MEDLIOR
HEALTH OUTCOMES RESEARCH

Project CADTH Rapid Response Evaluation Report

Date 30 March 2015

CADTH Contact

Lesley Dunfield
Director, HTA and Rapid Response
Tel: 613.226.2553, ext. 1291
lesleyd@cadth.ca
www.cadth.ca

MEDLIOR Contact

Tara Cowling
Principal Consultant
Tel: 403.460.2616
Fax (1 888) 427 1669
tara.cowling@medlior.com
www.medlior.com

Jodie Worrall
Senior Consultant
Fax (1 888) 427 1669
jodie.worrall@medlior.com
www.medlior.com

1. Executive Summary

For decades, systematic reviews have been the gold standard in evidence synthesis; however, decision-makers and other knowledge users often require more expedient reports [1]. Rapid reviews seek to provide synthesized evidence utilizing efficiencies in methodological processes. CADTH's Rapid Response Program was established in 2005 and currently delivers over 200 reports per year. Now approaching its' 10th anniversary, an external evaluation of the Rapid Response Program was commissioned as part of an ongoing focus on continual improvement.

Ten reports were selected by CADTH for independent evaluation across three Rapid Response product levels; each level was associated with specific evaluative tasks whereby selected elements of the methods used by the original reviewer and reproduced by the auditor were compared. The auditors did not seek to reproduce CADTH templates and reports in full, rather to challenge reproducibility of key elements in terms of methodological transparency. Discrepancies were evaluated in terms of impact on interpretation, conclusions and program improvements.

Level 1 Rapid Response reports are reference lists of the best available evidence with abstracts and links to full-text documents, if available. The audit included three selected Level 1 reports and specifically addressed research question definition and study selection. For two of three reports, the topic refinement form (TRF) did not sufficiently define the research question or provide enough context for the missing detail to be inferred by the auditor leading to some differences in study selection. The third evaluation confirmed that when the TRF was sufficiently detailed, the reviewer's and auditor's definition of research questions and subsequent study selection were aligned.

Level 1.5 Rapid Response reports build on the Level 1 reports with an additional summary based on the abstracts of the best available evidence; the auditor's tasks included Level 1 tasks plus an evaluation of the interpretation and summary of overall findings provided. Again, two of three reports did not sufficiently define the research question or key terms; for one report (RB0654) this led to difficulties in completing study selection. Accounting for differences in study selection, the reviewer and auditor summations were mostly aligned except for RB0520 where differences in interpretation between the auditor and reviewer were more difficult to explain, particularly given the minor inconsistencies between tables and text.

Level 2 Rapid Response reports are written summaries of the evidence from full text articles, with critical appraisal and policy implications, the auditor's tasks included most Level 1 and Level 1.5 tasks plus a comparative evaluation of critical appraisal, data extraction, data presentation, conclusions and interpretation. Two out of four reports did not sufficiently define the research question, key terms and/or provided insufficient context for the auditor to infer missing information. Where the reviewer and auditor evaluated the same studies, the critical appraisal, data extraction and presentation plus conclusions and interpretation were aligned (again there was some variation in level of detail and style of presentation) but few, if any, meaningful differences. A critical learning was observed in the evaluation of RC044 where a seemingly undocumented post-hoc change to the TRF effectively changed the scope of the project, meaning the auditor was not able to match the reviewers study selection which generated subsequent 'downstream' differences.

The detailed instructions to authors provided by CADTH meant that when studies were commonly considered there were minimal differences between reviewer and auditor in specific tasks (critical appraisal, data extraction and presentation as well as conclusions and interpretation[1]). Whether or not the same studies were considered was a direct consequence of the level of detail provided in the research question. The need to sufficiently refine the research question is a common goal for systematic reviewers and rapid reviewers alike; however, as frequently observed herein, when combined with limited background information, what may appear sufficient and even detailed can become inadequate when applied in isolation.

The original reviewer was likely involved in query and discussion with at least the topic requestor and information scientist making it likely the reviewer's knowledge of the background context to the project exceeded that captured on the TRF. This becomes problematic when external auditor queries arise, as this knowledge cannot be used to infer missing information in the same way. Requirements for detailed and complete terminology (including synonyms) should be anticipated particularly when the published literature is likely to be inconsistent. Additional assumptions or caveats around key terms should also be anticipated (for example, how to treat a mixed adult/paediatric population, a blurred diagnostic definition, or a composite outcome).

