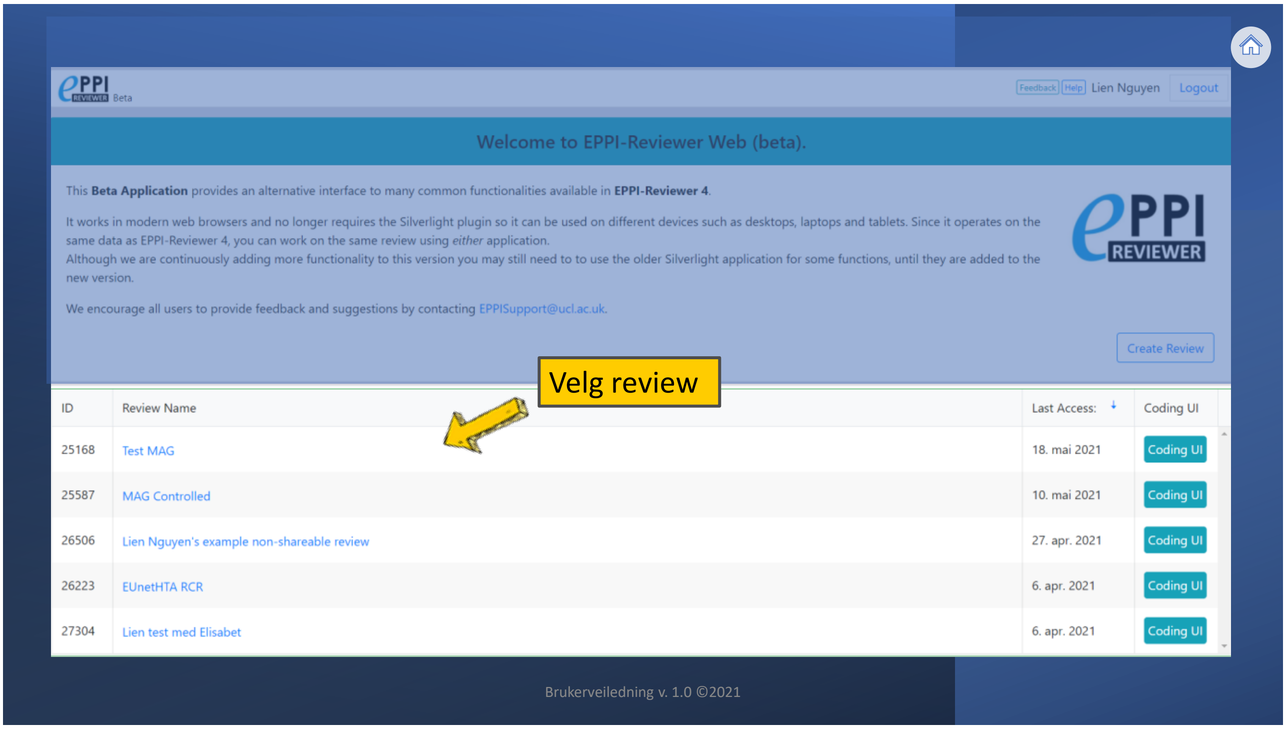
EPPI Reviewer is integrating access to 230 million OA bibliographic records of research articles, connected in a large network graph of concept & citation relationships: the Microsoft Academic Graph (MAG) - updated weekly.

WWGS 2020

We presented our Evidence Mapping Tools at the **What Works Global Summit 2020** Click to find out more...

New **Videos** for **ER Web**! Great for those new to EPPI Reviewer or switching from EPPI Reviewer 4.

30



Welcome to EPPI-Reviewer Web (beta).

This **Beta Application** provides an alternative interface to many common functionalities available in **EPPI-Reviewer 4**.

It works in modern web browsers and no longer requires the Silverlight plugin so it can be used on different devices such as desktops, laptops and tablets. Since it operates on the same data as EPPI-Reviewer 4, you can work on the same review using *either* application.
Although we are continuously adding more functionality to this version you may still need to to use the older Silverlight application for some functions, until they are added to the new version.

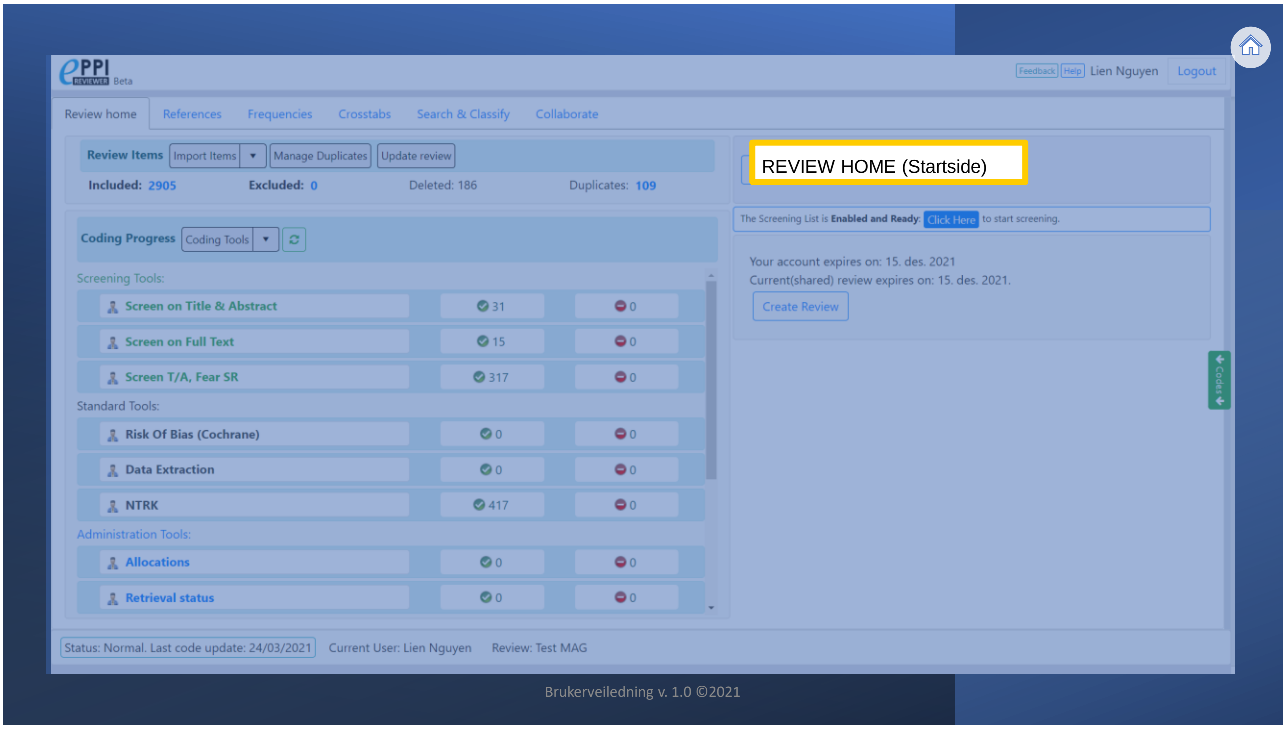
We encourage all users to provide feedback and suggestions by contacting EPPISupport@ucl.ac.uk.

Velg review

ID	Review Name	Last Access:	Coding UI
25168	Test MAG	18. mai 2021	Coding UI
25587	MAG Controlled	10. mai 2021	Coding UI
26506	Lien Nguyen's example non-shareable review	27. apr. 2021	Coding UI
26223	EUnetHTA RCR	6. apr. 2021	Coding UI
27304	Lien test med Elisabet	6. apr. 2021	Coding UI

Brukerveiledning v. 1.0 ©2021

31



REVIEW HOME (Startside)

Included: 2905 Excluded: 0 Deleted: 186 Duplicates: 109

The Screening List is **Enabled and Ready**: Click Here to start screening.

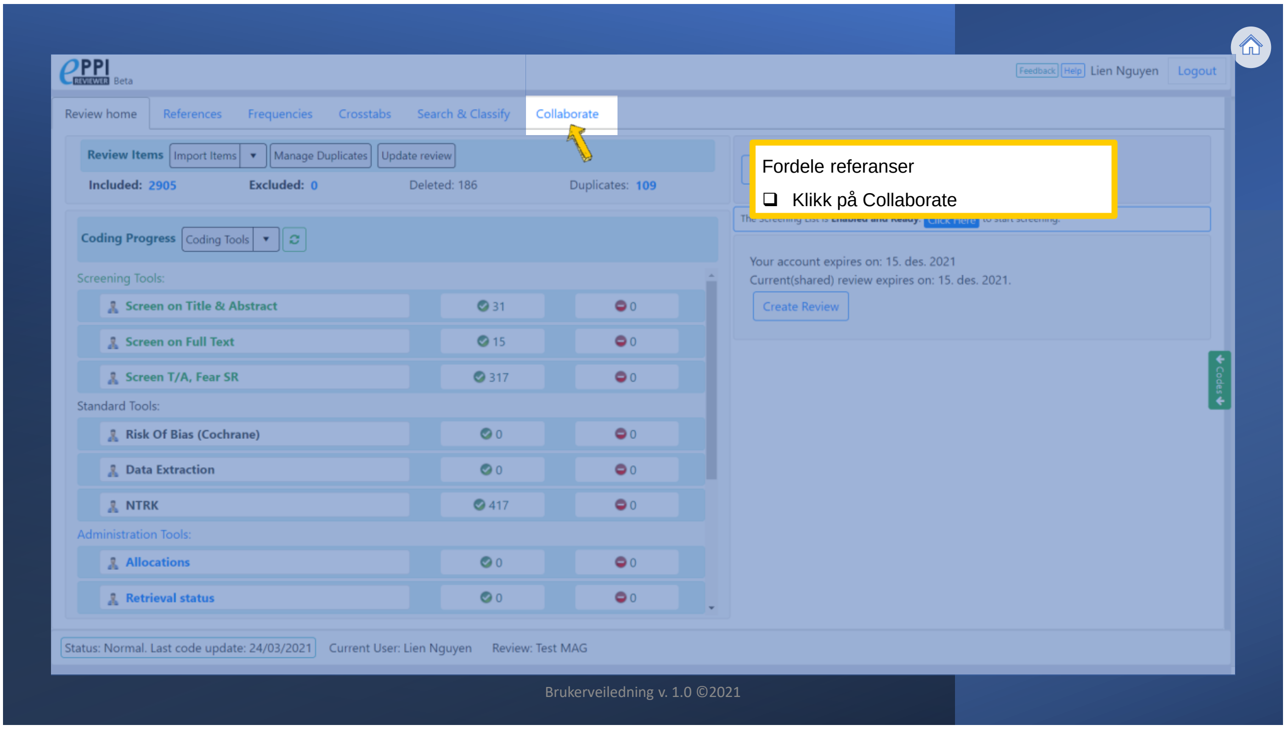
Your account expires on: 15. des. 2021
Current(shared) review expires on: 15. des. 2021.

Screening Tools:		
Screen on Title & Abstract	✓ 31	⊖ 0
Screen on Full Text	✓ 15	⊖ 0
Screen T/A, Fear SR	✓ 317	⊖ 0
Standard Tools:		
Risk Of Bias (Cochrane)	✓ 0	⊖ 0
Data Extraction	✓ 0	⊖ 0
NTRK	✓ 417	⊖ 0
Administration Tools:		
Allocations	✓ 0	⊖ 0
Retrieval status	✓ 0	⊖ 0

Status: Normal. Last code update: 24/03/2021 Current User: Lien Nguyen Review: Test MAG

Brukerveiledning v. 1.0 ©2021

32



Fordele referanser

❑ Klikk på Collaborate

Included: 2905 Excluded: 0 Deleted: 186 Duplicates: 109

Screening Tools:		
Screen on Title & Abstract	✓ 31	⊖ 0
Screen on Full Text	✓ 15	⊖ 0
Screen T/A, Fear SR	✓ 317	⊖ 0
Standard Tools:		
Risk Of Bias (Cochrane)	✓ 0	⊖ 0
Data Extraction	✓ 0	⊖ 0
NTRK	✓ 417	⊖ 0
Administration Tools:		
Allocations	✓ 0	⊖ 0
Retrieval status	✓ 0	⊖ 0

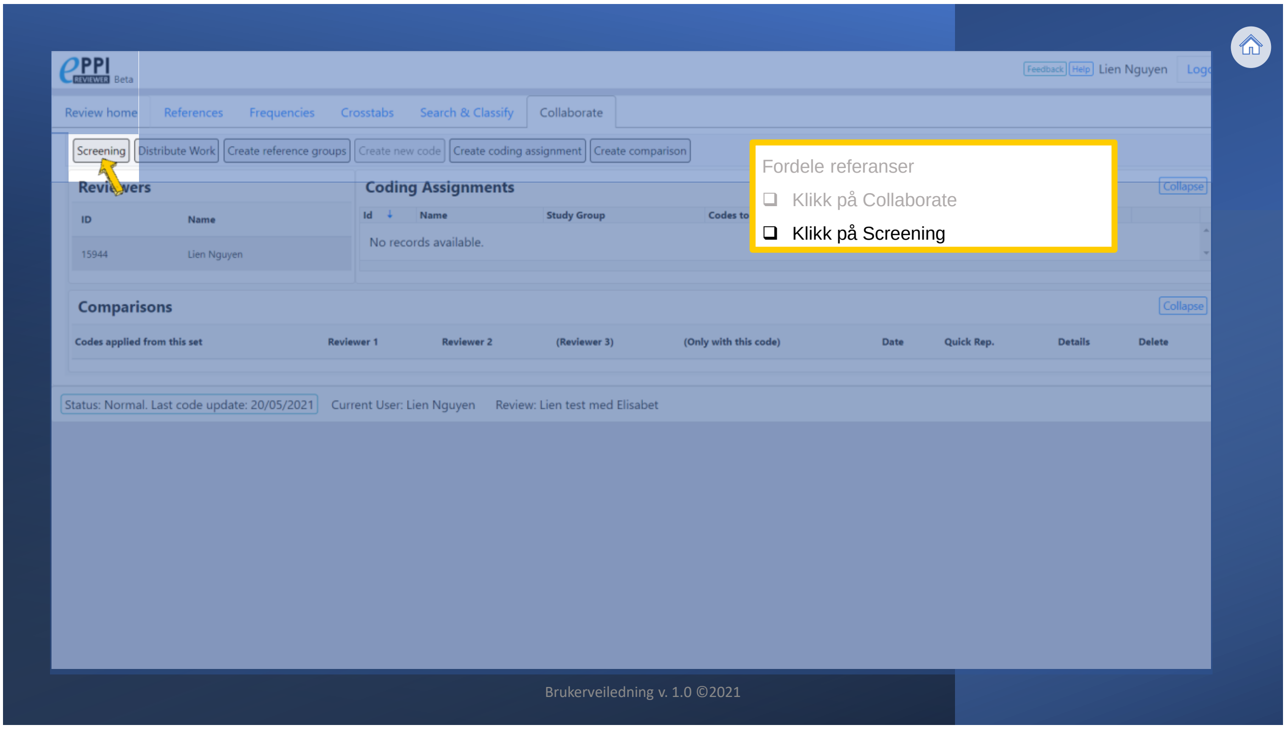
The screening list is **Enabled and Ready**. Click Here to start screening.

Your account expires on: 15. des. 2021
Current(shared) review expires on: 15. des. 2021.

Status: Normal. Last code update: 24/03/2021 Current User: Lien Nguyen Review: Test MAG

Brukerveiledning v. 1.0 ©2021

33



Fordele referanser

❑ Klikk på Collaborate

❑ Klikk på Screening

Reviewers

ID	Name
15944	Lien Nguyen

Coding Assignments

Id	Name	Study Group	Codes to

No records available.

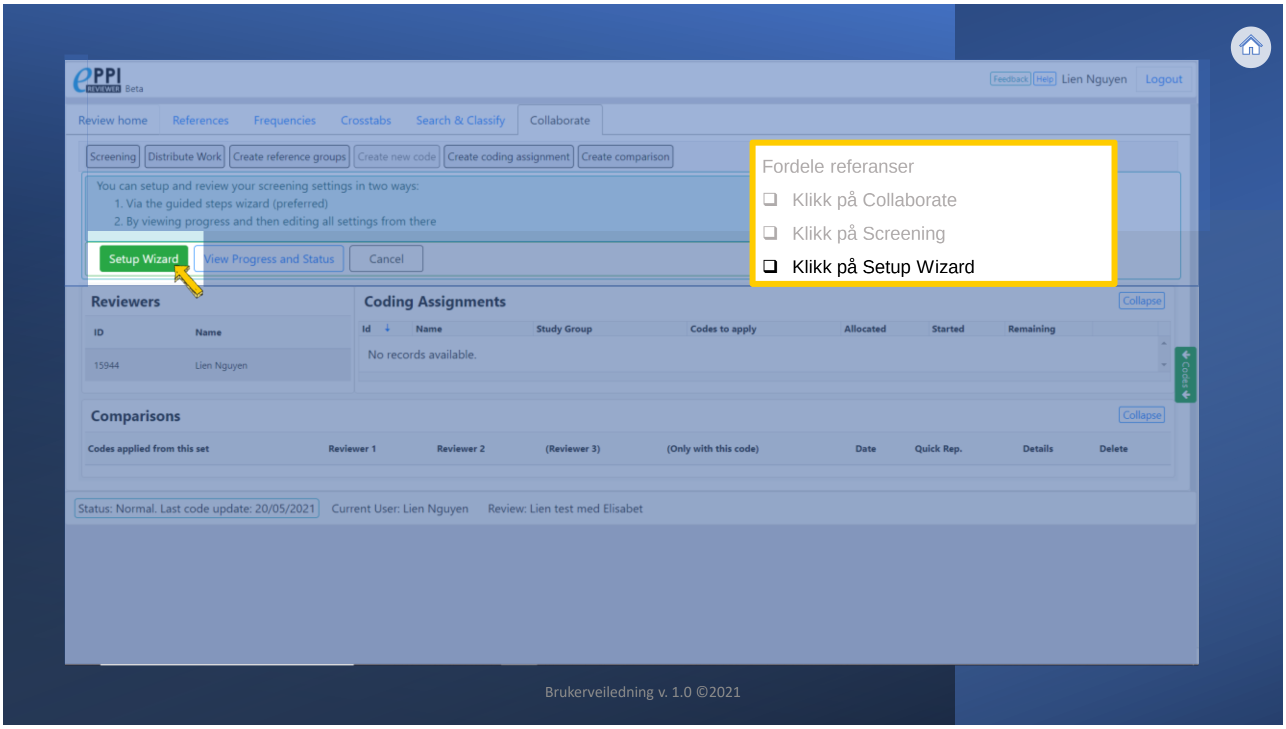
Comparisons

Codes applied from this set	Reviewer 1	Reviewer 2	(Reviewer 3)	(Only with this code)	Date	Quick Rep.	Details	Delete

Status: Normal. Last code update: 20/05/2021 Current User: Lien Nguyen Review: Lien test med Elisabet

Brukerveiledning v. 1.0 ©2021

34



You can setup and review your screening settings in two ways:
1. Via the guided steps wizard (preferred)
2. By viewing progress and then editing all settings from there

Reviewers

ID	Name
15944	Lien Nguyen

Coding Assignments

Id	Name	Study Group	Codes to apply	Allocated	Started	Remaining
No records available.						

Comparisons

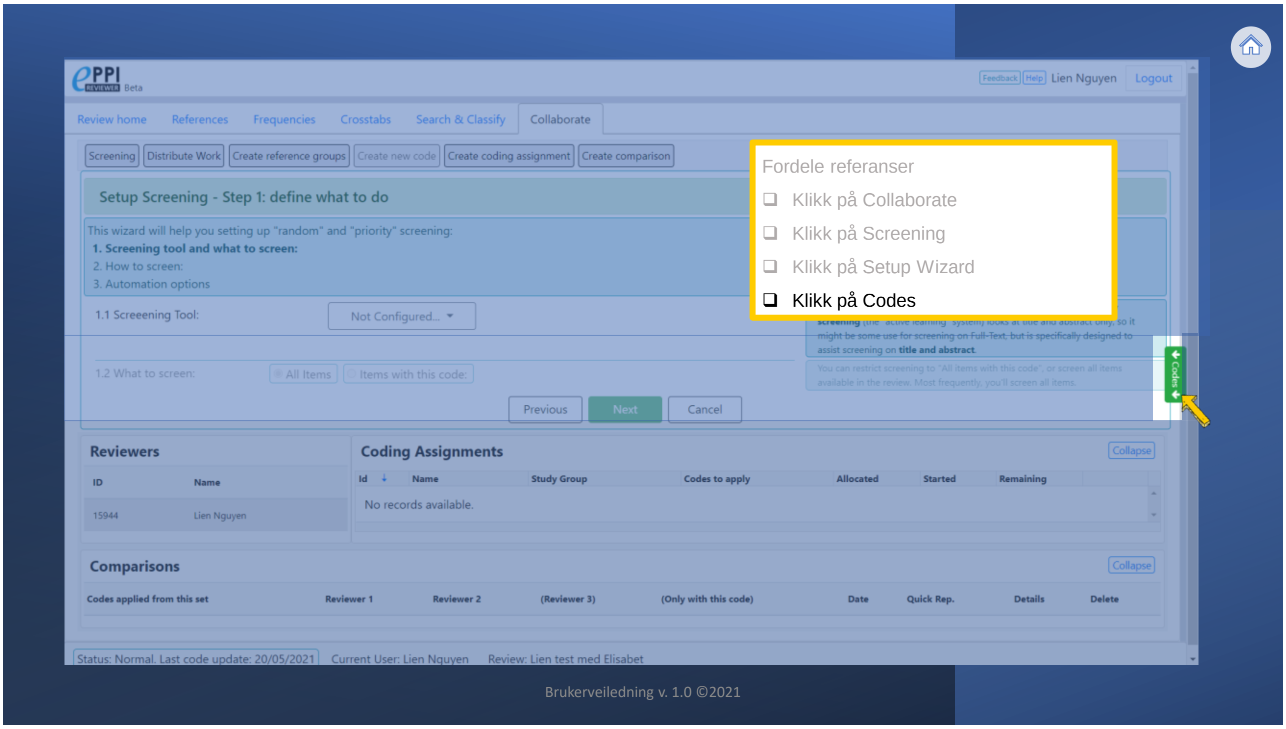
Codes applied from this set	Reviewer 1	Reviewer 2	(Reviewer 3)	(Only with this code)	Date	Quick Rep.	Details	Delete

Status: Normal. Last code update: 20/05/2021 Current User: Lien Nguyen Review: Lien test med Elisabet

Fordele referanser
- ❑ Klikk på Collaborate
- ❑ Klikk på Screening
- ❑ Klikk på Setup Wizard

Brukerveiledning v. 1.0 ©2021

35



Setup Screening - Step 1: define what to do

This wizard will help you setting up "random" and "priority" screening:
1. Screening tool and what to screen:
2. How to screen:
3. Automation options

1.1 Screening Tool: Not Configured...

1.2 What to screen: All Items Items with this code:

screening (the "active learning" system) looks at title and abstract only, so it might be some use for screening on Full-Text, but is specifically designed to assist screening on **title and abstract**.

You can restrict screening to "All items with this code", or screen all items available in the review. Most frequently, you'll screen all items.

Reviewers

ID	Name
15944	Lien Nguyen

Coding Assignments

Id	Name	Study Group	Codes to apply	Allocated	Started	Remaining
No records available.						

