Study Aims
- To study the difference in reviewer workload (ie. citations viewed) between systematic reviews of treatment and diagnosis
- Identify sources of such a difference to provide a basis for future work improving the efficiency of literature searches for systematic reviews

Background
- Systematic reviews are a key component in the formulation and dissemination of modern evidence based medical practice
- Reviewers conduct a comprehensive survey and meta-analysis of the literature to answer specific clinical questions
- High cost of errors mean that systematic reviews are largely conducted manually and can take months or even years in practice
- Comprehensive databases (eg. MEDLINE) have been developed to index available medical literature
- Citations in MEDLINE are annotated with one or more medical concepts from the MeSH (MEDical Subject Headings) ontology to facilitate efficient searching

Motivation
- Traditionally systematic reviews have focused on questions related to medical treatments
- Recently there has been an increased demand for reviews of other clinical areas such as diagnostic test accuracy (DTA)
- DTA reviews face several additional challenges compared to reviews of treatment
  - Developing reporting standards for diagnostic test accuracy studies
  - MeSH headings are still being developed to adequately describe DTA concepts (eg. there is currently no heading for DTA study)
- All systematic reviews follow the same general procedure (see figure at top right)

Methodology
- A set of systematic reviews of treatment (n=200) and diagnosis (n=6) published by Cochrane is analysed and summaries of their literature searches were compared
  - The number of citations screened at the first and second stages as well as the number included in the final analysis were recorded

Screening on Full Text
- More citations are obtained in full-text for DTA reviews
- The number of citations rejected for each accepted study is higher for DTA reviews
  - ie. It is harder to judge the relevance of a citation for DTA review on title and abstract alone

Screening on Title and Abstract
- Initial database searches for DTA reviews return far more citations than for searches for treatment
  - ie. Literature searches for DTA reviews have lower precision in order to meet the requirement for near perfect recall

Results
- The number of citations screened at all stages is far greater for DTA reviews than for treatment leading to significant increases in reviewer workload
- Despite low power caused by lack of available DTA studies most results were significant
  - Significant differences were observed for each stage for either the raw number of citations screened, or the number of rejected studies for each accepted citation

Conclusions
- Reviewer workload (ie. citations viewed) is much larger for DTA reviews
- Improvements in reporting standards and quality of meta-data could reduce the difference at the second and first screening stages respectively
- Meanwhile (and as the relative difficulties are unlikely to ever converge completely), future work must focus on dealing with the greater workload inherent for DTA reviews