The goal of this evaluative series was to challenge the reproducibility of key elements of published CADTH reports in terms of methodological quality and transparency. The auditor was not necessarily 'right' any more than the reviewer was necessarily 'wrong' (and vice versa); auditor-reviewer differences can be regarded as a signpost that more detail or additional information is needed (indeed within a systematic review these differences would be part of the process and may even involve a third reviewer as an adjudicator). In some instances minor differences could be regarded as artifactual and simply an inevitable consequence of two different reviewers acting independently.

Discrepancies considered in terms of impact on interpretation, and conclusions were ultimately found to be minimal. The majority of discrepancies that did arise, did so as a consequence of insufficient detail and definition from the outset (the research question). These concluding remarks should be reassuring but the importance (and challenge) associated with seemingly simple improvements in order to meet a goal of enhanced external reproducibility should perhaps not be underestimated.

[1] Some differences in level of detail and style of presentation were observed but this is to be expected

2 Table of Contents

Table of Tables

2. Background

Decision-makers and other knowledge users, however, often require expedient reports that methodically summarize available scientific evidence [1]. Rapid reviews seek to meet this requirement and provide synthesized evidence utilizing efficiencies in methodological processes. The increasing popularity and consequent growth of rapid reviews in healthcare decision-making is based upon these time- and resource-efficient processes; however, although no longer novel, rapid review remains poorly understood and represents an ill-defined set of disparate methodologies [2]. This lack of standardization demands a certain degree of caution when interpreting and utilizing the results. Although the number of rapid review "products" published annually continues to grow; quantity does not always equate with quality, and, without transparent methodology, the validity and appropriateness of these reviews often remains difficult to determine [3].

Rapid reviews often incorporate, with variable degree, the standardized methodology of systematic reviews but streamline aspects of the process to increase timeliness. The expediency of rapid reviews is therefore often dependent on how closely systematic review methodology is followed; for example, a modified approach to evidence identification, study selection, quality appraisal and/or evidence compilation is utilized to decrease time- and resource-burdens. The trade-off of such modification may be in terms of scope, level of detail, risk of bias and methods of synthesis. Furthermore, rapid reviews are often published in more reader-friendly formats by incorporating infographics or other data visualization techniques; however the usability of these formats is often associated with limitations in transparency and validity.

"…there is no universally accepted definition of what constitutes a rapid review."

CADTH introduced a Rapid Response Service in February 2005[2]. CADTH make the important distinction between the Rapid Response Service that was launched to support "*time-sensitive decisions*", versus comprehensive reviews required for "*important deliberations*". The topics for the CADTH Rapid Response Service are requested by healthcare decision-makers, to facilitate the "*appropriate and effective utilization of drugs and health technologies within health care systems across Canada*". The requests can be submitted through CADTH's online system or via a CADTH Liaison Officer.

There are several products within the CADTH Rapid Response Services [4], with variable timeframes depending on the methodology associated with the product. In order of increasing time and/or methodological rigor, the following products can be requested: reference lists, summary of abstracts, summary with critical appraisal, peer-review summary with critical appraisal, systematic review and meta-analysis and rapid health technology assessment. CADTH's Rapid Response Program is now approaching its 10th anniversary and delivers more than 200 reports per year, and is a highly valued service for CADTH customers. In light of the upcoming anniversary, an external evaluation of the program was requested to ensure continued improvement and customer needs.

[2] http://www.cadth.ca/en/products/rapid-response/rapid-response-service (accessed 25th March 2015)

3. Objectives

The objectives of the external validation were to:

- Assess and evaluate the methods used to prepare Rapid Response products (three Levels)
- Compare selected elements of methods to prepare each level of report conducted by the auditor and the original reviewer
- Assess discrepancies in terms of potential impact on the interpretation and conclusions of the report
- Describe recommendations for program improvements and communication of the product limitations

4. Methods

- Ten reports were selected by CADTH for independent evaluation across three Rapid Response product levels (Level 1, 1.5 and 2)
- Each level was associated with specific evaluative tasks
- The topic refinement form (TRF), a complete bibliography of search results as an unmarked list, and the final published CADTH report were provided to the auditors
- Grey literature search elements were not re-examined
- Selected elements of methods used by the original reviewer and reproduced by the auditor were compared
- Discrepancies were evaluated in terms of impact on interpretation, conclusions and program improvements

Ten Rapid Response reports were selected by CADTH for evaluation, each report covered a different research question(s) and different therapy areas (see Table 1 below).