Comparisons

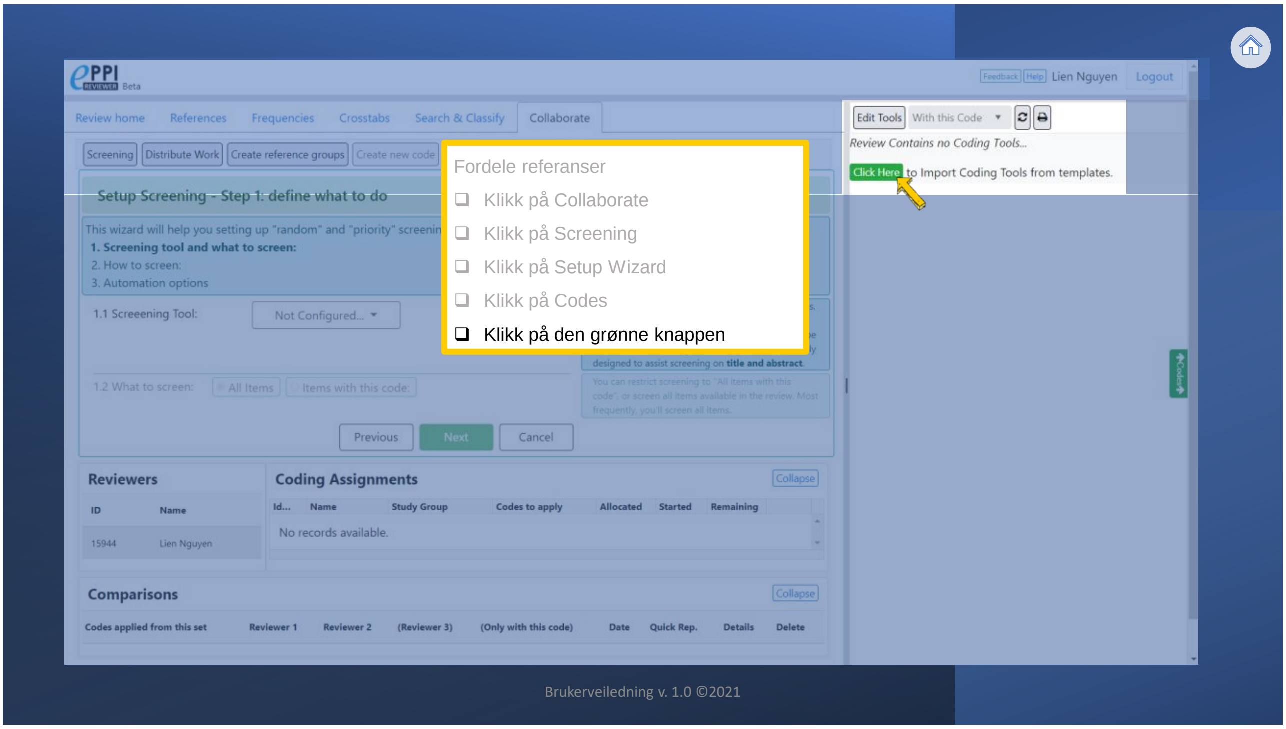
Codes applied from this set	Reviewer 1	Reviewer 2	(Reviewer 3)	(Only with this code)	Date	Quick Rep.	Details	Delete

Status: Normal. Last code update: 20/05/2021 Current User: Lien Nguyen Review: Lien test med Elisabet

Fordele referanser
- ❑ Klikk på Collaborate
- ❑ Klikk på Screening
- ❑ Klikk på Setup Wizard
- ❑ Klikk på Codes

Brukerveiledning v. 1.0 ©2021

36



Review Contains no Coding Tools...

Click Here to Import Coding Tools from templates.

Setup Screening - Step 1: define what to do

This wizard will help you setting up "random" and "priority" screenin

1. Screening tool and what to screen:
2. How to screen:
3. Automation options

1.1 Screeening Tool: Not Configured...

1.2 What to screen: ○ All Items ○ Items with this code:

designed to assist screening on **title and abstract**.

You can restrict screening to "All items with this code", or screen all items available in the review. Most frequently, you'll screen all items.

Fordele referanser

- ❑ Klikk på Collaborate
- ❑ Klikk på Screening
- ❑ Klikk på Setup Wizard
- ❑ Klikk på Codes
- ❑ Klikk på den grønne knappen

Reviewers

ID	Name
15944	Lien Nguyen

Coding Assignments

Id...	Name	Study Group	Codes to apply	Allocated	Started	Remaining

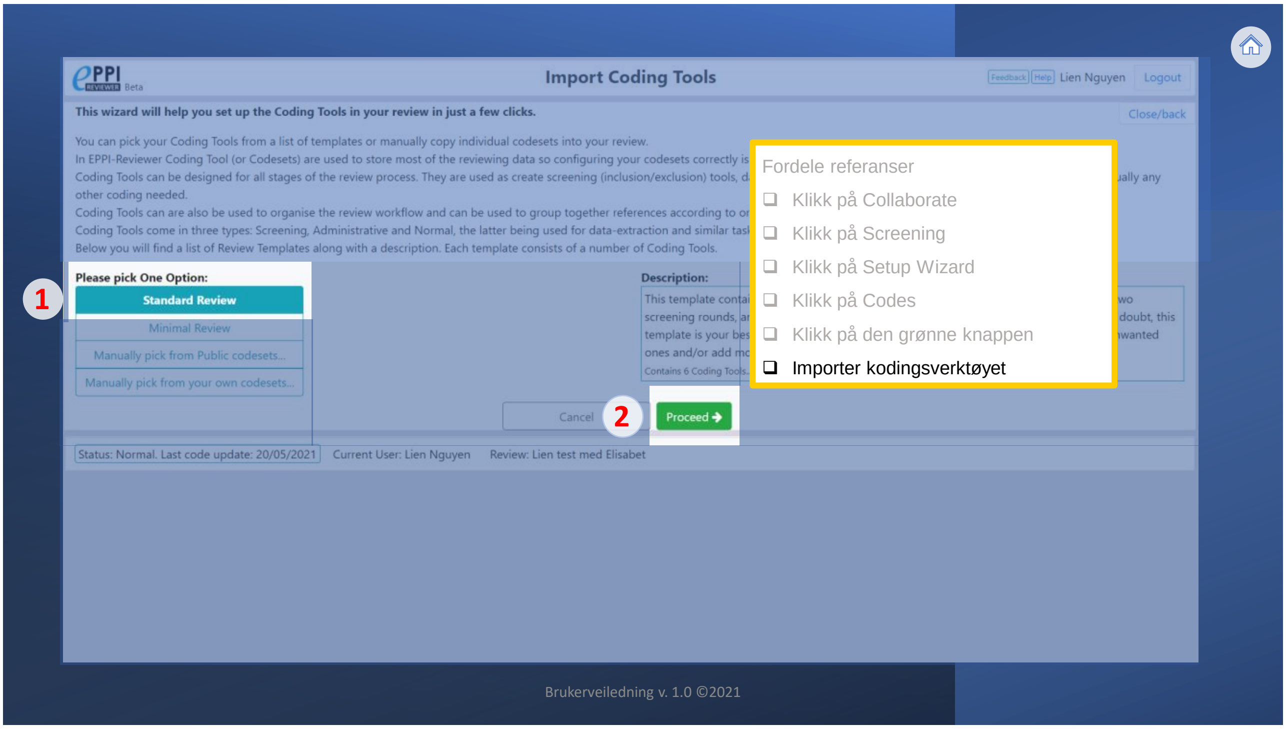
No records available.

Comparisons

Codes applied from this set	Reviewer 1	Reviewer 2	(Reviewer 3)	(Only with this code)	Date	Quick Rep.	Details	Delete

Brukerveiledning v. 1.0 ©2021

37



Import Coding Tools

This wizard will help you set up the Coding Tools in your review in just a few clicks.

You can pick your Coding Tools from a list of templates or manually copy individual codesets into your review.
In EPPI-Reviewer Coding Tool (or Codesets) are used to store most of the reviewing data so configuring your codesets correctly is
Coding Tools can be designed for all stages of the review process. They are used as create screening (inclusion/exclusion) tools, d ually any other coding needed.
Coding Tools can are also be used to organise the review workflow and can be used to group together references according to or
Coding Tools come in three types: Screening, Administrative and Normal, the latter being used for data-extraction and similar task
Below you will find a list of Review Templates along with a description. Each template consists of a number of Coding Tools.

Please pick One Option:

- Standard Review
- Minimal Review
- Manually pick from Public codesets...
- Manually pick from your own codesets...

Description:

This template contai ... wo screening rounds, ar ... doubt, this template is your bes ... wanted ones and/or add mo

Contains 6 Coding Tools...

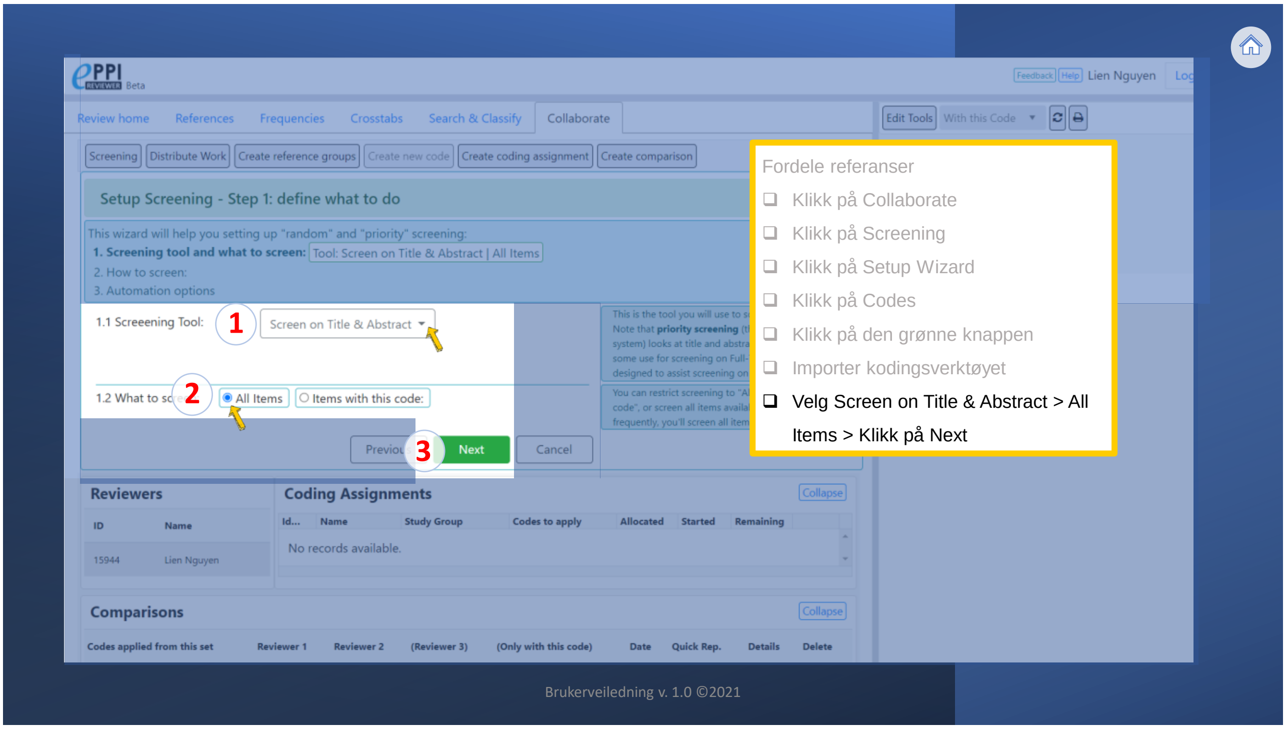
Fordele referanser

- ❑ Klikk på Collaborate
- ❑ Klikk på Screening
- ❑ Klikk på Setup Wizard
- ❑ Klikk på Codes
- ❑ Klikk på den grønne knappen
- ❑ Importer kodingsverktøyet

Status: Normal. Last code update: 20/05/2021 Current User: Lien Nguyen Review: Lien test med Elisabet

Brukerveiledning v. 1.0 ©2021

38



Setup Screening - Step 1: define what to do

This wizard will help you setting up "random" and "priority" screening:

1. Screening tool and what to screen: Tool: Screen on Title & Abstract | All Items

2. How to screen:

3. Automation options

1.1 Screeening Tool: Screen on Title & Abstract

1.2 What to sc[reen:] ● All Items ○ Items with this code:

This is the tool you will use to s… Note that **priority screening** (t… system) looks at title and abstra… some use for screening on Full-… designed to assist screening on…

You can restrict screening to "Al… code", or screen all items availa… frequently, you'll screen all item…

Reviewers

ID	Name
15944	Lien Nguyen

Coding Assignments

Id...	Name	Study Group	Codes to apply	Allocated	Started	Remaining
No records available.						

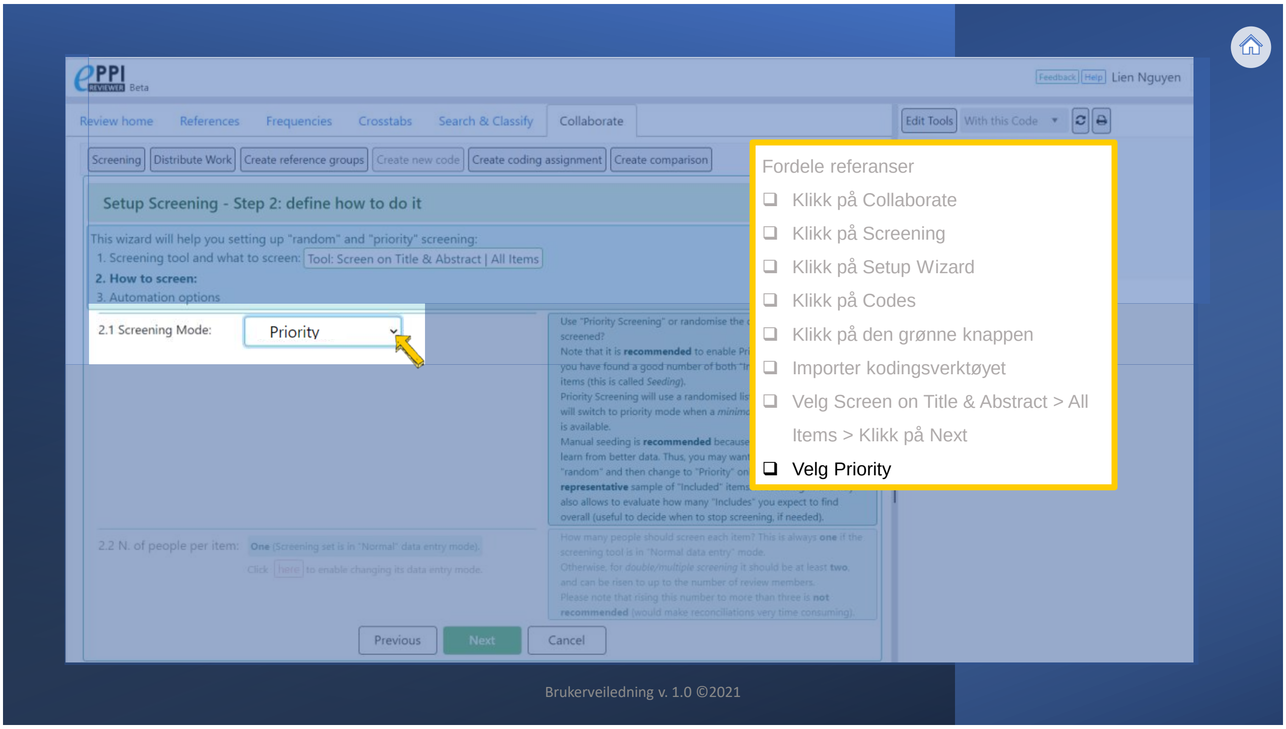
Comparisons

Codes applied from this set	Reviewer 1	Reviewer 2	(Reviewer 3)	(Only with this code)	Date	Quick Rep.	Details	Delete

Fordele referanser
- ❑ Klikk på Collaborate
- ❑ Klikk på Screening
- ❑ Klikk på Setup Wizard
- ❑ Klikk på Codes
- ❑ Klikk på den grønne knappen
- ❑ Importer kodingsverktøyet
- ❑ Velg Screen on Title & Abstract > All Items > Klikk på Next

Brukerveiledning v. 1.0 ©2021

39



Setup Screening - Step 2: define how to do it

This wizard will help you setting up "random" and "priority" screening:

1. Screening tool and what to screen: Tool: Screen on Title & Abstract | All Items

2. How to screen:

3. Automation options

2.1 Screening Mode: Priority

Use "Priority Screening" or randomise the o… screened?
Note that it is **recommended** to enable Pri… you have found a good number of both "In… items (this is called *Seeding*).
Priority Screening will use a randomised lis… will switch to priority mode when a *minimo*… is available.
Manual seeding is **recommended** because… learn from better data. Thus, you may want… "random" and then change to "Priority" on… **representative** sample of "Included" items… also allows to evaluate how many "Includes" you expect to find overall (useful to decide when to stop screening, if needed).

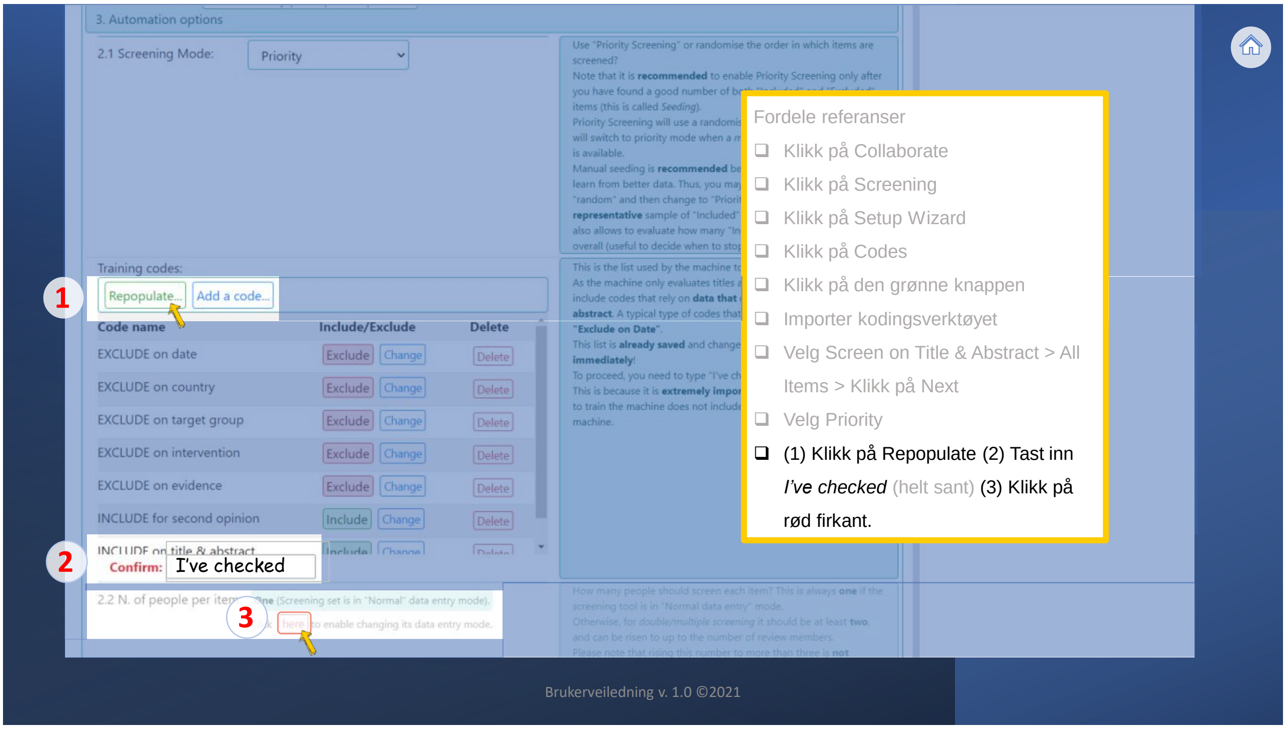
2.2 N. of people per item: **One** (Screening set is in "Normal" data entry mode).
Click here to enable changing its data entry mode.

How many people should screen each item? This is always **one** if the screening tool is in "Normal data entry" mode.
Otherwise, for *double/multiple screening* it should be at least **two**, and can be risen to up to the number of review members.
Please note that rising this number to more than three is **not recommended** (would make reconciliations very time consuming).

Fordele referanser
- ❑ Klikk på Collaborate
- ❑ Klikk på Screening
- ❑ Klikk på Setup Wizard
- ❑ Klikk på Codes
- ❑ Klikk på den grønne knappen
- ❑ Importer kodingsverktøyet
- ❑ Velg Screen on Title & Abstract > All Items > Klikk på Next
- ❑ Velg Priority

Brukerveiledning v. 1.0 ©2021

40



3. Automation options

2.1 Screening Mode: Priority

Use "Priority Screening" or randomise the order in which items are screened?
Note that it is **recommended** to enable Priority Screening only after you have found a good number of both "Included" and "Excluded" items (this is called *Seeding*).
Priority Screening will use a randomis… will switch to priority mode when a m… is available.
Manual seeding is **recommended** be… learn from better data. Thus, you may… "random" and then change to "Priorit… **representative** sample of "Included"… also allows to evaluate how many "In… overall (useful to decide when to stop…

Training codes:

1 Repopulate... Add a code...

Code name	Include/Exclude		Delete
EXCLUDE on date	Exclude	Change	Delete
EXCLUDE on country	Exclude	Change	Delete
EXCLUDE on target group	Exclude	Change	Delete
EXCLUDE on intervention	Exclude	Change	Delete
EXCLUDE on evidence	Exclude	Change	Delete
INCLUDE for second opinion	Include	Change	Delete
INCLUDE on title & abstract	Include	Change	Delete

This is the list used by the machine t… As the machine only evaluates titles a… include codes that rely on **data that**… **abstract**. A typical type of codes that… **"Exclude on Date"**.
This list is **already saved** and change… **immediately**!
To proceed, you need to type "I've ch… This is because it is **extremely impor**… to train the machine does not include… machine.

2 **Confirm:** I've checked

2.2 N. of people per item: One (Screening set is in "Normal" data entry mode).

3 Click here to enable changing its data entry mode.

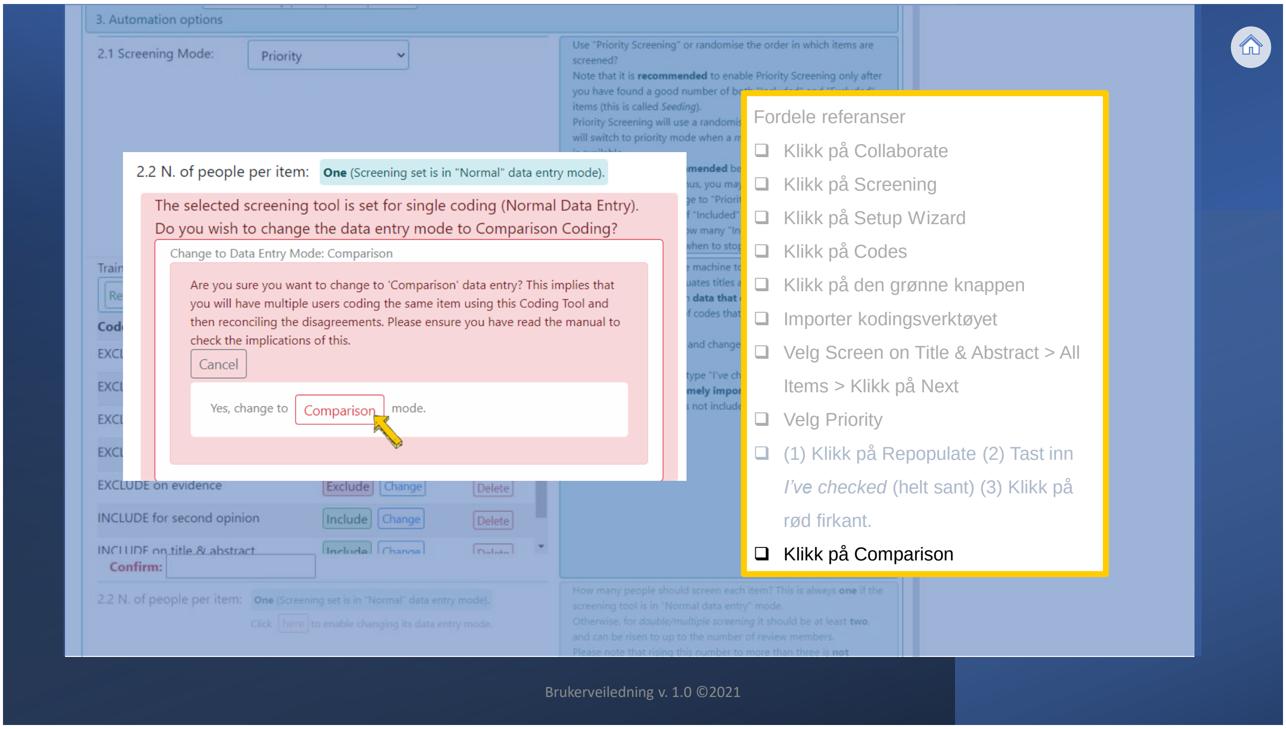
How many people should screen each item? This is always **one** if the screening tool is in "Normal data entry" mode.
Otherwise, for *double/multiple screening* it should be at least **two**, and can be risen to up to the number of review members.
Please note that rising this number to more than three is **not**

Fordele referanser

- ❑ Klikk på Collaborate
- ❑ Klikk på Screening
- ❑ Klikk på Setup Wizard
- ❑ Klikk på Codes
- ❑ Klikk på den grønne knappen
- ❑ Importer kodingsverktøyet
- ❑ Velg Screen on Title & Abstract > All Items > Klikk på Next
- ❑ Velg Priority
- ❑ (1) Klikk på Repopulate (2) Tast inn *I've checked* (helt sant) (3) Klikk på rød firkant.

Brukerveiledning v. 1.0 ©2021

41



3. Automation options

2.1 Screening Mode: Priority

2.2 N. of people per item: **One** (Screening set is in "Normal" data entry mode).

The selected screening tool is set for single coding (Normal Data Entry).
Do you wish to change the data entry mode to Comparison Coding?

Change to Data Entry Mode: Comparison

Are you sure you want to change to 'Comparison' data entry? This implies that you will have multiple users coding the same item using this Coding Tool and then reconciling the disagreements. Please ensure you have read the manual to check the implications of this.

Cancel

Yes, change to Comparison mode.

EXCLUDE on evidence Exclude Change Delete
INCLUDE for second opinion Include Change Delete
INCLUDE on title & abstract Include Change Delete
Confirm:

2.2 N. of people per item: One (Screening set is in "Normal" data entry mode).
Click here to enable changing its data entry mode.

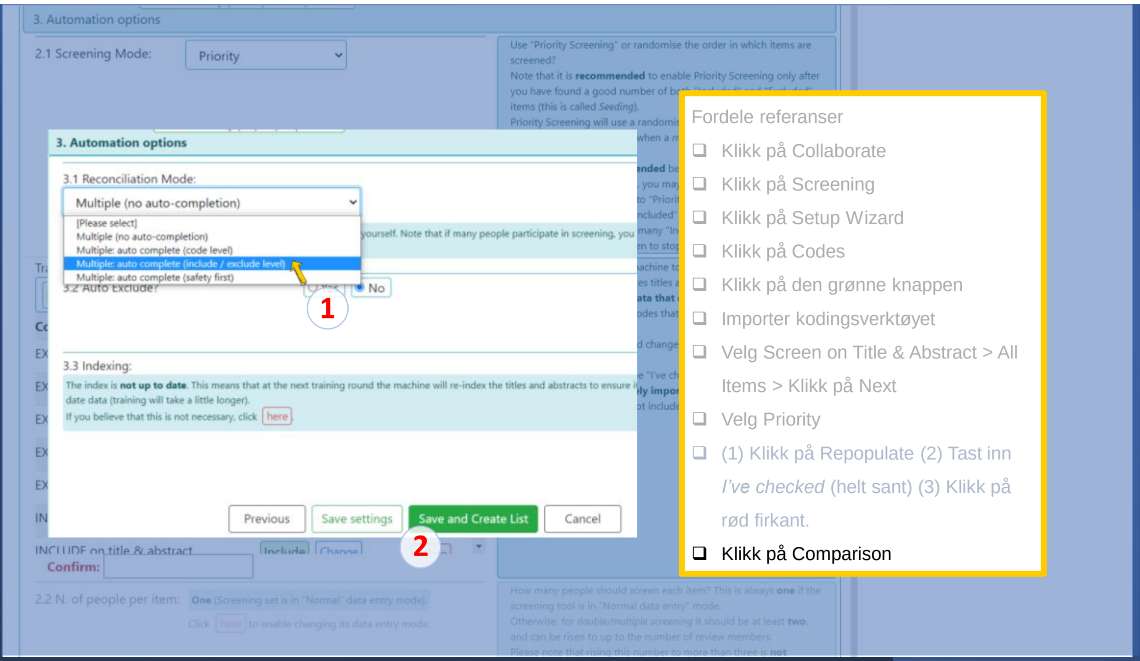
How many people should screen each item? This is always **one** if the screening tool is in "Normal data entry" mode.
Otherwise, for *double/multiple screening* it should be at least **two**, and can be risen to up to the number of review members.
Please note that rising this number to more than three is **not**

Fordele referanser

- ❑ Klikk på Collaborate
- ❑ Klikk på Screening
- ❑ Klikk på Setup Wizard
- ❑ Klikk på Codes
- ❑ Klikk på den grønne knappen
- ❑ Importer kodingsverktøyet
- ❑ Velg Screen on Title & Abstract > All Items > Klikk på Next
- ❑ Velg Priority
- ❑ (1) Klikk på Repopulate (2) Tast inn *I've checked* (helt sant) (3) Klikk på rød firkant.
- ❑ Klikk på Comparison

Brukerveiledning v. 1.0 ©2021

42



3. Automation options

2.1 Screening Mode: Priority

Use "Priority Screening" or randomise the order in which items are screened?
Note that it is **recommended** to enable Priority Screening only after you have found a good number of bo... items (this is called *Seeding*).
Priority Screening will use a randomis...

3. Automation options

3.1 Reconciliation Mode:

Multiple (no auto-completion)

- [Please select]
- Multiple (no auto-completion)
- Multiple: auto complete (code level)
- Multiple: auto complete (include / exclude level)
- Multiple: auto complete (safety first)

...yourself. Note that if many people participate in screening, you

3.2 Auto Exclude? ○ Yes ● No

3.3 Indexing:

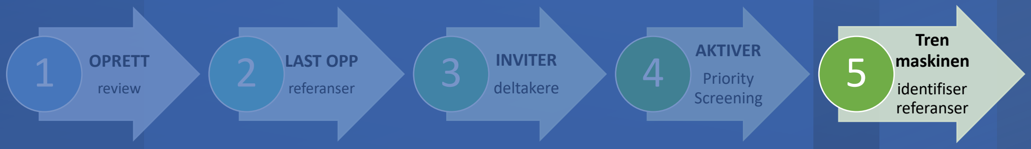
The index is **not up to date**. This means that at the next training round the machine will re-index the titles and abstracts to ensure i... date data (training will take a little longer).
If you believe that this is not necessary, click here.

2.2 N. of people per item: **One** (Screening set is in "Normal" data entry mode).
Click here to enable changing its data entry mode.

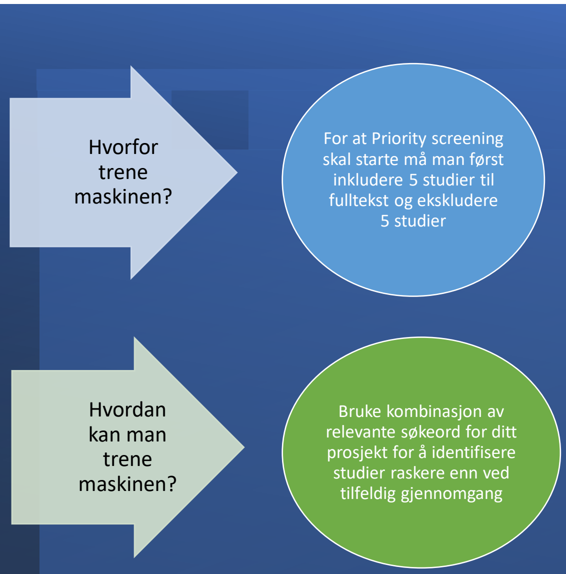
How many people should screen each item? This is always **one** if the screening tool is in "Normal data entry" mode.
Otherwise, for *double/multiple screening* it should be at least **two**, and can be risen to up to the number of review members.
Please note that rising this number to more than three is not

Fordele referanser

- ❑ Klikk på Collaborate
- ❑ Klikk på Screening
- ❑ Klikk på Setup Wizard
- ❑ Klikk på Codes
- ❑ Klikk på den grønne knappen
- ❑ Importer kodingsverktøyet
- ❑ Velg Screen on Title & Abstract > All Items > Klikk på Next
- ❑ Velg Priority
- ❑ (1) Klikk på Repopulate (2) Tast inn *I've checked* (helt sant) (3) Klikk på rød firkant.
- ❑ Klikk på Comparison



1 OPPRETT review
2 LAST OPP referanser
3 INVITER deltakere
4 AKTIVER Priority Screening
5 Tren maskinen identifiser referanser



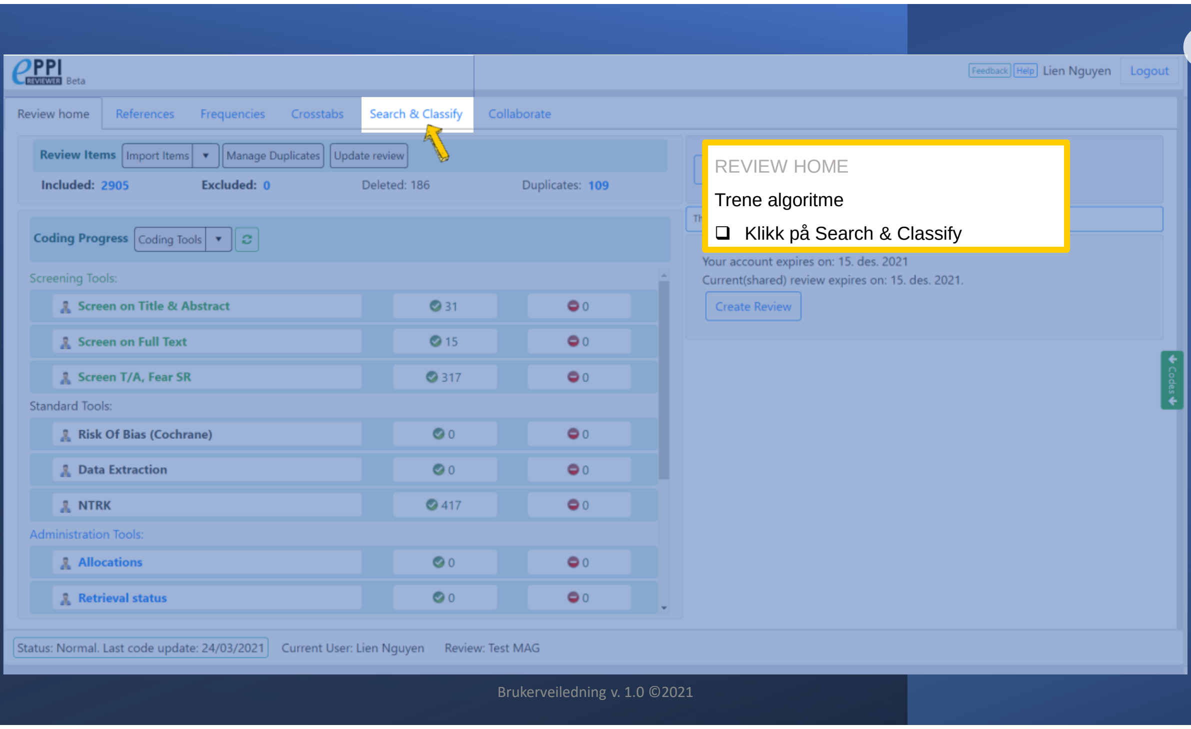
Hvorfor trene maskinen?

For at Priority screening skal starte må man først inkludere 5 studier til fulltekst og ekskludere 5 studier

Hvordan kan man trene maskinen?

Bruke kombinasjon av relevante søkeord for ditt prosjekt for å identifisere studier raskere enn ved tilfeldig gjennomgang

45



Included: 2905 Excluded: 0 Deleted: 186 Duplicates: 109

Screening Tools:

Screen on Title & Abstract	31	0
Screen on Full Text	15	0
Screen T/A, Fear SR	317	0

Standard Tools:

Risk Of Bias (Cochrane)	0	0
Data Extraction	0	0
NTRK	417	0

Administration Tools:

Allocations	0	0
Retrieval status	0	0

Your account expires on: 15. des. 2021

Current(shared) review expires on: 15. des. 2021.

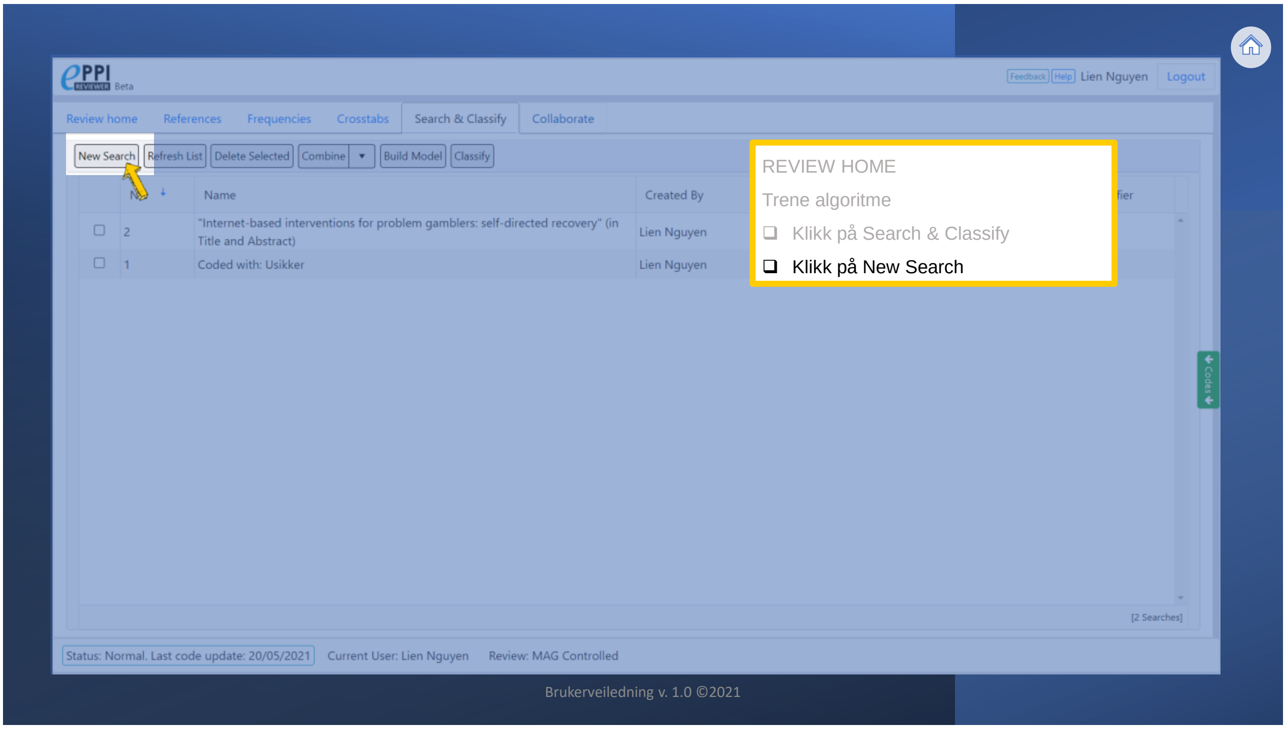
Status: Normal. Last code update: 24/03/2021 Current User: Lien Nguyen Review: Test MAG

REVIEW HOME

Trene algoritme

❑ Klikk på Search & Classify

46



REVIEW HOME

Trene algoritme

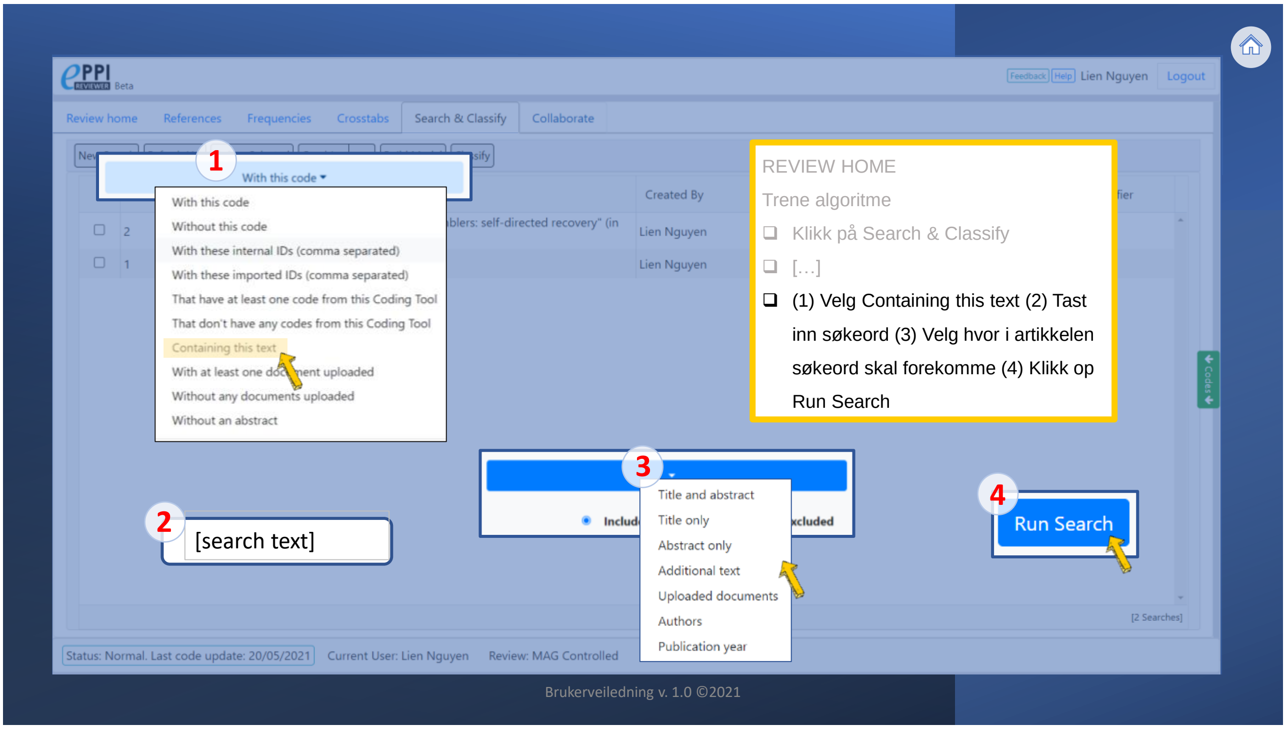
❑ Klikk på Search & Classify

❑ Klikk på New Search

	No.	Name	Created By
□	2	"Internet-based interventions for problem gamblers: self-directed recovery" (in Title and Abstract)	Lien Nguyen
□	1	Coded with: Usikker	Lien Nguyen

Status: Normal. Last code update: 20/05/2021 Current User: Lien Nguyen Review: MAG Controlled

47



REVIEW HOME

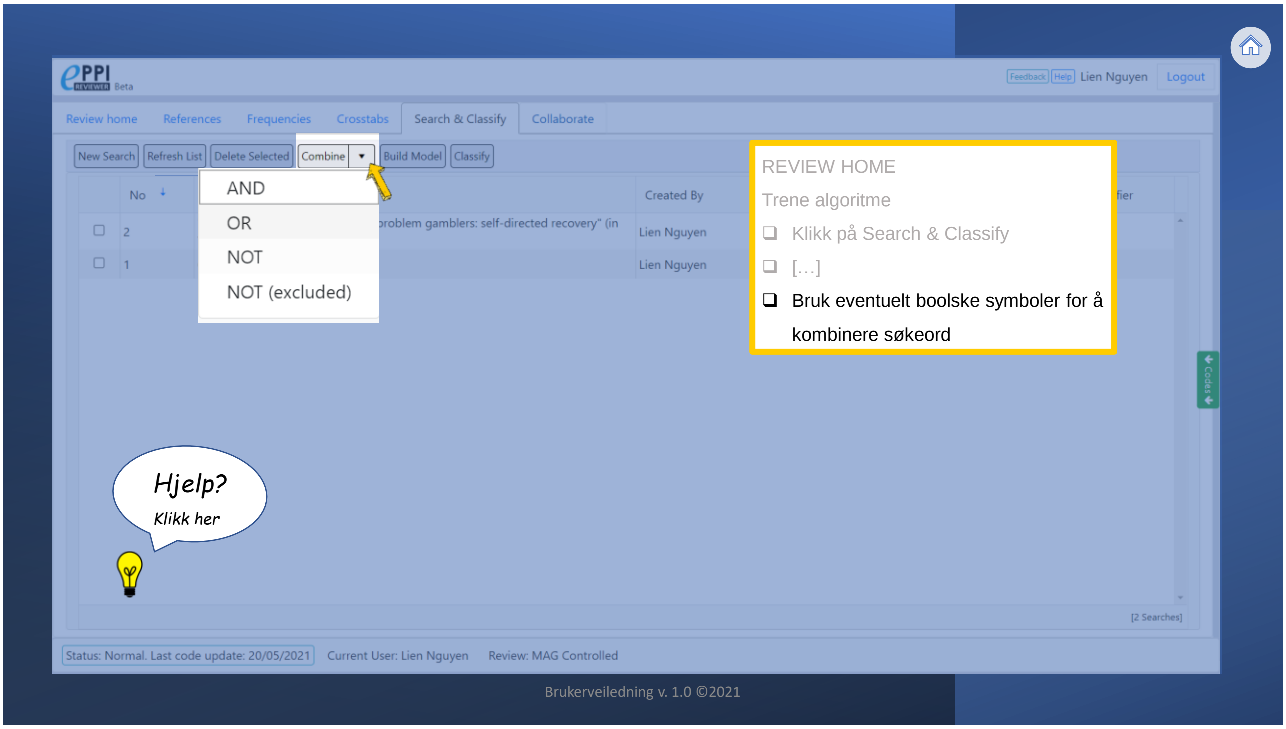
Trene algoritme

❑ Klikk på Search & Classify

❑ […]

❑ (1) Velg Containing this text (2) Tast inn søkeord (3) Velg hvor i artikkelen søkeord skal forekomme (4) Klikk op Run Search

48



REVIEW HOME

Trene algoritme

❑ Klikk på Search & Classify

❑ […]

❑ Bruk eventuelt boolske symboler for å kombinere søkeord

Brukerveiledning v. 1.0 ©2021

49

Machine learning classifiers – how to build your own in EPPI 4

What is a classifier?

Classification is the process of predicting data points. Classification predictive modelling is the task of predicting output variables from input variables. It belongs to the category of supervised learning where a human provides input data.

For example: Spam detection in email service providers can be identified as a classification problem. This is a binary classification since there are only 2 classes as spam and not spam. A classifier utilizes some training data to understand how given input variables relate to the class. In this case, known spam and non-spam emails have to be used as the training data. When the classifier is trained accurately, it can be used to detect an unknown email.

When is this relevant for you?

You have already coded a set of references in a dichotomous manner (e.g. *includes/excludes* from screening or priority screening). Now you want to see if you progress is sufficient to apply machine learning to further references to save time with screening or to prioritise your efforts on more relevant studies. With a decent model, you can expect to get a ranking of your further references by % likely relevance. This will also allow you to allocate references by % likely relevance to team members, or set yourself a cut-off percentage of % likely relevance to stop screening.

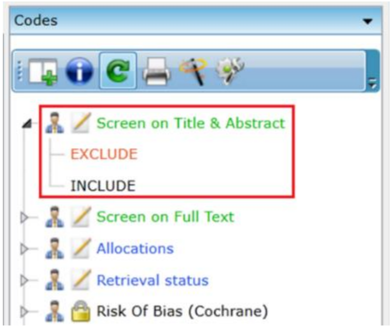
Note: a decent model can be built if you have enough *include/exclude* screening decisions to train the model with. The more, the better. You have to build your model before assessing how useful it is; see "How to interpret the results from your model?" at the end of the document for more detail.

How to set up your classifier:

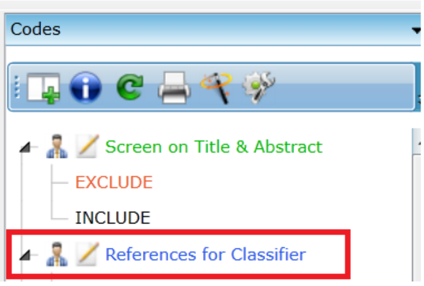
Before you get started you need a **training set** of known *includes /excludes* (e.g. your screening results). In addition, you need to create a code for all non-processed references to have them easily accessible.

1. Codesets

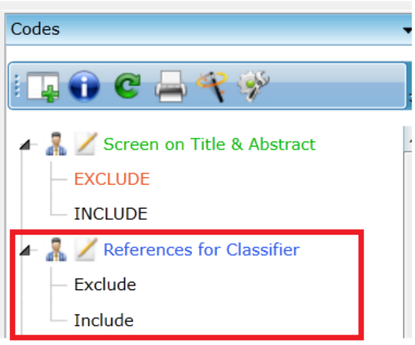
Have your includes/ excludes ready. To get most sensible predictions of likely relevance, you need to have a balanced ratio of includes/excludes (ideally, not exceeding 1:5). You will be guided in how to balance your studies.