Table 1 Rapid Response Reports selected for independent evaluation by CADTH

CADTH Product	Project code	Title
Level 1	RA0611	The Use of Medical Marijuana: Guidelines and Recommendations
	RA0627	Automation for the Preparation of Intravenous Solutions for Acute Care Patients: Cost-Effectiveness and Safety
	RA0677	Reprocessing of Single Use Medical Devices: Effectiveness, Safety, and Guidelines
Level 1.5	RB0520	Oncotype DX-Guided Treatment in Early Stage Breast Cancer: Cost-Effectiveness
	RB0654	Mobilization of Adult Inpatients in Hospitals or Long-Term/Chronic Care Facilities: Benefits and Harms, Safety, and Guidelines
	RB0721	Public Automated External Defibrillators and Cardiopulmonary Resuscitation Education: Clinical and Cost-Effectiveness
Level 2	L0161	Tinnitus Retraining Therapy: A Review of the Clinical Effectiveness
	L0227	Zoledronic Acid Intravenous Infusion: A Review of the Clinical Effectiveness and Guidelines
	RC0441	Rasburicase for Adults with Acute Tumor Lysis Syndrome: A Review of Clinical and Cost Effectiveness and Safety
	RC0570	Endovascular Thermal Ablation Technologies for Treatment of Varicose Veins: A Review of Clinical Effectiveness, Safety, Cost-Effectiveness and Guidelines – An Update

For each report CADTH provided the TRF, a complete bibliography of search results as an unmarked list, and the final published CADTH report. For Level 2, full papers for included studies were also provided.

Having reviewed the TRF, the auditor generated a new research question (including PICO[3] statement) and proceeded to select studies from the bibliography based on the study abstract alone. Following completion of this task, the research question and study selection between reviewer and auditor were compared. For Level 1.5 reports, the auditor also provided an interpretation and summary of commonly selected studies for comparison against the reviewers' interpretation and summary.

Level 2 reports are more detailed and methodologically closer to a traditional systematic review. In order to evaluate these additional elements, the auditor also undertook critical appraisal (using the AMSTAR [5], Downs and Black [6], or Drummond et al [7] checklists as appropriate) and data extraction for the studies by both the reviewer and the auditor. A summary of auditor tasks associated with each level of the Rapid Response products is provided in Table 2. Although rapid reviews typically include a search and evaluation of the grey literature, validation of study selection using grey literature[4] was not attempted as part of this evaluation. Studies from the grey literature included by the reviewer but not by the auditor were accounted for and are explained in full in the relevant report section.

Evaluating alignment and consistency for the different levels of CADTH Rapid Response reports was the focus for this project and any discrepancies between reviewer and independent auditor provided the opportunity for learning and methodological refinement. It is important to note that the auditors did not seek to reproduce CADTH templates and reports in full, rather to challenge reproducibility of key elements in terms of methodological transparency. This section describes results but maintains a clear focus on differences, discrepancies and potential learnings rather than reproducing information.

[3] PICO: Population, Intervention, Comparator and Outcomes
[4] Academic literature that is not formally published but can be accessed through searching other sources

Table 2 Description of auditor tasks for each Rapid Response project level

CADTH Product	Elements for assessment	Task description
Level 1	a. Research questions b. Study selection	a. Determine if research questions are clear, unambiguous, and comprehensive b. Determine if studies selected by auditor are in agreement with those selected by reviewer
Level 1.5	a. Research questions b. Study selection c. Interpretation d. Summary	a. Determine if research questions are clear, unambiguous, and comprehensive b. Determine if studies selected by auditor are in agreement with those selected by reviewer c. Determine if interpretation of auditor aligned with reviewer d. Determine if summary provided by auditor aligned with reviewer
Level 2	a. Research questions b. Study selection c. Critical appraisal d. Data extraction e. Synthesis of evidence f. Conclusions and implications for decision making	a. Determine if research questions are clear, unambiguous, and comprehensive b. Determine if studies selected by auditor are in agreement with those selected by reviewer c. Determine if auditor's critical appraisal is in agreement with reviewer d. Determine if data extraction by auditor is in agreement with reviewer e. Evaluate and determine if presentation of data are clear, missing or over-represented f. Determine if conclusions and interpretation are valid and appropriate.