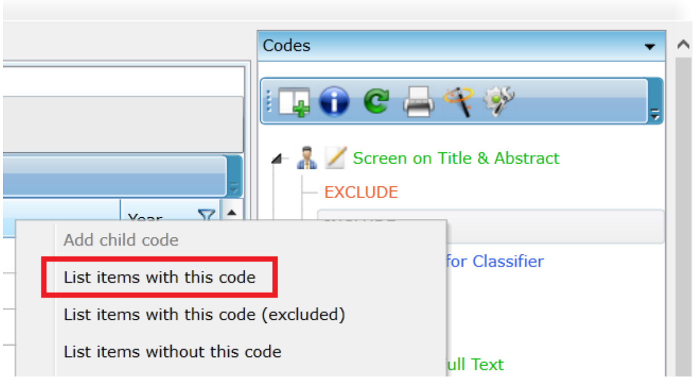
Create an **Administrative** codeset named: Reference for Classifier. Choose Codeset type: **Administrative**. It will then appear in blue.



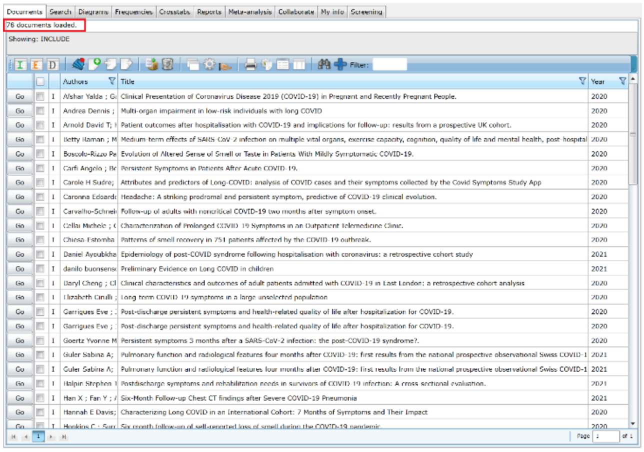
“Add a child code” via right clicking on the Reference for Classifier. One for *Includes* and one for *Excludes*



Check how many includes you have under “Screen on Title & Abstract”. Right-click on Include and “list items with this code”


Codes
Year
Screen on Title & Abstract
EXCLUDE
Add child code
List items with this code
List items with this code (excluded)
List items without this code
for Classifier
ull Text

In the example there are 78 includes. Remember/ write down your number of *Includes.*

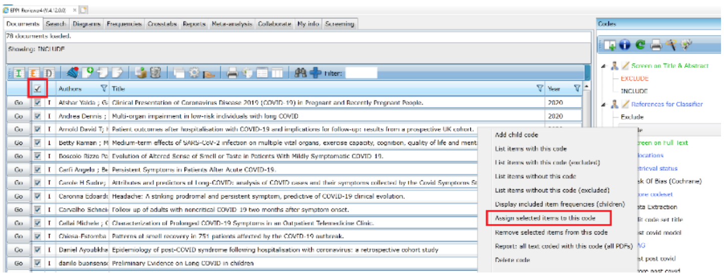


78 documents loaded.

Showing: INCLUDE

	☐		Authors	Title	Year
Go	☐	I	Afshar Yalda ; G	Clinical Presentation of Coronavirus Disease 2019 (COVID-19) in Pregnant and Recently Pregnant People.	2020
Go	☐	I	Andrea Dennis ;	Multi-organ impairment in low-risk individuals with long COVID	2020
Go	☐	I	Arnold David T;	Patient outcomes after hospitalisation with COVID-19 and implications for follow-up: results from a prospective UK cohort.	2020
Go	☐	I	Betty Raman ; M	Medium-term effects of SARS-CoV-2 infection on multiple vital organs, exercise capacity, cognition, quality of life and mental health, post-hospital	2020
Go	☐	I	Boscolo-Rizzo Pa	Evolution of Altered Sense of Smell or Taste in Patients With Mildly Symptomatic COVID-19.	2020
Go	☐	I	Carfi Angelo ; Bc	Persistent Symptoms in Patients After Acute COVID-19.	2020
Go	☐	I	Carole H Sudre;	Attributes and predictors of Long-COVID: analysis of COVID cases and their symptoms collected by the Covid Symptoms Study App	2020
Go	☐	I	Caronna Edoardo	Headache: A striking prodromal and persistent symptom, predictive of COVID-19 clinical evolution.	2020
Go	☐	I	Carvalho-Schnei	Follow-up of adults with noncritical COVID-19 two months after symptom onset.	2020
Go	☐	I	Cellai Michele ; (	Characterization of Prolonged COVID-19 Symptoms in an Outpatient Telemedicine Clinic.	2020
Go	☐	I	Chiesa-Estomba	Patterns of smell recovery in 751 patients affected by the COVID-19 outbreak.	2020
Go	☐	I	Daniel Ayoubkha	Epidemiology of post-COVID syndrome following hospitalisation with coronavirus: a retrospective cohort study	2021
Go	☐	I	danilo buonsens	Preliminary Evidence on Long COVID in children	2021
Go	☐	I	Daryl Cheng ; Cl	Clinical characteristics and outcomes of adult patients admitted with COVID-19 in East London: a retrospective cohort analysis	2020
Go	☐	I	Elizabeth Cirulli ;	Long term COVID-19 symptoms in a large unselected population	2020
Go	☐	I	Garrigues Eve ;	Post-discharge persistent symptoms and health-related quality of life after hospitalization for COVID-19.	2020
Go	☐	I	Garrigues Eve ;	Post-discharge persistent symptoms and health-related quality of life after hospitalization for COVID-19.	2020
Go	☐	I	Goertz Yvonne M	Persistent symptoms 3 months after a SARS-CoV-2 infection: the post-COVID-19 syndrome?.	2020
Go	☐	I	Guler Sabina A;	Pulmonary function and radiological features four months after COVID-19: first results from the national prospective observational Swiss COVID-1	2021
Go	☐	I	Guler Sabina A;	Pulmonary function and radiological features four months after COVID-19: first results from the national prospective observational Swiss COVID-1	2021
Go	☐	I	Halpin Stephen	Postdischarge symptoms and rehabilitation needs in survivors of COVID-19 infection: A cross-sectional evaluation.	2021
Go	☐	I	Han X ; Fan Y ;	Six-Month Follow-up Chest CT findings after Severe COVID-19 Pneumonia	2021
Go	☐	I	Hannah E Davis;	Characterizing Long COVID in an International Cohort: 7 Months of Symptoms and Their Impact	2020
Go	☐	I	Hopkins C ; Surr	Six month follow-up of self-reported loss of smell during the COVID-19 pandemic	2020

Select all references and assign them to childcode “Include” of the codeset Reference for Classifier.



78 documents loaded.

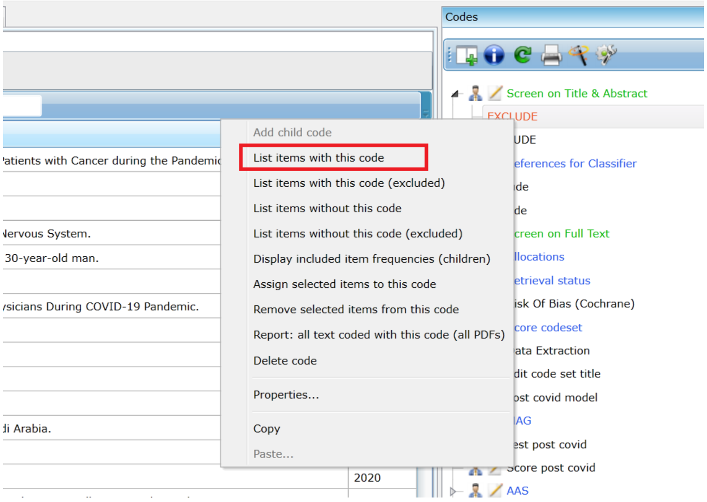
Showing: INCLUDE

	☑		Authors	Title	Year
Go	☑	I	Afshar Yalda ; G	Clinical Presentation of Coronavirus Disease 2019 (COVID-19) in Pregnant and Recently Pregnant People.	2020
Go	☑	I	Andrea Dennis ;	Multi-organ impairment in low-risk individuals with long COVID	2020
Go	☑	I	Arnold David T;	Patient outcomes after hospitalisation with COVID-19 and implications for follow-up: results from a prospective UK cohort.	
Go	☑	I	Betty Raman ; M	Medium-term effects of SARS-CoV-2 infection on multiple vital organs, exercise capacity, cognition, quality of life and ment	
Go	☑	I	Boscolo Rizzo Pa	Evolution of Altered Sense of Smell or Taste in Patients With Mildly Symptomatic COVID-19.	
Go	☑	I	Carfi Angelo ; Be	Persistent Symptoms in Patients After Acute COVID-19.	
Go	☑	I	Carole H Sudre;	Attributes and predictors of Long-COVID: analysis of COVID cases and their symptoms collected by the Covid Symptoms St	
Go	☑	I	Caronna Edoardo	Headache: A striking prodromal and persistent symptom, predictive of COVID-19 clinical evolution.	
Go	☑	I	Carvalho Schnei	Follow up of adults with noncritical COVID-19 two months after symptom onset.	
Go	☑	I	Cellai Michele ; C	Characterization of Prolonged COVID-19 Symptoms in an Outpatient Telemedicine Clinic.	
Go	☑	I	Chiesa-Estomba	Patterns of smell recovery in 751 patients affected by the COVID-19 outbreak.	
Go	☑	I	Daniel Ayoubkha	Epidemiology of post-COVID syndrome following hospitalisation with coronavirus: a retrospective cohort study	
Go	☑	I	danilo buonsenso	Preliminary Evidence on Long COVID in children	

EPPI Reviewer4 (V.4.12.0.0)
Documents
Search
Diagrams
Frequencies
Crosstabs
Reports
Meta-analysis
Collaborate
My info
Screening
Filter:
Codes
Screen on Title & Abstract
EXCLUDE
INCLUDE
References for Classifier
Exclude
Add child code
List items with this code
List items with this code (excluded)
List items without this code
List items without this code (excluded)
Display included item frequencies (children)
Assign selected items to this code
Remove selected items from this code
Report: all text coded with this code (all PDFs)
Delete code

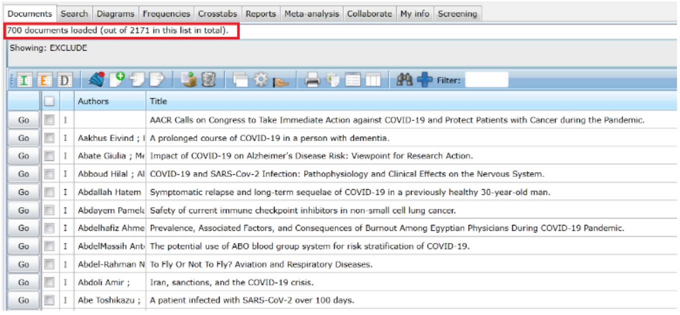
You now need to assign a selection of your Excludes to childcode “Exclude” of the codeset Reference for Classifier.

Right-click on “Exclude”, and then “list items with this code”.



Codes
Screen on Title & Abstract
EXCLUDE
UDE
eferences for Classifier
de
de
creen on Full Text
llocations
etrieval status
isk Of Bias (Cochrane)
core codeset
ata Extraction
dit code set title
ost covid model
AG
est post covid
Score post covid
AAS
Add child code
List items with this code
List items with this code (excluded)
List items without this code
List items without this code (excluded)
Display included item frequencies (children)
Assign selected items to this code
Remove selected items from this code
Report: all text coded with this code (all PDFs)
Delete code
Properties...
Copy
Paste...
atients with Cancer during the Pandemic
Nervous System.
30-year-old man.
ysicians During COVID-19 Pandemic.
di Arabia.
2020

In the example there are 2171 references coded as “Exclude”.



700 documents loaded (out of 2171 in this list in total).

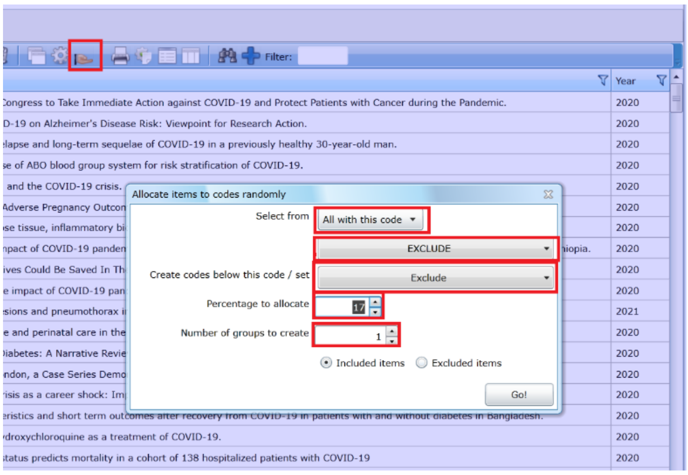
Showing: EXCLUDE

	□		Authors	Title
Go	□	I		AACR Calls on Congress to Take Immediate Action against COVID-19 and Protect Patients with Cancer during the Pandemic.
Go	□	I	Aakhus Eivind ; I	A prolonged course of COVID-19 in a person with dementia.
Go	□	I	Abate Giulia ; Me	Impact of COVID-19 on Alzheimer's Disease Risk: Viewpoint for Research Action.
Go	□	I	Abboud Hilal ; Al	COVID-19 and SARS-Cov-2 Infection: Pathophysiology and Clinical Effects on the Nervous System.
Go	□	I	Abdallah Hatem	Symptomatic relapse and long-term sequelae of COVID-19 in a previously healthy 30-year-old man.
Go	□	I	Abdayem Pamela	Safety of current immune checkpoint inhibitors in non-small cell lung cancer.
Go	□	I	Abdelhafiz Ahme	Prevalence, Associated Factors, and Consequences of Burnout Among Egyptian Physicians During COVID-19 Pandemic.
Go	□	I	AbdelMassih Ant	The potential use of ABO blood group system for risk stratification of COVID-19.
Go	□	I	Abdel-Rahman N	To Fly Or Not To Fly? Aviation and Respiratory Diseases.
Go	□	I	Abdoli Amir ;	Iran, sanctions, and the COVID-19 crisis.
Go	□	I	Abe Toshikazu ;	A patient infected with SARS-CoV-2 over 100 days.

To allocate a selection of “Exclude” not more than 1:5 of Include (i.e. 5 x 78 = 390), click the hand symbol to “Allocate items to codes randomly”.

To not exceed the 1:5 ratio, calculate the correct amount percentage you need to assign.

In the example: (5 x 78) / (2171/100)= 17.97. So you need to allocate 17% in one group to the childcode “Exclude” of the codeset Reference for Classifier.



Allocate items to codes randomly

Select from: All with this code

EXCLUDE

Create codes below this code / set: Exclude

Percentage to allocate: 17

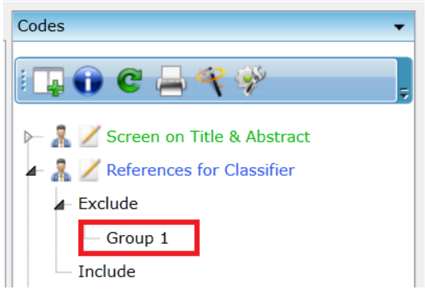
Number of groups to create: 1

● Included items ○ Excluded items

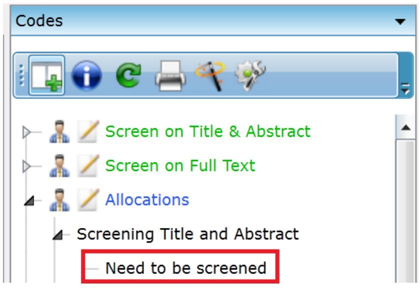
Go!

Title	Year
Congress to Take Immediate Action against COVID-19 and Protect Patients with Cancer during the Pandemic.	2020
D-19 on Alzheimer's Disease Risk: Viewpoint for Research Action.	2020
lapse and long-term sequelae of COVID-19 in a previously healthy 30-year-old man.	2020
se of ABO blood group system for risk stratification of COVID-19.	2020
and the COVID-19 crisis.	2020
Adverse Pregnancy Outcon	2020
se tissue, inflammatory bi	2020
npact of COVID-19 panden … iopia.	2020
ives Could Be Saved In Th	2020
e impact of COVID-19 pan	2020
sions and pneumothorax i	2021
e and perinatal care in the	2020
Diabetes: A Narrative Revie	2020
ndon, a Case Series Demo	2020
risis as a career shock: Im	2020
eristics and short term outcomes after recovery from COVID-19 in patients with and without diabetes in Bangladesh.	2020
ydroxychloroquine as a treatment of COVID-19.	2020
status predicts mortality in a cohort of 138 hospitalized patients with COVID-19	2020

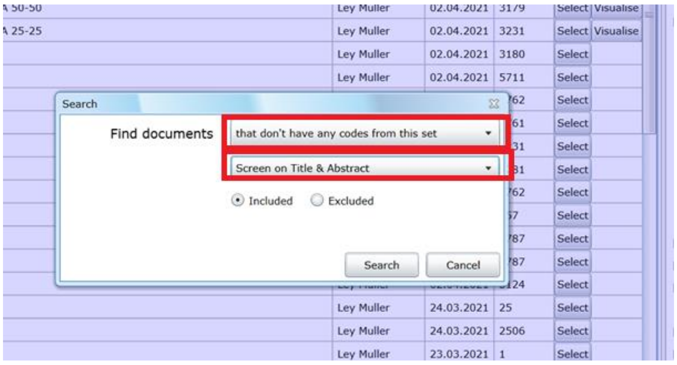
Under the codeset Reference for Classifier/ Exclude you find "*Group 1*" – your random selection of excludes.



Your references which haven't been screened need to have their own code too. For example, you can code them to a code "*need to be screened*" under the allocations codeset.



To find the not screened studies, go to the search tab, and search for studies "*that don't have any codes from this set*" "*Screen on Title & Abstract*". Assign these studies to your "Need to be screened" code.

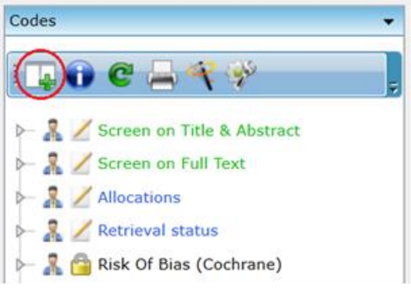

Search

Find documents: that don't have any codes from this set

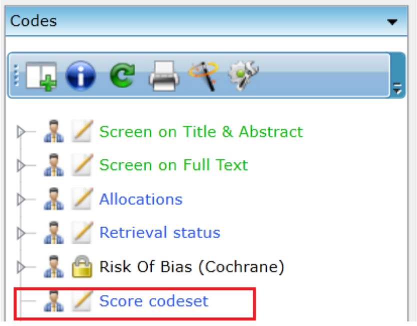
Screen on Title & Abstract

● Included ○ Excluded

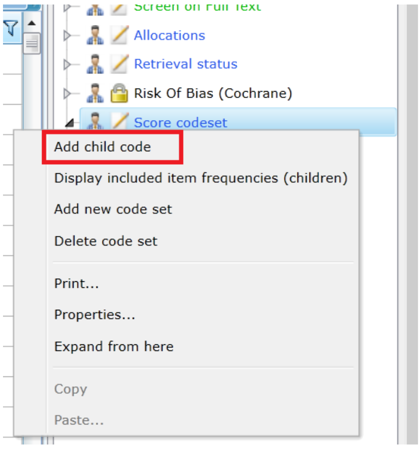
1.a) Create an **Administrative** codeset named: Score codeset



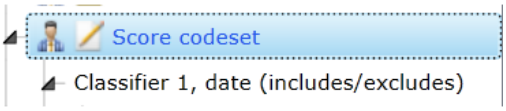
1. b) Check that your: Score codeset is visible and blue



1. c) “Add a child code” via right clicking on the score codeset

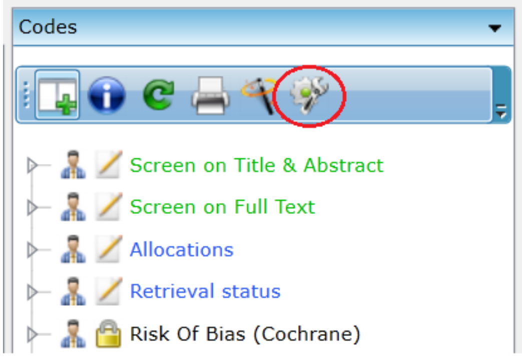


1. d) Name the childe code: Number and date it, and provide the information on how many includes/excludes you have ready

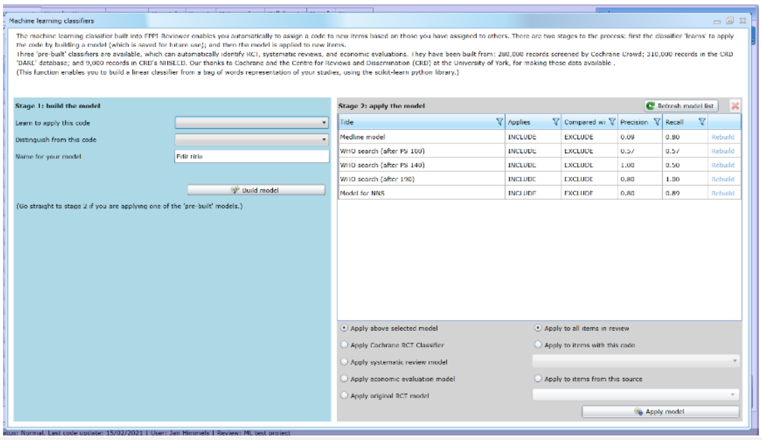


2. The Classifier menu

Click on the spanner “classifier” icon to get the Machine building classifier menu



The Machine building classifier menu



Machine learning classifiers

The machine learning classifier built into EPPI Reviewer enables you automatically to assign a code to new items based on those you have assigned to others. There are two stages to the process: first the classifier 'learns' to apply the code by building a model (which is saved for future use); and then the model is applied to new items.
Three 'pre-built' classifiers are available, which can automatically identify RCT, systematic reviews, and economic evaluations. They have been built from: 280,000 records screened by Cochrane Crowd; 310,000 records in the CRD 'DARE' database; and 9,000 records in CRD's NHSEED. Our thanks to Cochrane and the Centre for Reviews and Dissemination (CRD) at the University of York, for making these data available .
(This function enables you to build a linear classifier from a bag of words representation of your studies, using the scikit-learn python library.)

Stage 1: build the model

Learn to apply this code

Distinguish from this code

Name for your model: Edit title

Build model

(Go straight to stage 2 if you are applying one of the 'pre-built' models.)

Stage 2: apply the model

Refresh model list

Title	Applies	Compared wi	Precision	Recall	
Medline model	INCLUDE	EXCLUDE	0.09	0.80	Rebuild
WHO search (after PS 100)	INCLUDE	EXCLUDE	0.57	0.57	Rebuild
WHO search (after PS 140)	INCLUDE	EXCLUDE	1.00	0.50	Rebuild
WHO search (after 190)	INCLUDE	EXCLUDE	0.80	1.00	Rebuild
Model for NNS	INCLUDE	EXCLUDE	0.80	0.89	Rebuild

● Apply above selected model
○ Apply Cochrane RCT Classifier
○ Apply systematic review model
○ Apply economic evaluation model
○ Apply original RCT model

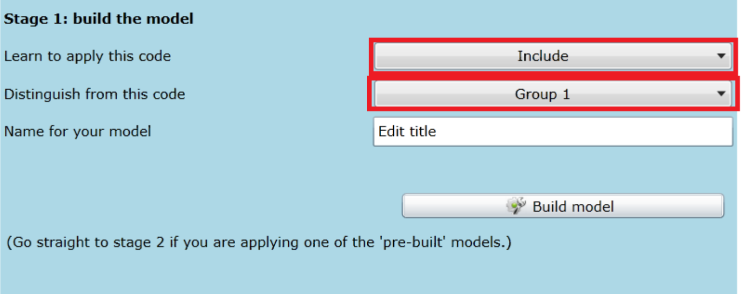
● Apply to all items in review
○ Apply to items with this code
○ Apply to items from this source

Apply model

atus: Normal. Last code update: 15/02/2021 | User: Jan Himmels | Review: ML test project

3. Build the model

Apply the **Reference for Classifier**: **Include** code from **Reference for Classifier/ Exclude/ Group 1** code.



Stage 1: build the model

Learn to apply this code: Include

Distinguish from this code: Group 1

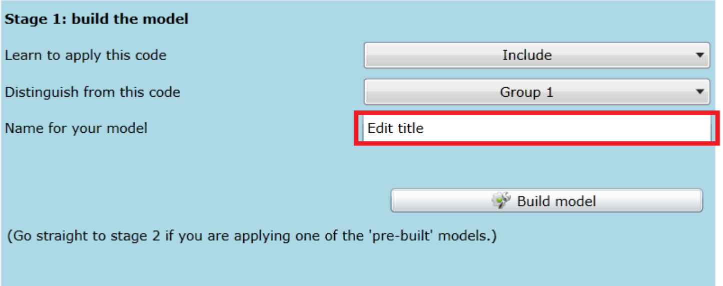
Name for your model: Edit title

Build model

(Go straight to stage 2 if you are applying one of the 'pre-built' models.)

Name the model “Classifier INCLUDE vs EXCLUDE, [number of include – number of exclude]”

Example: “Classifier INCLUDE vs EXCLUDE, 50-200” shows that this model has been trained by 50 included studies and 200 excluded studies.



Stage 1: build the model

Learn to apply this code: Include

Distinguish from this code: Group 1

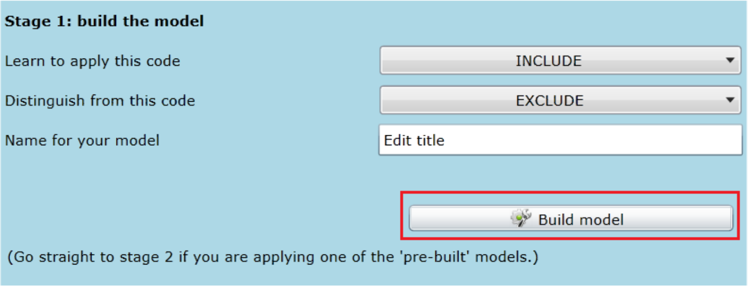
Name for your model: Edit title

Build model

(Go straight to stage 2 if you are applying one of the 'pre-built' models.)

Build the model

(Wait a few minutes.)



Stage 1: build the model

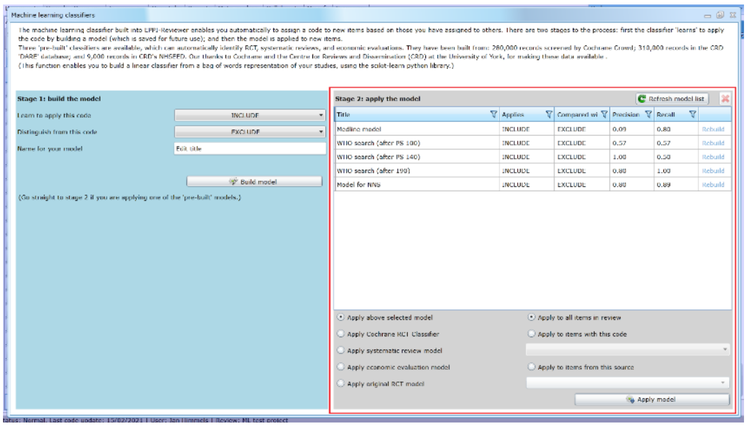
Learn to apply this code	INCLUDE
Distinguish from this code	EXCLUDE
Name for your model	Edit title

(Go straight to stage 2 if you are applying one of the 'pre-built' models.)

Your **model is ready** based on your *includes* and *exludes*!

4. Apply the model

Go to Stage 2 (right side): Applying the model to un-coded/not screened studies



The machine learning classifier built into EPPI-Reviewer enables you automatically to assign a code to new items based on those you have assigned to others. There are two stages to the process: first the classifier 'learns' to apply the code by building a model (which is saved for future use); and then the model is applied to new items.
Three 'pre-built' classifiers are available, which can automatically identify RCT, systematic reviews, and economic evaluations. They have been built from: 280,000 records screened by Cochrane Crowd; 310,000 records in the CRD 'DARE' database; and 9,000 records in CRD's NHSEED. Our thanks to Cochrane and the Centre for Reviews and Dissemination (CRD) at the University of York, for making these data available .
(This function enables you to build a linear classifier from a bag of words representation of your studies, using the scikit-learn python library.)

Stage 1: build the model

Learn to apply this code	INCLUDE
Distinguish from this code	EXCLUDE
Name for your model	Edit title

(Go straight to stage 2 if you are applying one of the 'pre-built' models.)

Stage 2: apply the model

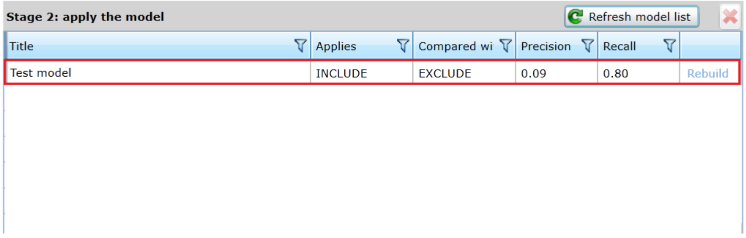
Title	Applies	Compared wi	Precision	Recall	
Medline model	INCLUDE	EXCLUDE	0.09	0.80	Rebuild
WHO search (after PS 100)	INCLUDE	EXCLUDE	0.57	0.57	Rebuild
WHO search (after PS 140)	INCLUDE	EXCLUDE	1.00	0.50	Rebuild
WHO search (after 190)	INCLUDE	EXCLUDE	0.80	1.00	Rebuild
Model for NNS	INCLUDE	EXCLUDE	0.80	0.89	Rebuild

● Apply above selected model
○ Apply Cochrane RCT Classifier
○ Apply systematic review model
○ Apply economic evaluation model
○ Apply original RCT model

● Apply to all items in review
○ Apply to items with this code
○ Apply to items from this source

tatus: Normal. Last code update: 15/02/2021 | User: Jan Hummels | Review: ML test project

4.a) Select the model you just built

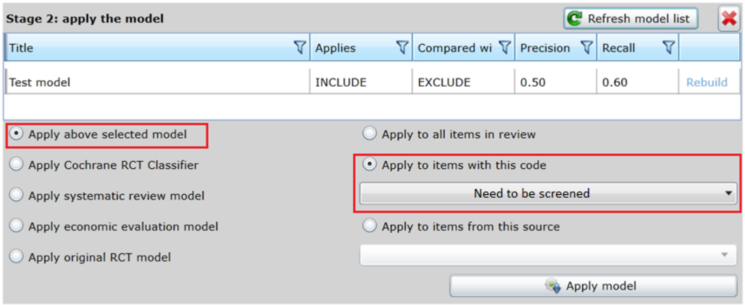


Stage 2: apply the model

Title	Applies	Compared wi	Precision	Recall	
Test model	INCLUDE	EXCLUDE	0.09	0.80	Rebuild

4.b) Select the studies to apply the model to:

specific code (that describes your un-processed studies, i.e. "need to be screened" *(the code specified in point 1)* **or** a *specific source (i.e. a RIS-file)*



Stage 2: apply the model

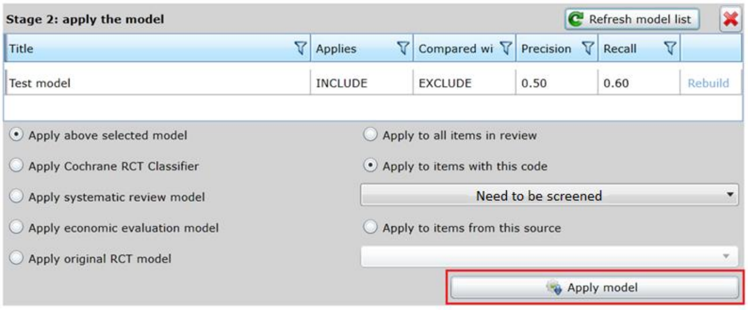
Title	Applies	Compared wi	Precision	Recall	
Test model	INCLUDE	EXCLUDE	0.50	0.60	Rebuild

● Apply above selected model
○ Apply Cochrane RCT Classifier
○ Apply systematic review model
○ Apply economic evaluation model
○ Apply original RCT model

○ Apply to all items in review
● Apply to items with this code
Need to be screened
○ Apply to items from this source

Now: Apply model

Wait for a few minutes.



Stage 2: apply the model

Title	Applies	Compared wi	Precision	Recall	
Test model	INCLUDE	EXCLUDE	0.50	0.60	Rebuild

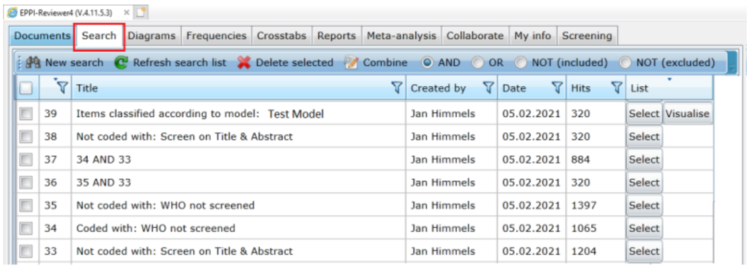
● Apply above selected model
○ Apply Cochrane RCT Classifier
○ Apply systematic review model
○ Apply economic evaluation model
○ Apply original RCT model

○ Apply to all items in review
● Apply to items with this code
Need to be screened
○ Apply to items from this source

5. Find the results of your model

Choose the "*Search*" tab to see the results.

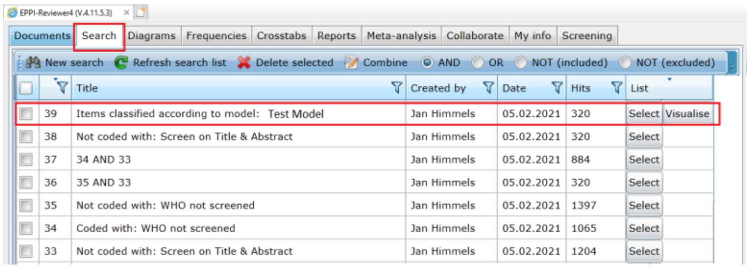
You will likely have to click "*Refresh search list*" a few times



		Title	Created by	Date	Hits	List	
□	39	Items classified according to model: Test Model	Jan Himmels	05.02.2021	320	Select	Visualise
□	38	Not coded with: Screen on Title & Abstract	Jan Himmels	05.02.2021	320	Select	
□	37	34 AND 33	Jan Himmels	05.02.2021	884	Select	
□	36	35 AND 33	Jan Himmels	05.02.2021	320	Select	
□	35	Not coded with: WHO not screened	Jan Himmels	05.02.2021	1397	Select	
□	34	Coded with: WHO not screened	Jan Himmels	05.02.2021	1065	Select	
□	33	Not coded with: Screen on Title & Abstract	Jan Himmels	05.02.2021	1204	Select	

By clicking "*Visualise*", you get a distribution chart. By clicking on "*Select*", you get a list of the references with ranking by relevance.

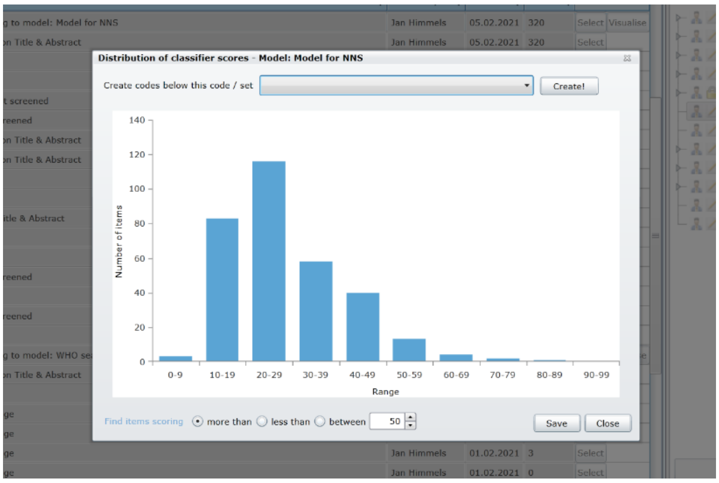
You want to **visualise** your results



		Title	Created by	Date	Hits	List	
□	39	Items classified according to model: Test Model	Jan Himmels	05.02.2021	320	Select	Visualise
□	38	Not coded with: Screen on Title & Abstract	Jan Himmels	05.02.2021	320	Select	
□	37	34 AND 33	Jan Himmels	05.02.2021	884	Select	
□	36	35 AND 33	Jan Himmels	05.02.2021	320	Select	
□	35	Not coded with: WHO not screened	Jan Himmels	05.02.2021	1397	Select	
□	34	Coded with: WHO not screened	Jan Himmels	05.02.2021	1065	Select	
□	33	Not coded with: Screen on Title & Abstract	Jan Himmels	05.02.2021	1204	Select	

After clicking "*Visualise*", a distribution chart **pops up**

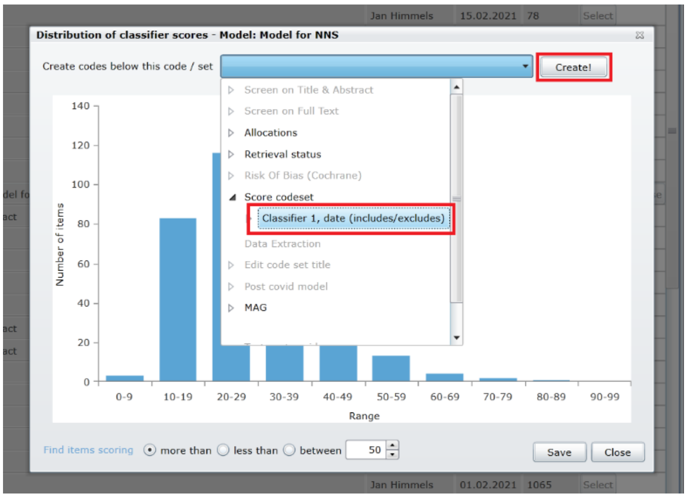
In the example to the right, about 120 studies are ranked as 20-29% likely included; only a few are ranked as 0-9% likely.



6. Saving your results as codes by % likely relevance

Select the child code under the “Score codeset” to save each bar as a code (the **child code** you created in Step 1.d).

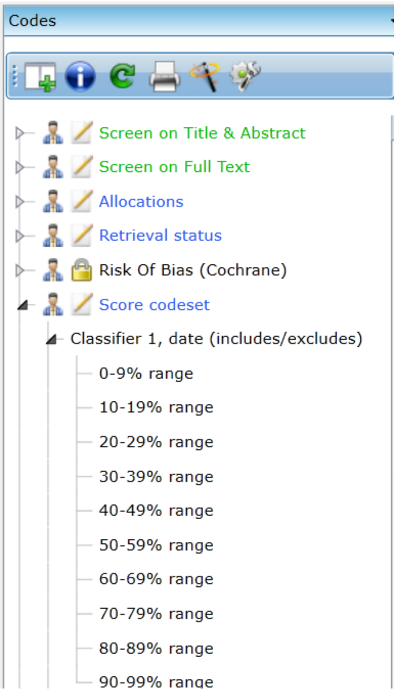
Click “**Create!**”



Distribution of classifier scores - Model: Model for NNS

Create codes below this code / set

Find items scoring ● more than ○ less than ○ between 50



Codes

- Screen on Title & Abstract
- Screen on Full Text
- Allocations
- Retrieval status
- Risk Of Bias (Cochrane)
- Score codeset
  - Classifier 1, date (includes/excludes)
    - 0-9% range
    - 10-19% range
    - 20-29% range
    - 30-39% range
    - 40-49% range
    - 50-59% range
    - 60-69% range
    - 70-79% range
    - 80-89% range
    - 90-99% range

Under your administrative Score codeset, and under the **child code** you can find each bar from the chart, as its own code.

7. Using your results

Consider your options:

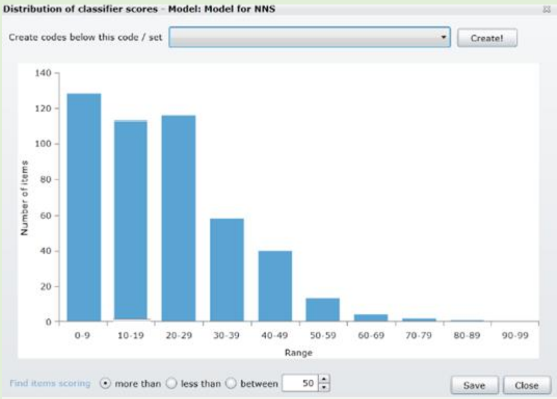
With your results ready, you need to assess their usefulness, if you are satisfied with the results you may want to code studies with low/high likely relevance as includes or excludes, or you can allocate them to a member of your team so screen them.

Interpret your results

A decent model

Your results, visualised in the bar chart, reflect the strength of the model. The example below shows a distribution with few studies having a high % likely relevance, and gradually more with less likely relevance. The example reflects a rather good model, with the most relevant studies already having been identified.

In this case you can continue on to changing your screening procedures.



Distribution of classifier scores - Model: Model for NNS

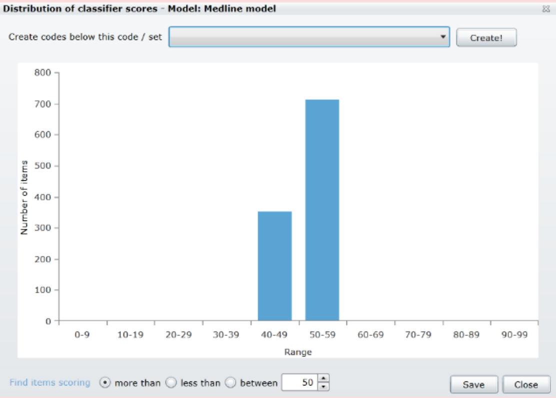
Create codes below this code / set

Find items scoring ● more than ○ less than ○ between 50

A model that needs to be trained more or adjusted

The results of a less successful model are depicted below. The model was not able to be very certain in which studies were most likely relevant, or which studies were unlikely relevant. This indicates that the classifier had too little data available to make more certain predictions.

In this case, you should continue screening, and rebuild the model once you have screened more studies (rule of thumb: 50-100 studies).



Distribution of classifier scores - Model: Medline model

Create codes below this code / set

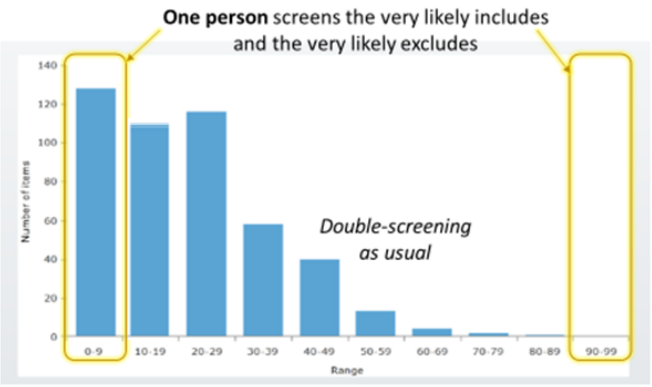
Find items scoring ● more than ○ less than ○ between 50

If after you continue screening your model is still clustered around 50-60%, try making your *includes* and *excludes* more balanced. This likely means picking a new, smaller random selection of *excludes.*

Changing your screening procedures based on your classifier

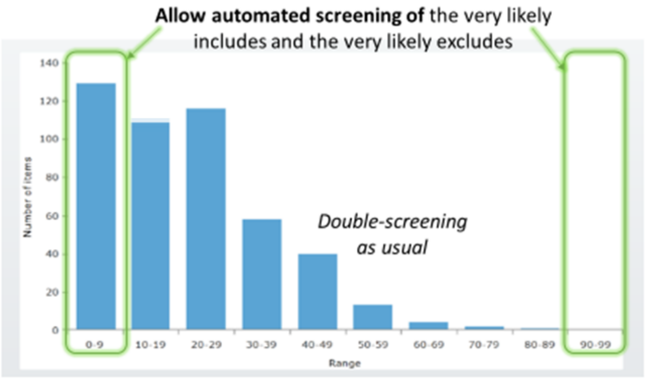
If you have built a decent classifier, you have several options. Some examples:

One person, instead of two, can confirm the studies classified as very likely (90-99%) and as very unlikely (0-9%).



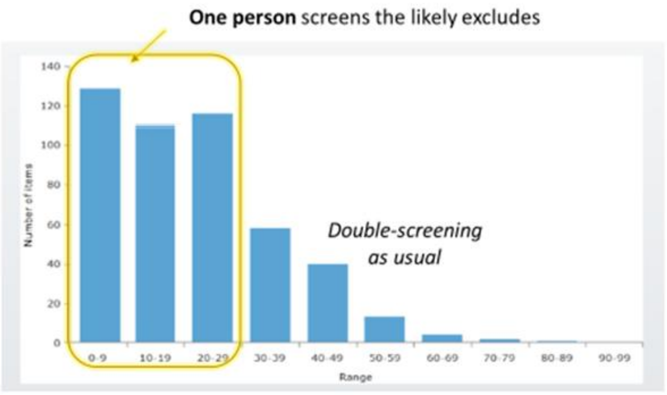
One person screens the very likely includes and the very likely excludes

Without manual confirmation, you can screen the studies classified as very likely (90-99%) and as very unlikely (0-9%) according to the classifier's prediction.



Allow automated screening of the very likely includes and the very likely excludes

One person, instead of two, can confirm the studies classified as less likely (0-29%).



One person screens the likely excludes

Or other combinations.

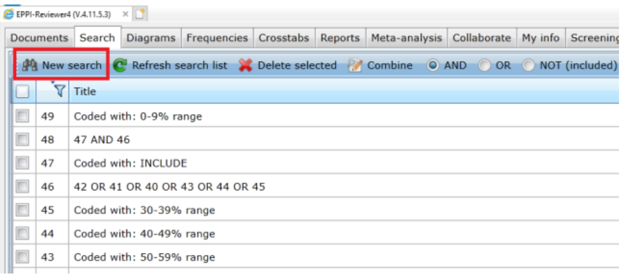
You could also de-prioritize the screening of least likely studies, so that the team proceeds with other tasks, and these least-likely studies are screened whenever people have time.

8. a) How to accept the classifier's screening predictions

If you want to accept the classifier's prediction of a screening code (without a human screener), you must still be the one to actually assign a code.

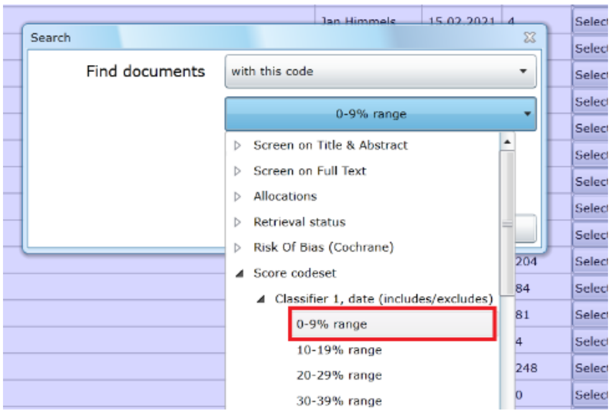
You can do this by searching and coding in bulk. E.g. you decide to exclude all studies with less than 10% likely relevance.

Open the search tab and create a new search.



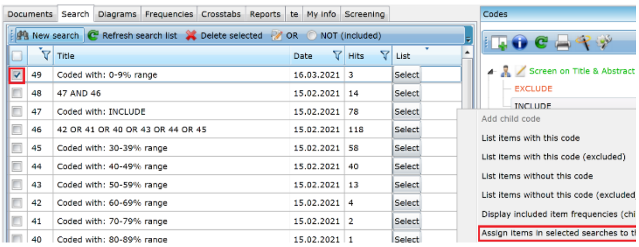
□		Title
□	49	Coded with: 0-9% range
□	48	47 AND 46
□	47	Coded with: INCLUDE
□	46	42 OR 41 OR 40 OR 43 OR 44 OR 45
□	45	Coded with: 30-39% range
□	44	Coded with: 40-49% range
□	43	Coded with: 50-59% range

Search for the % range you want to assign the include/ exclude code to (e.g. 0-9% range)



Select the search result via the checkbox

In the Codes menu on the right side, right-click on "EXCLUDE", then click "Assign items in selected searches to this code."