5. Results: Level 1 (Reference lists)

- Three Level 1 Rapid Response reports were selected for evaluation: RA0611, RA0627, RA0677
- The evaluation of RA0611 highlighted the importance of sufficiently defining the research question particularly when the published literature is likely to be inconsistent
- The evaluation of RA0627 similarly highlighted the importance of sufficiently defining the research question
- The evaluation of RA0677 confirmed that when the TRF provided adequate information the reviewer's and auditor's definition of research questions and study selection were aligned.

Level 1 Rapid Response reports are reference lists of the best available evidence with abstracts and links to full-text documents, if available, completed within 5-10 business days. As described in Table 2, the audit of Level 1 reports addressed research question definition and study selection.

5.1. The Use of Medical Marijuana: Guidelines and Recommendations (RA0611)

This topic was requested as part of preparations for a multi-stakeholder meeting. The stated rationale for the request (indicated on the TRF) was to gather more information on medical marijuana with regards to patient indications.

Auditor Comments

Refining the research question

The TRF did not indicate whether the terms 'medical marijuana/cannabis' included or excluded constituent cannabinoids such as tetrahydrocannabinol (THC) and cannabidiol (CBD) and synthetic cannabinoids such as nabilone (Table 3). These terms were not always used consistently in the literature, making study selection on the basis of prospectively-defined interventions more difficult for some publications, for example Lynch et al 2011 [8] used the phrase *"Cannabinoids studied included smoked cannabis, oromucosal extracts of cannabis based medicine, nabilone, dronabinol and a novel THC analogue"*, Curtis et al 2009 [9] "*any cannabinoid preparation"*, Martin-Sanchez et al 2009 [10] *"any cannabis preparation"* and Wang et al 2008 [11] "*medical cannabinoids"*.

For the purposes of auditor study selection it was assumed that studies of cannabis in a plant-based form that was subsequently vaporized, smoked, consumed (extract or capsule), administered via oral spray or similar were to be included[5]. Studies of constituent or synthetic cannabinoids were excluded. Studies where the nature of the cannabis used remained unclear or unreported were also excluded.

[5] As indicated either by the description of pre-specified inclusion criteria (what the authors sought to include) or, where inclusion criteria were unhelpful, by descriptions of interventions given in results (what the authors found)

The overall significance of systematic reviews and meta-analyses was also uncertain. The research title indicated a focus on guidelines and recommendations; however, as the TRF required the original reviewer to select 'Health Technology Assessment (HTA)/Systematic review/Meta-analysis' it was unclear whether systematic reviews and meta-analyses were actually of interest (in so far as they may make clinical recommendations) or if this was an arbitrary consequence of selecting HTA reports (which typically make recommendations). For the purposes of auditor study selection it was assumed that systematic reviews and meta-analyses making clinical recommendations were relevant and included. A summary of the research question as defined by the original reviewer and the auditor is presented in Table 3.

Table 3 A comparison of PICO(D) as captured by original reviewer and the auditor for RA0611

PICO(D)	Original reviewer	Auditor	Comment
Population	Patients requiring marijuana for specific medical conditions	Patients requiring marijuana for specific medical conditions	-
Intervention	Medical marijuana/ cannabis	Cannabis in a plant-based form that was subsequently vaporized, smoked, consumed (extract or capsule), administered via oral spray or similar	Unclear (relevance of cannabis constituents and other synthetic cannabinoids)
Comparator	N/A	N/A	-
Outcomes	Guidelines and Recommendations (what patient populations should be receiving treatment, what kind of marijuana should be administered)	Any	-
(Design)	HTA/ Systematic review/Meta-analysis, Guidelines (all countries)[6]	HTA/ Systematic review/Meta-analysis, Guidelines (all countries)	Unclear (relevance of systematic review and meta-analysis given focus of research title on guidelines and recommendations)