Documents
Search
Diagrams
Frequencies
Crosstabs
Reports
te
My Info
Screening
New search
Refresh search list
Delete selected
OR
NOT (included)
Codes
Screen on Title & Abstract
EXCLUDE
INCLUDE
Add child code
List items with this code
List items with this code (excluded)
List items without this code
List items without this code (excluded
Display included item frequencies (chi
Assign items in selected searches to t

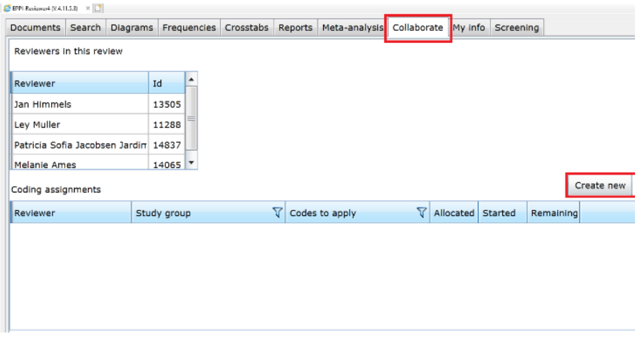
□		Title	Date	Hits	List
☑	49	Coded with: 0-9% range	16.03.2021	3	Select
□	48	47 AND 46	15.02.2021	14	Select
□	47	Coded with: INCLUDE	15.02.2021	78	Select
□	46	42 OR 41 OR 40 OR 43 OR 44 OR 45	15.02.2021	118	Select
□	45	Coded with: 30-39% range	15.02.2021	58	Select
□	44	Coded with: 40-49% range	15.02.2021	40	Select
□	43	Coded with: 50-59% range	15.02.2021	13	Select
□	42	Coded with: 60-69% range	15.02.2021	4	Select
□	41	Coded with: 70-79% range	15.02.2021	2	Select
□	40	Coded with: 80-89% range	15.02.2021	1	Select

NB! If your **Screening on T/A codeset** is set up to require two persons' coding ("Comparison coding"), and you want to keep this set-up rather than change to allow single-person coding ("Normal coding"), then a second person needs to screen these studies. Allocate this same range to a second person with instruction to bulk-screen them as you did, then make a comparison as you normally would to confirm screening.

8. b) How to allocate studies by likely relevance

If you want certain team members to prioritize screening of certain studies based on likely relevance, you can create specific allocations in the "**Collaborate**" tab.

Click **"Create new"**



Reviewers in this review

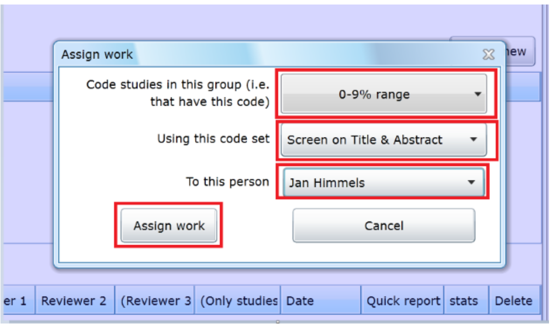
Reviewer	Id
Jan Himmels	13505
Ley Muller	11288
Patricia Sofia Jacobsen Jardin	14837
Melanie Ames	14065

Coding assignments

Reviewer	Study group	Codes to apply	Allocated	Started	Remaining

i. Select the range you want to allocate.
ii. Select the codeset you want the individual to code
iii. Select the person to allocate to

iv. Assign the work.

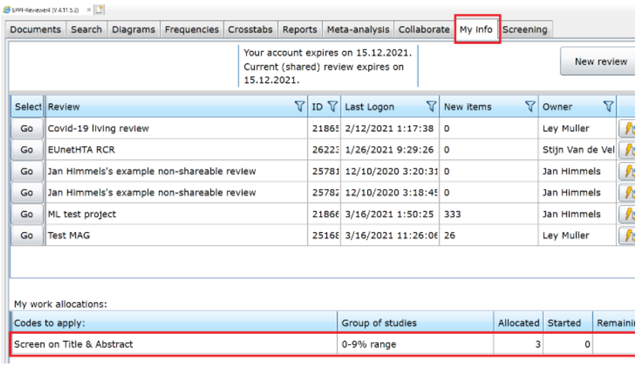


Code studies in this group (i.e. that have this code) 0-9% range

Using this code set Screen on Title & Abstract

To this person Jan Himmels

The person to whom you allocated to will see the assigned references, in the "**My info**" tab, and there under "**My work allocations**".



Your account expires on 15.12.2021. Current (shared) review expires on 15.12.2021.

Select	Review	ID	Last Logon	New items	Owner
Go	Covid-19 living review	21865	2/12/2021 1:17:38	0	Ley Muller
Go	EUnetHTA RCR	26223	1/26/2021 9:29:26	0	Stijn Van de Vel
Go	Jan Himmels's example non-shareable review	25781	12/10/2020 3:20:31	0	Jan Himmels
Go	Jan Himmels's example non-shareable review	25782	12/10/2020 3:18:45	0	Jan Himmels
Go	ML test project	21866	3/16/2021 1:50:25	333	Jan Himmels
Go	Test MAG	25168	3/16/2021 11:26:06	26	Ley Muller

My work allocations:

Codes to apply:	Group of studies	Allocated	Started	Remaini
Screen on Title & Abstract	0-9% range	3	0	

Risk of Bias assessments with machine learning

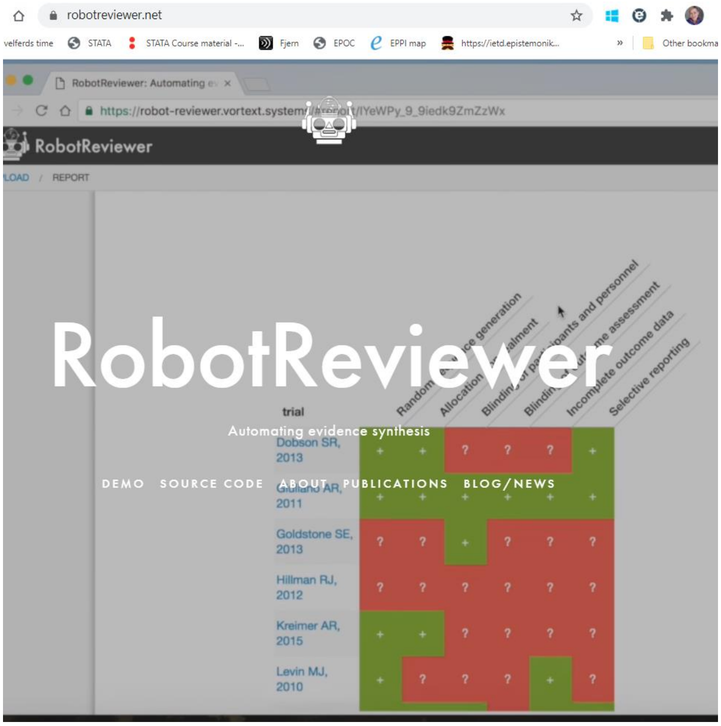
In EPPI-Reviewer

Instructions for ***team leaders***

1

Technology

- www.robotreviewer.net


robotreviewer.net
velferds time
STATA
STATA Course material -...
Fjern
EPOC
EPPI map
https://ietd.epistemonik...
Other bookma
RobotReviewer: Automating ev
https://robot-reviewer.vortext.systems/#report/lYeWPy_9_9iedk9ZmZzWx
RobotReviewer
LOAD / REPORT
RobotReviewer
Automating evidence synthesis
DEMO
SOURCE CODE
ABOUT
PUBLICATIONS
BLOG/NEWS

trial	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting
Dobson SR, 2013	+	+	?	?	?	+
Giuliano AR, 2011	+	+	+	+	+	+
Goldstone SE, 2013	?	?	+	?	?	?
Hillman RJ, 2012	?	?	?	?	?	?
Kreimer AR, 2015	+	+	?	?	?	?
Levin MJ, 2010	+	?	?	?	+	?

2

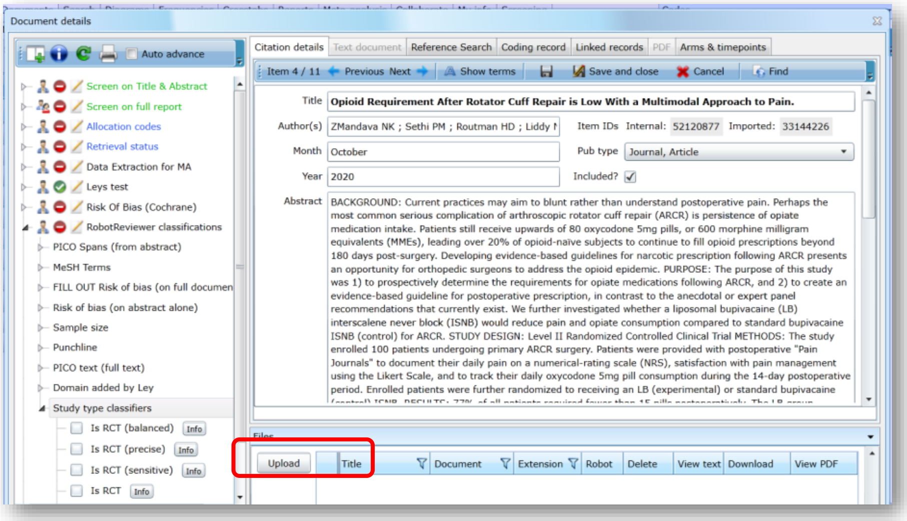
Before you begin RoB assessments

- Request Silverlight access from NHN for you and your team (as early as possible)
- Set your team up in EPPI
- Call in Ley/someone from the machine learning team to talk through possible procedures, such as:
  - Should your team members be blinded to RobotReviewer?
  - Do you want to compare to not using machine learning?
  - Can we collect some data?
- Set up a 1-hour training meeting with your team and Ley, to explain procedures
- Recommended: another 1-hour meeting with your team and Ley, for them to begin assessments

6 steps

1. Upload pdfs

http://eppi.ioe.ac.uk/eppireviewer4/eppireviewer4.aspx



Title: **Opioid Requirement After Rotator Cuff Repair is Low With a Multimodal Approach to Pain.**

Author(s): ZMandava NK ; Sethi PM ; Routman HD ; Liddy N

Item IDs Internal: 52120877 Imported: 33144226

Month: October

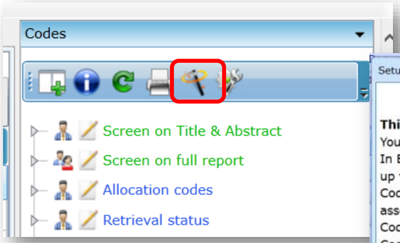
Pub type: Journal, Article

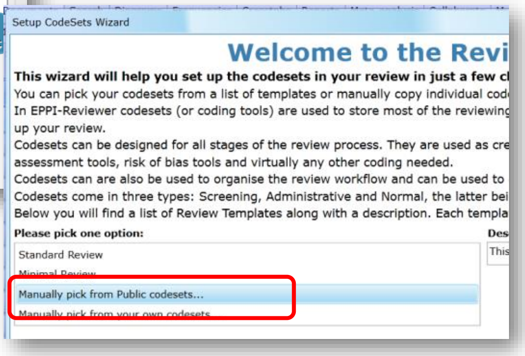
Year: 2020

Included? ☑

Abstract: BACKGROUND: Current practices may aim to blunt rather than understand postoperative pain. Perhaps the most common serious complication of arthroscopic rotator cuff repair (ARCR) is persistence of opiate medication intake. Patients still receive upwards of 80 oxycodone 5mg pills, or 600 morphine milligram equivalents (MMEs), leading over 20% of opioid-naïve subjects to continue to fill opioid prescriptions beyond 180 days post-surgery. Developing evidence-based guidelines for narcotic prescription following ARCR presents an opportunity for orthopedic surgeons to address the opioid epidemic. PURPOSE: The purpose of this study was 1) to prospectively determine the requirements for opiate medications following ARCR, and 2) to create an evidence-based guideline for postoperative prescription, in contrast to the anecdotal or expert panel recommendations that currently exist. We further investigated whether a liposomal bupivacaine (LB) interscalene never block (ISNB) would reduce pain and opiate consumption compared to standard bupivacaine ISNB (control) for ARCR. STUDY DESIGN: Level II Randomized Controlled Clinical Trial METHODS: The study enrolled 100 patients undergoing primary ARCR surgery. Patients were provided with postoperative "Pain Journals" to document their daily pain on a numerical-rating scale (NRS), satisfaction with pain management using the Likert Scale, and to track their daily oxycodone 5mg pill consumption during the 14-day postoperative period. Enrolled patients were further randomized to receiving an LB (experimental) or standard bupivacaine

2. Add RobotReviewer code set




Setup CodeSets Wizard

Welcome to the Revi

This wizard will help you set up the codesets in your review in just a few cl

You can pick your codesets from a list of templates or manually copy individual cod

In EPPI-Reviewer codesets (or coding tools) are used to store most of the reviewing

up your review.

Codesets can be designed for all stages of the review process. They are used as cre

assessment tools, risk of bias tools and virtually any other coding needed.

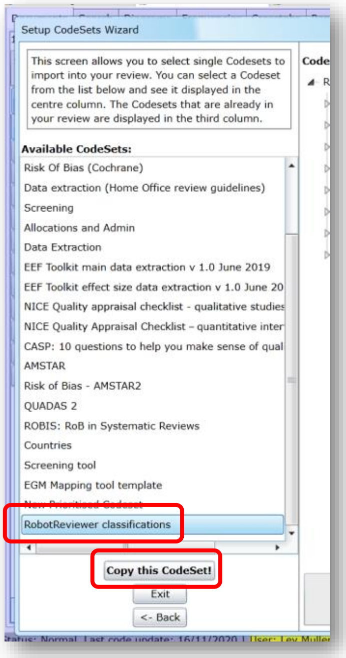
Codesets can are also be used to organise the review workflow and can be used to

Codesets come in three types: Screening, Administrative and Normal, the latter bei

Below you will find a list of Review Templates along with a description. Each templa

Please pick one option:

Des
This
Standard Review
Minimal Review
Manually pick from Public codesets...
Manually pick from your own codesets


Setup CodeSets Wizard

This screen allows you to select single Codesets to import into your review. You can select a Codeset from the list below and see it displayed in the centre column. The Codesets that are already in your review are displayed in the third column.

Available CodeSets:

- Risk Of Bias (Cochrane)
- Data extraction (Home Office review guidelines)
- Screening
- Allocations and Admin
- Data Extraction
- EEF Toolkit main data extraction v 1.0 June 2019
- EEF Toolkit effect size data extraction v 1.0 June 20
- NICE Quality appraisal checklist - qualitative studies
- NICE Quality Appraisal Checklist – quantitative inter
- CASP: 10 questions to help you make sense of qual
- AMSTAR
- Risk of Bias - AMSTAR2
- QUADAS 2
- ROBIS: RoB in Systematic Reviews
- Countries
- Screening tool
- EGM Mapping tool template
- New Prioritised Codeset
- RobotReviewer classifications

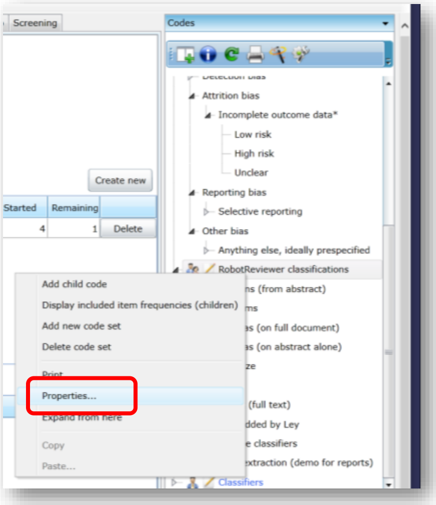
Code
Copy this CodeSet!
Exit
<- Back
Status: Normal Last code update: 16/11/2020 | User: Lev Muller

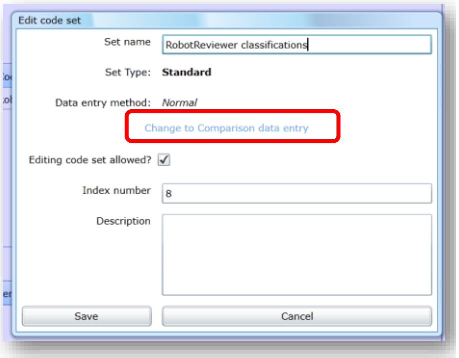
5

3. Change RobotReviewer codeset to «comparison» type

so that each researcher's asessments are tracked but not immediately visible to others

Right click on codeset → Properties




Edit code set

Set name: RobotReviewer classifications

Set Type: Standard

Data entry method: Normal

Editing code set allowed? ☑

Index number: 8

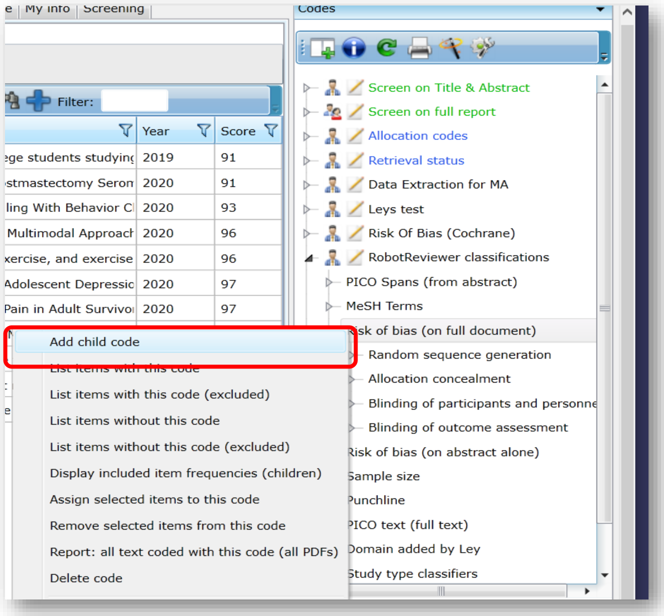
Description:

6

RobotReviewer only completes the first 4 domains, so you need to add the rest.

4. Add remaining 3 domains to the ***Risk of bias (on full document)*** codeset.

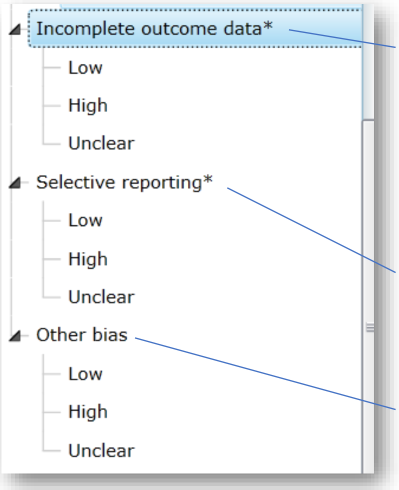
Right click on this code → Add child code.


e | My info | Screening
Filter:
Codes
Screen on Title & Abstract
Screen on full report
Allocation codes
Retrieval status
Data Extraction for MA
Leys test
Risk Of Bias (Cochrane)
RobotReviewer classifications
PICO Spans (from abstract)
MeSH Terms
isk of bias (on full document)
Random sequence generation
Allocation concealment
Blinding of participants and personne
Blinding of outcome assessment
Risk of bias (on abstract alone)
Sample size
Punchline
PICO text (full text)
Domain added by Ley
Study type classifiers
Add child code
List items with this code
List items with this code (excluded)
List items without this code
List items without this code (excluded)
Display included item frequencies (children)
Assign selected items to this code
Remove selected items from this code
Report: all text coded with this code (all PDFs)
Delete code

	Year	Score
ege students studyin	2019	91
stmastectomy Seron	2020	91
ling With Behavior C	2020	93
Multimodal Approach	2020	96
xercise, and exercise	2020	96
Adolescent Depressic	2020	97
Pain in Adult Survivo	2020	97

Add Cochrane's instructions to each code description:

This is how your 3 new domains should look:



Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition or exclusions where reported, and any reinclusions in analyses for the review

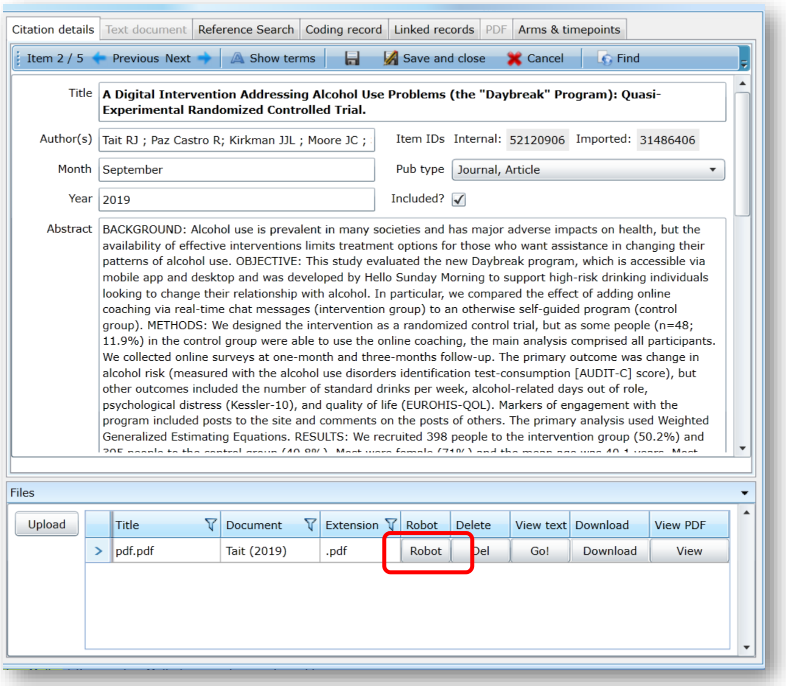
*Assessments should be made for each main outcome or class of outcomes.

State how selective outcome reporting was examined and what was found.
*Assessments should be made for each main outcome or class of outcomes.

State any important concerns about bias not covered in the other domains in the tool

5. Run each pdf through RobotReviewer

If you get an error message at the end, just click through it.



Title: A Digital Intervention Addressing Alcohol Use Problems (the "Daybreak" Program): Quasi-Experimental Randomized Controlled Trial.

Author(s): Tait RJ ; Paz Castro R; Kirkman JJL ; Moore JC ;

Item IDs Internal: 52120906 Imported: 31486406

Month: September

Pub type: Journal, Article

Year: 2019

Included? ☑

Abstract: BACKGROUND: Alcohol use is prevalent in many societies and has major adverse impacts on health, but the availability of effective interventions limits treatment options for those who want assistance in changing their patterns of alcohol use. OBJECTIVE: This study evaluated the new Daybreak program, which is accessible via mobile app and desktop and was developed by Hello Sunday Morning to support high-risk drinking individuals looking to change their relationship with alcohol. In particular, we compared the effect of adding online coaching via real-time chat messages (intervention group) to an otherwise self-guided program (control group). METHODS: We designed the intervention as a randomized control trial, but as some people (n=48; 11.9%) in the control group were able to use the online coaching, the main analysis comprised all participants. We collected online surveys at one-month and three-months follow-up. The primary outcome was change in alcohol risk (measured with the alcohol use disorders identification test-consumption [AUDIT-C] score), but other outcomes included the number of standard drinks per week, alcohol-related days out of role, psychological distress (Kessler-10), and quality of life (EUROHIS-QOL). Markers of engagement with the program included posts to the site and comments on the posts of others. The primary analysis used Weighted Generalized Estimating Equations. RESULTS: We recruited 398 people to the intervention group (50.2%) and 395 people to the control group (49.8%). Most were female (71%) and the mean age was 40.1 years. Most

	Title	Document	Extension	Robot	Delete	View text	Download	View PDF
>	pdf.pdf	Tait (2019)	.pdf	Robot	Del	Go!	Download	View

6. Allocate to your team members as appropriate, making it clear who is the primary researcher who fills out the entire form and who is checking their work.

Send them the instructions for team members document.

Schedule a 1.5-2 hours meeting with your team and someone from the machine learning team, to train and begin assessing together.

Risk of Bias assessments with machine learning

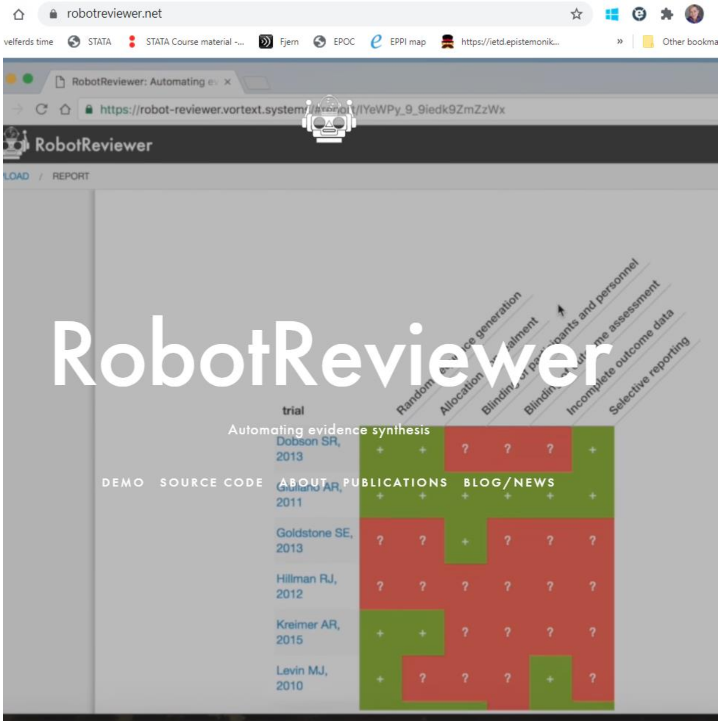
In EPPI-Reviewer

Instructions for ***team members***

1

Technology

- www.robotreviewer.net (drag and drop a pdf of an RCT to see whathappens)
- EPPI Reviewer has RR's technology built it, so researchers can skip the website.
- RR completes the first 4 of 7 domains in Cochrane's Risk of Bias. Developers suggest using RR as a support, not as an independent researcher.
- What is potentially even more helpful, is that it provides the text it used to make each judgement. That text by itself can be used by researchers.

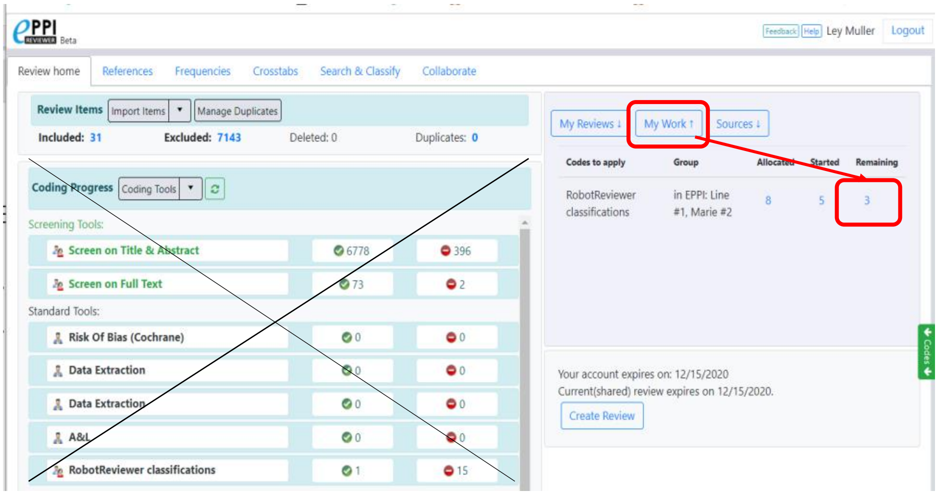

robotreviewer.net
velferds time
STATA
STATA Course material -...
Fjern
EPOC
EPPI map
https://ietd.epistemonik...
Other bookma
RobotReviewer: Automating e
https://robot-reviewer.vortext.systems/#report/IYeWPy_9_9iedk9ZmZzWx
RobotReviewer
LOAD / REPORT
RobotReviewer
Automating evidence synthesis
DEMO
SOURCE CODE
ABOUT
PUBLICATIONS
BLOG/NEWS

trial	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting
Dobson SR, 2013	+	+	?	?	?	+
Giuliano AR, 2011	+	+	+	+	+	+
Goldstone SE, 2013	?	?	+	?	?	?
Hillman RJ, 2012	?	?	?	?	?	?
Kreimer AR, 2015	+	+	?	?	?	?
Levin MJ, 2010	+	?	?	?	+	?

2

- Web version in Chrome, Firefox, Edge, Safari https://eppi.ioe.ac.uk/eppireviewer-web/
- Find your assignment

- Are you using version 4? [Skip to those instructions]()



ePPI REVIEWER Beta

Included: 31 Excluded: 7143 Deleted: 0 Duplicates: 0

Screening Tools:

Screen on Title & Abstract	✓ 6778	● 396
Screen on Full Text	✓ 73	● 2

Standard Tools:

Risk Of Bias (Cochrane)	✓ 0	● 0
Data Extraction	✓ 0	● 0
Data Extraction	✓ 0	● 0
A&L	✓ 0	● 0
RobotReviewer classifications	✓ 1	● 15

Codes to apply	Group	Allocated	Started	Remaining
RobotReviewer classifications	in EPPI: Line #1, Marie #2	8	5	3

Your account expires on: 12/15/2020
Current(shared) review expires on 12/15/2020.

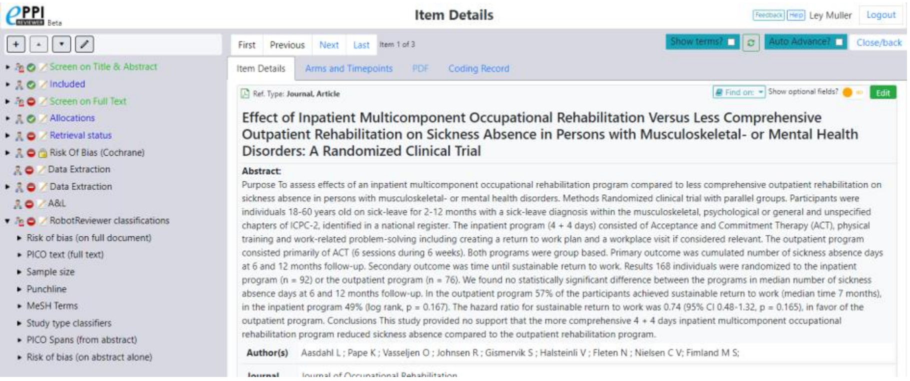
3

How has your project leader told you to assess RoB?

- Blinded to your other team members (but not blinded to machine learning)
  - Use [slides 5-9]()
- Not blinded to your other team members
  - Use [slides 10-12]()

4

Blinded to your team members



Ref. Type: **Journal, Article**

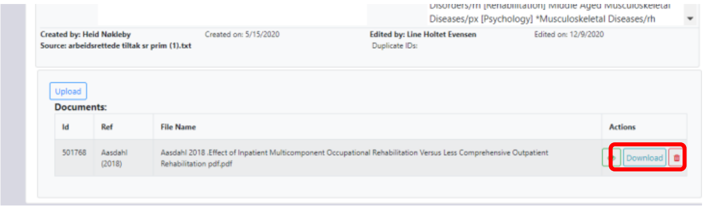
Effect of Inpatient Multicomponent Occupational Rehabilitation Versus Less Comprehensive Outpatient Rehabilitation on Sickness Absence in Persons with Musculoskeletal- or Mental Health Disorders: A Randomized Clinical Trial

Abstract:
Purpose To assess effects of an inpatient multicomponent occupational rehabilitation program compared to less comprehensive outpatient rehabilitation on sickness absence in persons with musculoskeletal- or mental health disorders. Methods Randomized clinical trial with parallel groups. Participants were individuals 18-60 years old on sick-leave for 2-12 months with a sick-leave diagnosis within the musculoskeletal, psychological or general and unspecified chapters of ICPC-2, identified in a national register. The inpatient program (4 + 4 days) consisted of Acceptance and Commitment Therapy (ACT), physical training and work-related problem-solving including creating a return to work plan and a workplace visit if considered relevant. The outpatient program consisted primarily of ACT (6 sessions during 6 weeks). Both programs were group based. Primary outcome was cumulated number of sickness absence days at 6 and 12 months follow-up. Secondary outcome was time until sustainable return to work. Results 168 individuals were randomized to the inpatient program (n = 92) or the outpatient program (n = 76). We found no statistically significant difference between the programs in median number of sickness absence days at 6 and 12 months follow-up. In the outpatient program 57% of the participants achieved sustainable return to work (median time 7 months), in the inpatient program 49% (log rank, p = 0.167). The hazard ratio for sustainable return to work was 0.74 (95% CI 0.48-1.32, p = 0.165), in favor of the outpatient program. Conclusions This study provided no support that the more comprehensive 4 + 4 days inpatient multicomponent occupational rehabilitation program reduced sickness absence compared to the outpatient rehabilitation program.

Author(s) Aasdahl L ; Pape K ; Vasseljen O ; Johnsen R ; Gismervik S ; Halsteinli V ; Fleten N ; Nielsen C V; Fimland M S;

Journal Journal of Occupational Rehabilitation

- First, download the pdf and move it to a different window



Disorders/rh [Rehabilitation] Middle Aged Musculoskeletal Diseases/px [Psychology] *Musculoskeletal Diseases/rh

Created by: Heid Nøkleby Created on: 5/15/2020 **Edited by: Line Holtet Evensen** Edited on: 12/9/2020
Source: arbeidsrettede tiltak sr prim (1).txt Duplicate IDs:

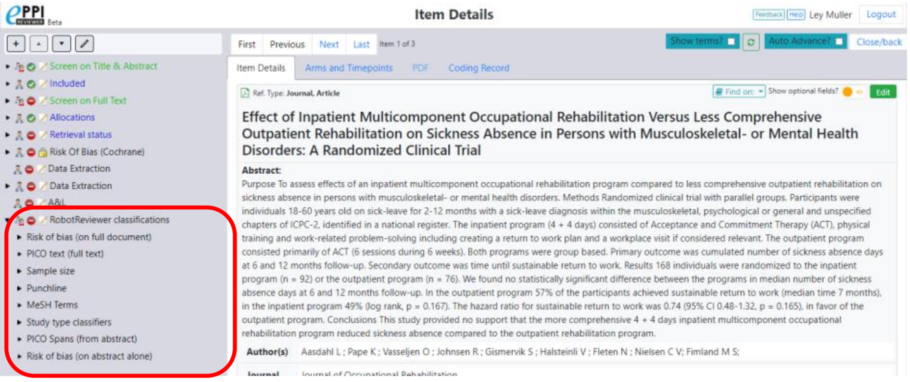
Documents:

Id	Ref	File Name	Actions
501768	Aasdahl (2018)	Aasdahl 2018 .Effect of Inpatient Multicomponent Occupational Rehabilitation Versus Less Comprehensive Outpatient Rehabilitation pdf.pdf	Download

5

Option 1: Blinded to your team members

- Codeset you are interested in: RobotReviewer classifications



Ref. Type: **Journal, Article**

Effect of Inpatient Multicomponent Occupational Rehabilitation Versus Less Comprehensive Outpatient Rehabilitation on Sickness Absence in Persons with Musculoskeletal- or Mental Health Disorders: A Randomized Clinical Trial

Abstract:
Purpose To assess effects of an inpatient multicomponent occupational rehabilitation program compared to less comprehensive outpatient rehabilitation on sickness absence in persons with musculoskeletal- or mental health disorders. Methods Randomized clinical trial with parallel groups. Participants were individuals 18-60 years old on sick-leave for 2-12 months with a sick-leave diagnosis within the musculoskeletal, psychological or general and unspecified chapters of ICPC-2, identified in a national register. The inpatient program (4 + 4 days) consisted of Acceptance and Commitment Therapy (ACT), physical training and work-related problem-solving including creating a return to work plan and a workplace visit if considered relevant. The outpatient program consisted primarily of ACT (6 sessions during 6 weeks). Both programs were group based. Primary outcome was cumulated number of sickness absence days at 6 and 12 months follow-up. Secondary outcome was time until sustainable return to work. Results 168 individuals were randomized to the inpatient program (n = 92) or the outpatient program (n = 76). We found no statistically significant difference between the programs in median number of sickness absence days at 6 and 12 months follow-up. In the outpatient program 57% of the participants achieved sustainable return to work (median time 7 months), in the inpatient program 49% (log rank, p = 0.167). The hazard ratio for sustainable return to work was 0.74 (95% CI 0.48-1.32, p = 0.165), in favor of the outpatient program. Conclusions This study provided no support that the more comprehensive 4 + 4 days inpatient multicomponent occupational rehabilitation program reduced sickness absence compared to the outpatient rehabilitation program.

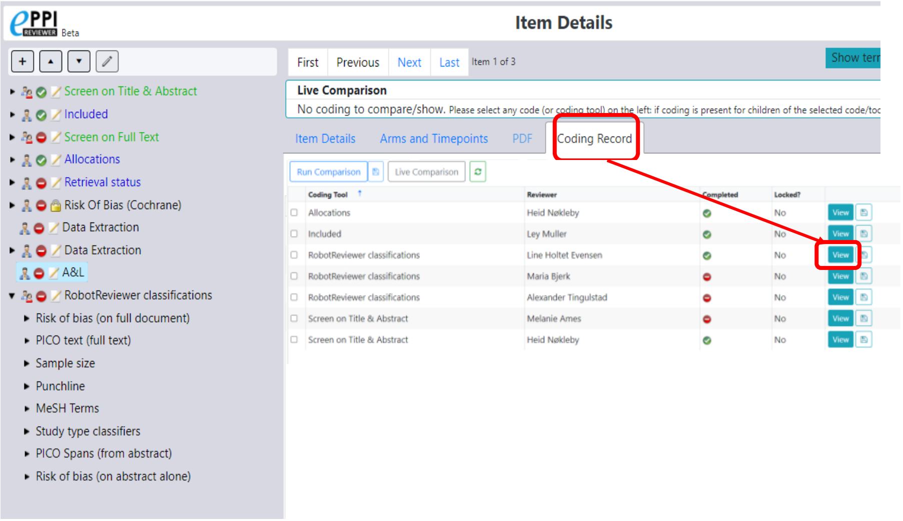
Author(s) Aasdahl L ; Pape K ; Vasseljen O ; Johnsen R ; Gismervik S ; Halsteinli V ; Fleten N ; Nielsen C V; Fimland M S;

Journal Journal of Occupational Rehabilitation

6

Option 1: Blinded to your team members

- Turn on live comparisons to see machine learning assessments (this breaks blinding):
- Coding record → View the person whose codes represent machine learning (your project leader will tell you).



Item Details

+ ▲ ▼

- Screen on Title & Abstract
- Included
- Screen on Full Text
- Allocations
- Retrieval status
- Risk Of Bias (Cochrane)
- Data Extraction
- Data Extraction
- A&L
- RobotReviewer classifications
  - Risk of bias (on full document)
  - PICO text (full text)
  - Sample size
  - Punchline
  - MeSH Terms
  - Study type classifiers
  - PICO Spans (from abstract)
  - Risk of bias (on abstract alone)

First Previous Next Last Item 1 of 3

Live Comparison

No coding to compare/show. Please select any code (or coding tool) on the left: if coding is present for children of the selected code/toc

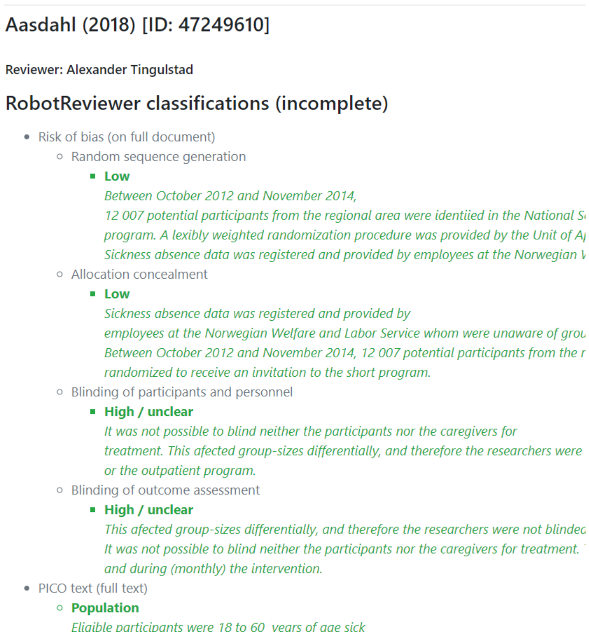
Item Details | Arms and Timepoints | PDF | Coding Record

Run Comparison | Live Comparison

	Coding Tool	Reviewer	Completed	Locked?	
□	Allocations	Heid Nøkleby	✓	No	View
□	Included	Ley Muller	✓	No	View
□	RobotReviewer classifications	Line Holtet Evensen	✓	No	View
□	RobotReviewer classifications	Maria Bjerk	⊖	No	View
□	RobotReviewer classifications	Alexander Tingulstad	⊖	No	View
□	Screen on Title & Abstract	Melanie Ames	⊖	No	View
□	Screen on Title & Abstract	Heid Nøkleby	✓	No	View

Option 1: Blinded to your team members

- A new window will pop up displaying the automated risk of bias assessments.
- Any text extracted will be in italics



Aasdahl (2018) [ID: 47249610]

Reviewer: Alexander Tingulstad

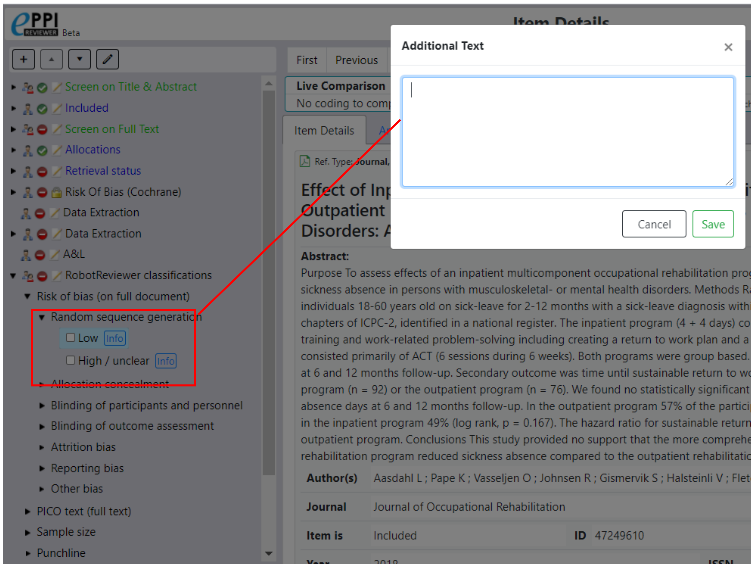
RobotReviewer classifications (incomplete)

- Risk of bias (on full document)
  - Random sequence generation
    - **Low**
      Between October 2012 and November 2014,
      12 007 potential participants from the regional area were identiied in the National S
      program. A lexibly weighted randomization procedure was provided by the Unit of A
      Sickness absence data was registered and provided by employees at the Norwegian V
  - Allocation concealment
    - **Low**
      Sickness absence data was registered and provided by
      employees at the Norwegian Welfare and Labor Service whom were unaware of grou
      Between October 2012 and November 2014, 12 007 potential participants from the r
      randomized to receive an invitation to the short program.
  - Blinding of participants and personnel
    - **High / unclear**
      It was not possible to blind neither the participants nor the caregivers for
      treatment. This afected group-sizes differentially, and therefore the researchers were
      or the outpatient program.
  - Blinding of outcome assessment
    - **High / unclear**
      This afected group-sizes differentially, and therefore the researchers were not blinded
      It was not possible to blind neither the participants nor the caregivers for treatment.
      and during (monthly) the intervention.
- PICO text (full text)
  - **Population**
    Eligible participants were 18 to 60 years of age sick

Option 1: Blinded to your team members

Your assignments

1. Fill out all 7 domains in ***Risk of bias (on full document)***
   a) Check the correct code (Low or High/unclear)
   b) Click on **Info** and add in support for your assessment. Copy the text extracted by machine learning, if you agree, otherwise copy from the pdf, or write in your own text. Specify «high» vs «unclear» in the info box.



No coding to comp

Ref. Type: Journal,

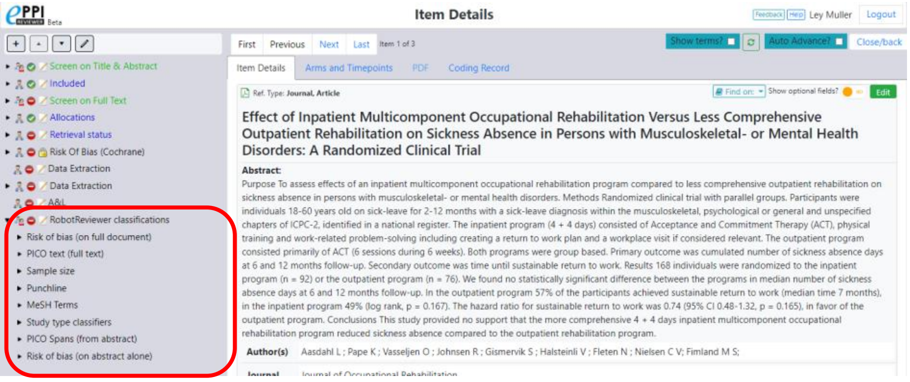
Effect of Inp
Outpatient
Disorders: A

Abstract:
Purpose To assess effects of an inpatient multicomponent occupational rehabilitation prog
sickness absence in persons with musculoskeletal- or mental health disorders. Methods R
individuals 18-60 years old on sick-leave for 2-12 months with a sick-leave diagnosis with
chapters of ICPC-2, identified in a national register. The inpatient program (4 + 4 days) co
training and work-related problem-solving including creating a return to work plan and a
consisted primarily of ACT (6 sessions during 6 weeks). Both programs were group based.
at 6 and 12 months follow-up. Secondary outcome was time until sustainable return to wo
program (n = 92) or the outpatient program (n = 76). We found no statistically significant
absence days at 6 and 12 months follow-up. In the outpatient program 57% of the partici
in the inpatient program 49% (log rank, p = 0.167). The hazard ratio for sustainable return
outpatient program. Conclusions This study provided no support that the more comprehe
rehabilitation program reduced sickness absence compared to the outpatient rehabilitatic

Author(s)	Aasdahl L ; Pape K ; Vasseljen O ; Johnsen R ; Gismervik S ; Halsteinli V ; Flet		
Journal	Journal of Occupational Rehabilitation		
Item is	Included	**ID**	47249610

Option 2: Not blinded to your team members

- Codeset you are interested in: RobotReviewer classifications
- Open pdf: Download



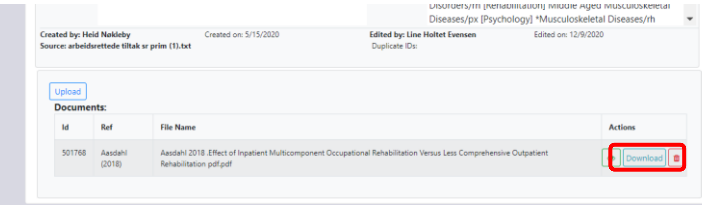
Ref. Type: **Journal, Article**

Effect of Inpatient Multicomponent Occupational Rehabilitation Versus Less Comprehensive Outpatient Rehabilitation on Sickness Absence in Persons with Musculoskeletal- or Mental Health Disorders: A Randomized Clinical Trial

Abstract:
Purpose To assess effects of an inpatient multicomponent occupational rehabilitation program compared to less comprehensive outpatient rehabilitation on sickness absence in persons with musculoskeletal- or mental health disorders. Methods Randomized clinical trial with parallel groups. Participants were individuals 18-60 years old on sick-leave for 2-12 months with a sick-leave diagnosis within the musculoskeletal, psychological or general and unspecified chapters of ICPC-2, identified in a national register. The inpatient program (4 + 4 days) consisted of Acceptance and Commitment Therapy (ACT), physical training and work-related problem-solving including creating a return to work plan and a workplace visit if considered relevant. The outpatient program consisted primarily of ACT (6 sessions during 6 weeks). Both programs were group based. Primary outcome was cumulated number of sickness absence days at 6 and 12 months follow-up. Secondary outcome was time until sustainable return to work. Results 168 individuals were randomized to the inpatient program (n = 92) or the outpatient program (n = 76). We found no statistically significant difference between the programs in median number of sickness absence days at 6 and 12 months follow-up. In the outpatient program 57% of the participants achieved sustainable return to work (median time 7 months), in the inpatient program 49% (log rank, p = 0.167). The hazard ratio for sustainable return to work was 0.74 (95% CI 0.48-1.32, p = 0.165), in favor of the outpatient program. Conclusions This study provided no support that the more comprehensive 4 + 4 days inpatient multicomponent occupational rehabilitation program reduced sickness absence compared to the outpatient rehabilitation program.