Study selection

The final CADTH report confirmed that no HTA reports, systematic reviews, meta-analyses, or evidence-based guidelines were identified regarding the use of medical marijuana for specific medical conditions. Additional references for reader information are provided listing ten systematic reviews, two randomized controlled trials (RCTs), three reviews and two additional references. In contrast, the auditor identified three abstracts [10] [12] [13] that appeared to meet the inclusion criteria insofar as they described

[6] What patient populations should be receiving treatment, what kind of marijuana should be administered

systematic reviews of patients using cannabis (plant-based) for specific medical conditions (these systematic reviews are summarized in Table 4). All three conclude to a lesser or greater extent that cannabis cannot be recommended for routine use as the likely harms outweigh potential benefits.

There are two likely reasons for this discrepancy:

1. The reviewer and auditor used different applications of the terms 'medical marijuana/cannabis'
2. The reviewer and auditor used different interpretations of the relevance of systematic reviews and subsequent importance of clinical recommendations arising from systematic reviews (as opposed to HTA reports or guidelines); none of the studies selected by the auditor were guidelines making population-level recommendations, they were all systematic reviews making a clinical recommendation relating to a particular patient population.

Learnings

It is likely that, at the time, these issues were not apparent to the reviewers and through interaction with requestor, these points were clarified and the task at hand was clear. However, to an auditor with no other information, it is critically important to sufficiently refine the PICO question and to include information to assist the reviewer in interpreting the language likely to be found in the literature.

Table 4 Studies selected by the auditor for inclusion

Reference	PICO(D)	Details	Author conclusions
Richards et al Neuromodulators for pain management in rheumatoid arthritis. Cochrane Database Syst Rev. 2012;1:CD008921.	Population	Rheumatoid arthritis	*"One small, low quality trial assessed oromucosal cannabis against placebo and found a small, significant difference favouring cannabis in the verbal rating score 'pain at present' (MD -0.72, 95% CI -1.31 to -0.13) after five weeks… Until further research is available… oromucosal cannabis… [has a] more significant side effect profile … potential harms seem to outweigh any modest benefit achieved"*
	Intervention	Neuromodulators	
	Comparator	Another therapy	
	Outcomes	At least one clinically relevant	
	(Design)	Systematic review of RCTs	
Phillips et al. Pharmacological treatment of painful HIV-associated sensory neuropathy: a systematic review and meta-analysis of randomised controlled trials. PLoS ONE. 2010;5(12):e14433.	Population	HIV-associated sensory neuropathy	*"Interventions demonstrating greater efficacy than placebo were smoked cannabis NNT 3.38 95%CI(1.38 to 4.10)… smoked cannabis cannot be recommended as routine therapy."*
	Intervention	Pharmacological treatment	
	Comparator	Not specified	
	Outcomes	Not specified	

	(Design)	Systematic review of RCTs	
Martin-Sanchez E, Furukawa TA, Taylor J, Martin JL. Systematic review and meta-analysis of cannabis treatment for chronic pain. Pain Med. 2009 Nov;10(8):1353-68.	Population	Chronic pain	*"The efficacy analysis displayed a difference in standardized means in favor of the cannabis arm of -0.61 (-0.84 to -0.37)…[however]…Currently available evidence suggests that cannabis treatment is moderately efficacious for treatment of chronic pain, but beneficial effects may be partially (or completely) offset by potentially serious harms."*
	Intervention	Cannabis preparations	
	Comparator	Placebo	
	Outcomes	Not specified	
	(Design)	Systematic review of RCTs	

5.2. Automation for the Preparation of Intravenous Solutions for Acute Care Patients: Cost-Effectiveness and Safety (RA0627)

This topic was requested in order to help determine the potential return on investment for moving from a manual to an automated system for the preparation of intravenous (IV) solutions; in particular, is there evidence of reduced cost and improved safety in terms of medication error rates.