Author(s) Aasdahl L ; Pape K ; Vasseljen O ; Johnsen R ; Gismervik S ; Halsteinli V ; Fleten N ; Nielsen C V; Fimland M S;

Journal Journal of Occupational Rehabilitation



Disorders/rh [Rehabilitation] Middle Aged Musculoskeletal Diseases/px [Psychology] *Musculoskeletal Diseases/rh

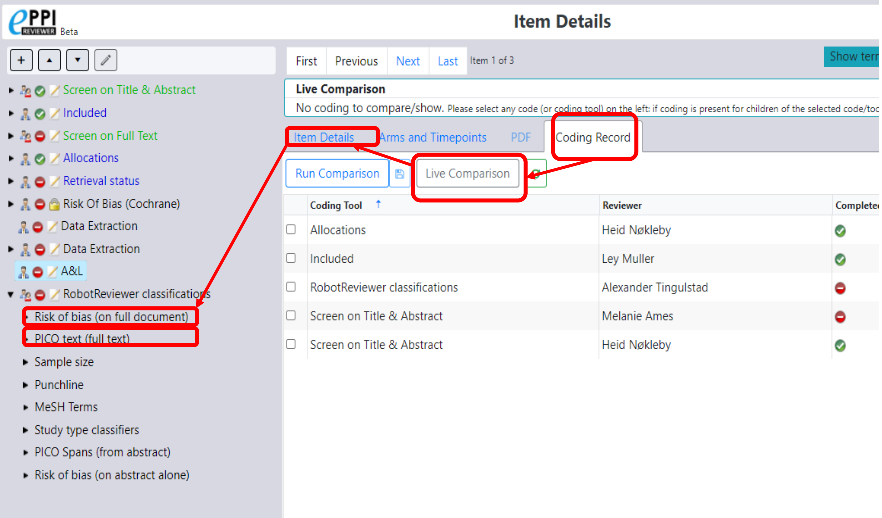
Created by: Heid Nøkleby Created on: 5/15/2020 **Edited by: Line Holtet Evensen** Edited on: 12/9/2020
Source: arbeidsrettede tiltak sr prim (1).txt Duplicate IDs:

Documents:

Id	Ref	File Name	Actions
501768	Aasdahl (2018)	Aasdahl 2018 .Effect of Inpatient Multicomponent Occupational Rehabilitation Versus Less Comprehensive Outpatient Rehabilitation pdf.pdf	Download

Option 2: Not blinded to your team members

- Turn on live comparisons to see machine learning assessments (this breaks blinding):
- Coding record → Live comparison → Citation details → click on the specific code you want to see. The child-codes immediate subordinate will be shown, so you might have to use the arrows to expand a code.



Item Details

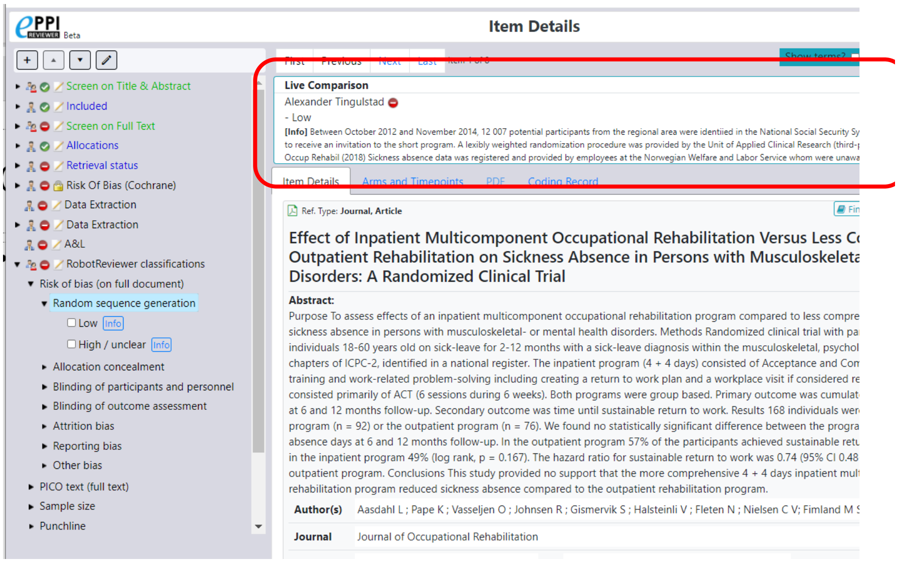
Live Comparison

No coding to compare/show. Please select any code (or coding tool) on the left: if coding is present for children of the selected code/too

	Coding Tool ↑	Reviewer	Complete
□	Allocations	Heid Nøkleby	✓
□	Included	Ley Muller	✓
□	RobotReviewer classifications	Alexander Tingulstad	⊖
□	Screen on Title & Abstract	Melanie Ames	⊖
□	Screen on Title & Abstract	Heid Nøkleby	✓

Option 2: Not blinded to your team members

- Take a look at the information already available.
- The code relevant to you is ***Risk of bias (on full document)***, while this can be helpful: ***PICO text (full-text)***. But the others also have interesting info.
- NB! You won't see any coding on the left side, because the assessment isn't completed yet. Look at the top of the screen for RR's coding, which will be marked under your team leader's name (or someone else on the machine learning team).
- Any text extracted will be displayed after **[Info]**



Item Details

Live Comparison

Alexander Tingulstad ⊖

- Low

[Info] Between October 2012 and November 2014, 12 007 potential participants from the regional area were identiied in the National Social Security Sy
to receive an invitation to the short program. A lexibly weighted randomization procedure was provided by the Unit of Applied Clinical Research (third-
Occup Rehabil (2018) Sickness absence data was registered and provided by employees at the Norwegian Welfare and Labor Service whom were unawa

Ref. Type: **Journal, Article**

**Effect of Inpatient Multicomponent Occupational Rehabilitation Versus Less Co
Outpatient Rehabilitation on Sickness Absence in Persons with Musculoskeleta
Disorders: A Randomized Clinical Trial**

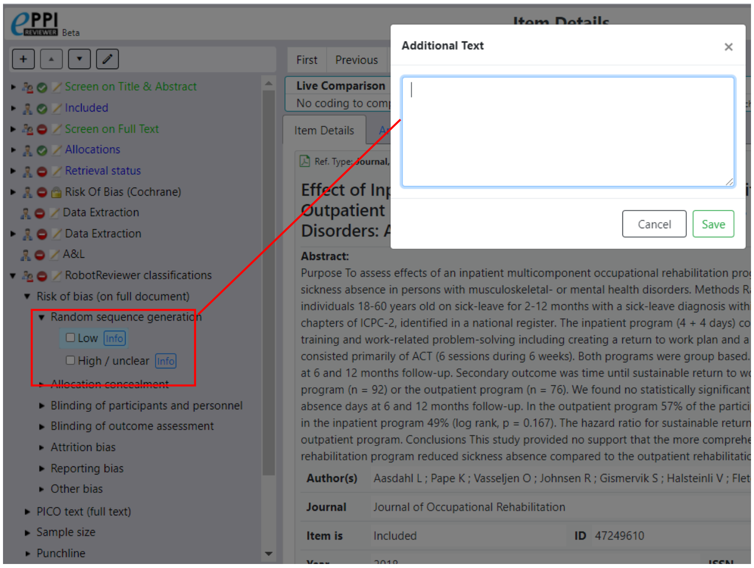
Abstract:
Purpose To assess effects of an inpatient multicomponent occupational rehabilitation program compared to less compre
sickness absence in persons with musculoskeletal- or mental health disorders. Methods Randomized clinical trial with pa
individuals 18-60 years old on sick-leave for 2-12 months with a sick-leave diagnosis within the musculoskeletal, psychol
chapters of ICPC-2, identified in a national register. The inpatient program (4 + 4 days) consisted of Acceptance and Con
training and work-related problem-solving including creating a return to work plan and a workplace visit if considered re
consisted primarily of ACT (6 sessions during 6 weeks). Both programs were group based. Primary outcome was cumulat
at 6 and 12 months follow-up. Secondary outcome was time until sustainable return to work. Results 168 individuals wer
program (n = 92) or the outpatient program (n = 76). We found no statistically significant difference between the progra
absence days at 6 and 12 months follow-up. In the outpatient program 57% of the participants achieved sustainable retu
in the inpatient program 49% (log rank, p = 0.167). The hazard ratio for sustainable return to work was 0.74 (95% CI 0.48
outpatient program. Conclusions This study provided no support that the more comprehensive 4 + 4 days inpatient mul
rehabilitation program reduced sickness absence compared to the outpatient rehabilitation program.

Author(s)	Aasdahl L ; Pape K ; Vasseljen O ; Johnsen R ; Gismervik S ; Halsteinli V ; Fleten N ; Nielsen C V; Fimland M S
Journal	Journal of Occupational Rehabilitation

Option 2: Not blinded to your team members

Your assignments

1. Fill out all 7 domains in ***Risk of bias (on full document)***
   a) Check the correct code (Low or High/unclear)
   b) Click on **Info** and add in support for your assessment. Copy the text extracted by machine learning, if you agree, otherwise copy from the pdf, or write in your own text. Specify «high» vs «unclear» in the info box.



- Screen on Title & Abstract
- Included
- Screen on Full Text
- Allocations
- Retrieval status
- Risk Of Bias (Cochrane)
- Data Extraction
- Data Extraction
- A&L
- RobotReviewer classifications
  - Risk of bias (on full document)
    - Random sequence generation
      - □ Low Info
      - □ High / unclear Info
    - Allocation concealment
    - Blinding of participants and personnel
    - Blinding of outcome assessment
    - Attrition bias
    - Reporting bias
    - Other bias
  - PICO text (full text)
  - Sample size
  - Punchline

Live Comparison
No coding to comp

Ref. Type: Journal,

Effect of Inp
Outpatient
Disorders: A

Abstract:
Purpose To assess effects of an inpatient multicomponent occupational rehabilitation pro
sickness absence in persons with musculoskeletal- or mental health disorders. Methods R
individuals 18-60 years old on sick-leave for 2-12 months with a sick-leave diagnosis with
chapters of ICPC-2, identified in a national register. The inpatient program (4 + 4 days) co
training and work-related problem-solving including creating a return to work plan and a
consisted primarily of ACT (6 sessions during 6 weeks). Both programs were group based.
at 6 and 12 months follow-up. Secondary outcome was time until sustainable return to w
program (n = 92) or the outpatient program (n = 76). We found no statistically significant
absence days at 6 and 12 months follow-up. In the outpatient program 57% of the partic
in the inpatient program 49% (log rank, p = 0.167). The hazard ratio for sustainable return
outpatient program. Conclusions This study provided no support that the more comprehe
rehabilitation program reduced sickness absence compared to the outpatient rehabilitati

Author(s)	Aasdahl L ; Pape K ; Vasseljen O ; Johnsen R ; Gismervik S ; Halsteinli V ; Flet		
Journal	Journal of Occupational Rehabilitation		
Item is	Included	**ID**	47249610

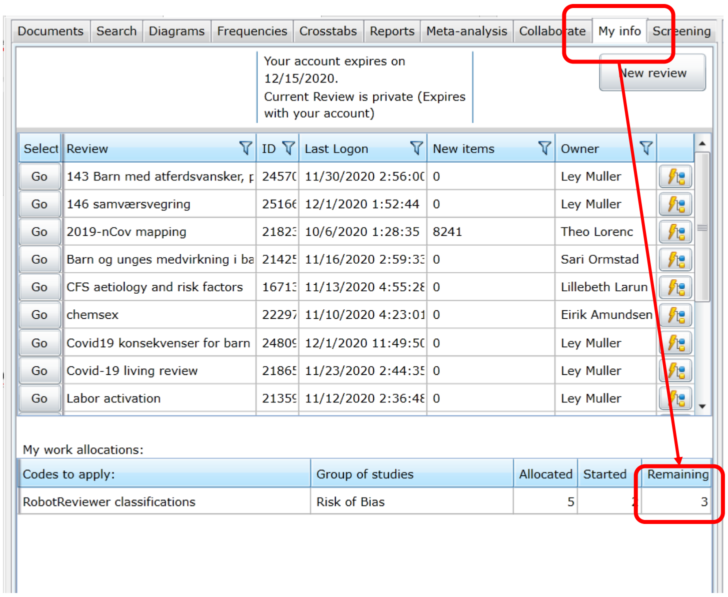
Additional Text ×

Cancel Save

13

EPPI version 4 Interface

- Version 4 in internet explorer: http://eppi.ioe.ac.uk/eppireviewer4/eppireviewer4.aspx
- Find your assignment



Your account expires on 12/15/2020.
Current Review is private (Expires with your account)

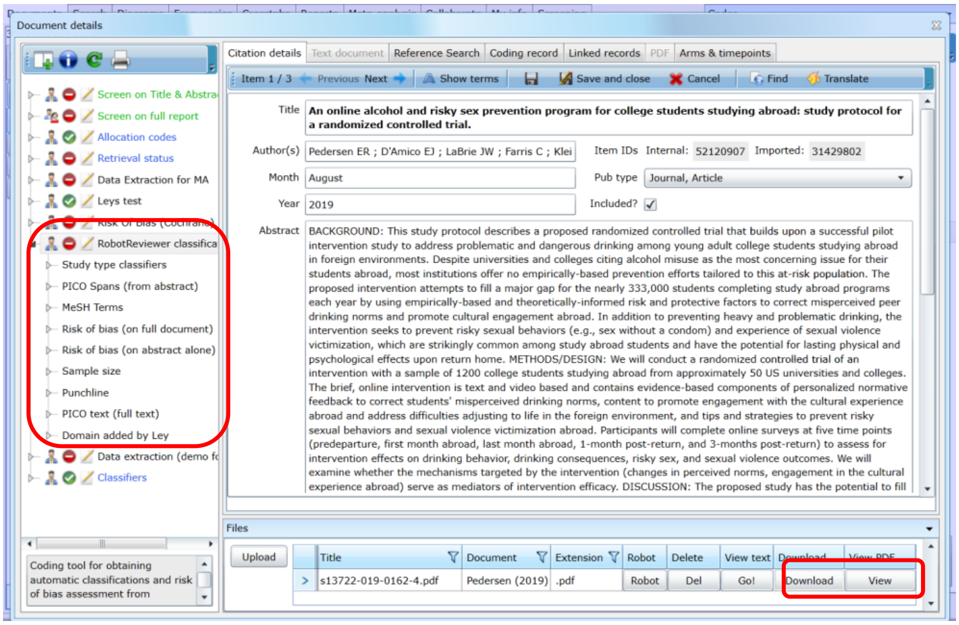
Select	Review	ID	Last Logon	New items	Owner
Go	143 Barn med atferdsvansker, p	2457(	11/30/2020 2:56:0(	0	Ley Muller
Go	146 samværsvegring	2516(	12/1/2020 1:52:44	0	Ley Muller
Go	2019-nCov mapping	2182:	10/6/2020 1:28:35	8241	Theo Lorenc
Go	Barn og unges medvirkning i ba	2142!	11/16/2020 2:59:3:	0	Sari Ormstad
Go	CFS aetiology and risk factors	1671:	11/13/2020 4:55:2ŧ	0	Lillebeth Larun
Go	chemsex	2229?	11/10/2020 4:23:01	0	Eirik Amundsen
Go	Covid19 konsekvenser for barn	2480?	12/1/2020 11:49:5(	0	Ley Muller
Go	Covid-19 living review	2186!	11/23/2020 2:44:3!	0	Ley Muller
Go	Labor activation	2135?	11/12/2020 2:36:4ŧ	0	Ley Muller

My work allocations:

Codes to apply:	Group of studies	Allocated	Started	Remaining
RobotReviewer classifications	Risk of Bias	5	2	3

14

- Codeset you are interested in: RobotReviewer classifications
- Open pdf: Download



Title: An online alcohol and risky sex prevention program for college students studying abroad: study protocol for a randomized controlled trial.

Author(s): Pedersen ER ; D'Amico EJ ; LaBrie JW ; Farris C ; Klei

Item IDs Internal: 52120907 Imported: 31429802

Month: August

Pub type: Journal, Article

Year: 2019

Included? ☑

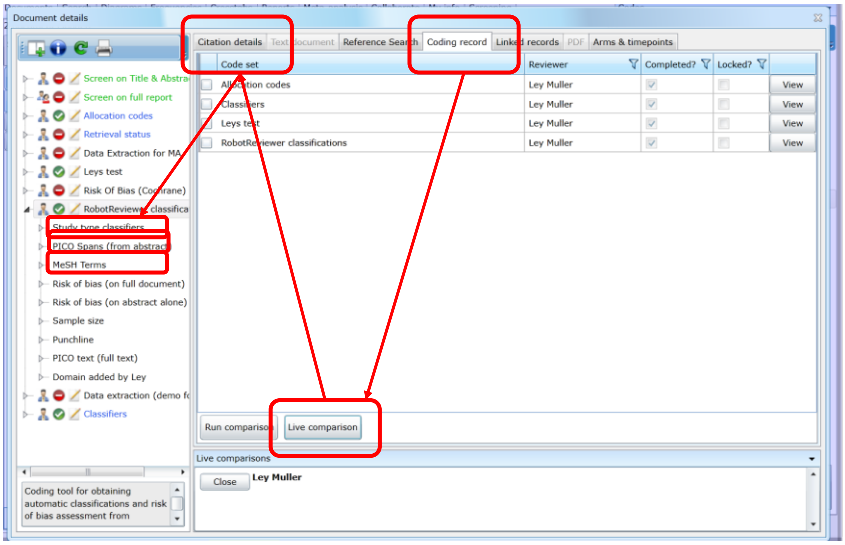
Abstract: BACKGROUND: This study protocol describes a proposed randomized controlled trial that builds upon a successful pilot intervention study to address problematic and dangerous drinking among young adult college students studying abroad in foreign environments. Despite universities and colleges citing alcohol misuse as the most concerning issue for their students abroad, most institutions offer no empirically-based prevention efforts tailored to this at-risk population. The proposed intervention attempts to fill a major gap for the nearly 333,000 students completing study abroad programs each year by using empirically-based and theoretically-informed risk and protective factors to correct misperceived peer drinking norms and promote cultural engagement abroad. In addition to preventing heavy and problematic drinking, the intervention seeks to prevent risky sexual behaviors (e.g., sex without a condom) and experience of sexual victimization, which are strikingly common among study abroad students and have the potential for lasting physical and psychological effects upon return home. METHODS/DESIGN: We will conduct a randomized controlled trial of an intervention with a sample of 1200 college students studying abroad from approximately 50 US universities and colleges. The brief, online intervention is text and video based and contains evidence-based components of personalized normative feedback to correct students' misperceived drinking norms, content to promote engagement with the cultural experience abroad and address difficulties adjusting to life in the foreign environment, and tips and strategies to prevent risky sexual behaviors and sexual violence victimization abroad. Participants will complete online surveys at five time points (predeparture, first month abroad, last month abroad, 1-month post-return, and 3-months post-return) to assess for intervention effects on drinking behavior, drinking consequences, risky sex, and sexual violence outcomes. We will examine whether the mechanisms targeted by the intervention (changes in perceived norms, engagement in the cultural experience abroad) serve as mediators of intervention efficacy. DISCUSSION: The proposed study has the potential to fill

Title	Document	Extension	Robot	Delete	View text	Download	View PDF
s13722-019-0162-4.pdf	Pedersen (2019)	.pdf	Robot	Del	Go!	Download	View

Coding tool for obtaining automatic classifications and risk of bias assessment from

15

- Turn on live comparisons to see machine learning assessments (this breaks blinding):
- Coding record → Live comparison → Citation details → click on the specific code you want to see. The child-codes immediately subordinate will be shown, so you might have to use the arrows to expand a code.



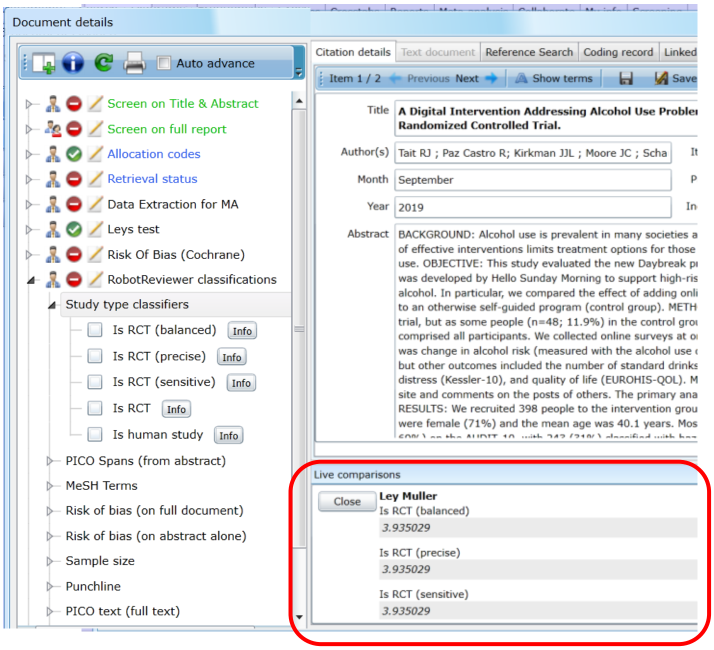
Code set	Reviewer	Completed?	Locked?	
Allocation codes	Ley Muller	☑	□	View
Classifiers	Ley Muller	☑	□	View
Leys test	Ley Muller	☑	□	View
RobotReviewer classifications	Ley Muller	☑	□	View

Ley Muller

Coding tool for obtaining automatic classifications and risk of bias assessment from

16

- Take a look at the information already available.
- The code relevant to you is ***Risk of bias (on full document)***, while this can be helpful: ***PICO text (full-text)***. But the others also have interesting info.
- NB! You won't see any coding on the left side, because the assessment isn't completed yet. Look at the bottom of the screen for RR's coding, which will be marked under your team leader's name (or someone else on the machine learning team).
- Any text extracted will be in italics.



Title **A Digital Intervention Addressing Alcohol Use Probler Randomized Controlled Trial.**

Author(s) Tait RJ ; Paz Castro R; Kirkman JJL ; Moore JC ; Scha

Month September

Year 2019

Abstract BACKGROUND: Alcohol use is prevalent in many societies a of effective interventions limits treatment options for those use. OBJECTIVE: This study evaluated the new Daybreak pi was developed by Hello Sunday Morning to support high-ris alcohol. In particular, we compared the effect of adding onli to an otherwise self-guided program (control group). METH trial, but as some people (n=48; 11.9%) in the control grou comprised all participants. We collected online surveys at oi was change in alcohol risk (measured with the alcohol use c but other outcomes included the number of standard drinks distress (Kessler-10), and quality of life (EUROHIS-QOL). M site and comments on the posts of others. The primary ana RESULTS: We recruited 398 people to the intervention grou were female (71%) and the mean age was 40.1 years. Mos

Live comparisons

Close

Ley Muller

Is RCT (balanced)

3.935029

Is RCT (precise)

3.935029

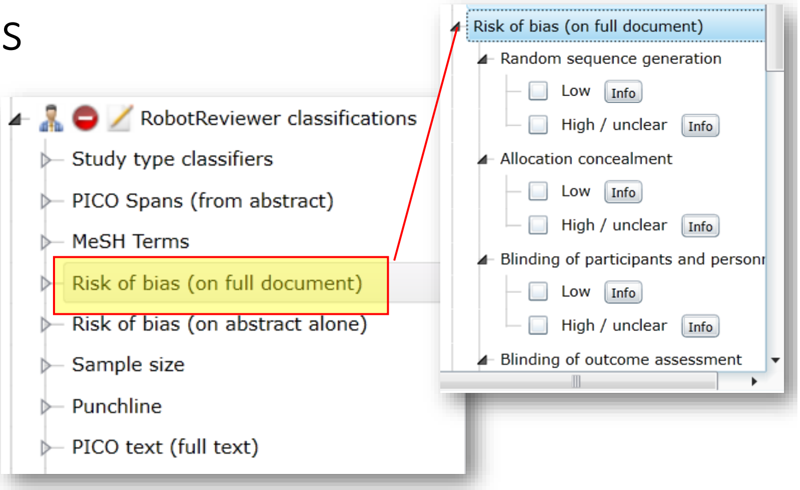
Is RCT (sensitive)

3.935029

Your assignments

1. Fill out all 7 domains in ***Risk of bias (on full document)***
   a) Check the correct code
   b) Click on Info and add in support for your assessment. Copy from RR, if you agree, otherwise copy from the pdf, or write in your own text. Specify «high» vs «unclear» in the info box.

.


RobotReviewer classifications
Study type classifiers
PICO Spans (from abstract)
MeSH Terms
Risk of bias (on full document)
Risk of bias (on abstract alone)
Sample size
Punchline
PICO text (full text)
Risk of bias (on full document)
Random sequence generation
Low
Info
High / unclear
Info
Allocation concealment
Low
Info
High / unclear
Info
Blinding of participants and personn
Low
Info
High / unclear
Info
Blinding of outcome assessment

Published by the Norwegian Institute of Public Health
September 2021
P.O.B 4404 Nydalen
NO-0403 Oslo
Phone: + 47-21 07 70 00
The report can be downloaded as pdf
at www.fhi.no/en/publ/