Auditor Comments

Refining the research question

The TRF did not sufficiently define the term IV solutions. IV solutions could refer to any drug in solution administered intravenously, or perhaps more typically, to 'IV solutions' (hereafter known as 'IV fluids' to avoid confusion), that are used to maintain fluid balance, replace fluid losses, and treat electrolyte imbalances. The TRF also lacked clarity around the intended scope of the term 'acute care'; in particular, the relevance of the surgical setting (for example the automation of IV anaesthesia) and the oncology subgroup.

For the purposes of auditor study selection it was assumed that the term 'IV solution' could refer to any treatment delivered intravenously, not only to 'IV fluids'. It was also assumed that acute care was defined as secondary care for short-term treatment of time-sensitive conditions such as severe injury, severe episode of illness, or during recovery from surgery (the intraoperative period was excluded)[7]. A summary of the research question as defined by the original reviewer and the auditor is presented in Table 5.

[7] Definition provided by Hirshon et al 2012, *"The term acute care encompasses a range of clinical health-care functions, including emergency medicine, trauma care, pre-hospital emergency care, acute care surgery, critical care, urgent care and short-term inpatient stabilization" Bulletin of the World Health Organization 2013;91:386-388* http://www.who.int/bulletin/volumes/91/5/12-112664/en/ accessed 12th Feb 2015

Table 5 A comparison of PICO(D) as captured by original reviewer and the auditor for RA0627

PICO(D)	Original reviewer	Auditor	Comment
Population	Any patient in acute care requiring IV solutions Subgroup: Oncology patients requiring IV solutions	Acute care patients requiring IV solutions	Unclear (the acute setting was not clarified, the significance of the oncology subgroup was not explained)
Intervention	Any automated (or robotic system) IV solution preparation system	Any automated/robotic IV solution preparation system	Unclear (IV solution was not defined sufficiently)
Comparator	Manual IV solution preparation	Manual IV solution preparation	-
Outcomes	Cost-effectiveness, medication error rates	Cost-effectiveness, medication error rates	-
(Design)	HTA/systematic review/meta-analysis; RCTs; NRS only if few other study types found; economic evaluations	HTA/systematic review/meta-analysis; RCTs; NRS only if few other study types found; economic evaluations	-

Study selection

The final CADTH report confirmed that no HTA reports, systematic reviews, meta-analyses, evidence-based guidelines, or economic evaluations were identified regarding the use of automation for the preparation of IV solutions for acute care patients. Two non-randomized studies were included with five additional references are provided (these additional references were not classified by study type, unlike the list provided for RA0611). The auditor also identified the same two non-randomized studies for inclusion [14] [15].

Learnings

Pragmatically, the majority of references could ultimately have been excluded for other reasons, most often for not reporting outcomes related to cost-effectiveness or medication error rates; however, it is unlikely the auditor would have identified the same two studies had different assumptions been made regarding the definition of ‘IV solutions’ and/or a different interpretation of ‘acute care’. Both of the included studies dealt with drugs/interventions as opposed to IV fluids, Seger et al 2012 [14] described automated preparation of antineoplastic and adjuvant drug preparation whilst Dehmel et al 2011 [15] described the automated production of solutions containing amiodarone, noradrenaline and hydrocortisone. It is critically important to sufficiently refine the PICO question and to include information to assist the reviewer in interpreting the language likely to be found in the literature.

5.3. Reprocessing of Single Use Medical Devices: Effectiveness, Safety, and Guidelines (RA0677)

A requestor was reviewing policy around reprocessing of single use medical devices (SUDs) and requested an updated search to the CADTH 2008 HTA report to determine whether any new literature has been published.

Auditor Comments

Refining the research question

The TRF adequately defined the research question and the auditor made no additional assumptions in conducting study selection. A summary of the research question as defined by the original reviewer and the auditor is presented in Table 6.

Table 6 A comparison of PICO(D) as captured by original reviewer and the auditor for RA0677

PICO(D)	Original reviewer	Auditor	Comment
Population	Any patient	Any patient	-
Intervention	Use of medical devices manufactured for and labelled as single-use that had undergone reprocessing by an institutional health care provider or by a third-party reprocessor, and the use of SUDs that had been previously opened but not used	Any SUD reprocessed for subsequent use or previously opened but not used	-
Comparator	One time use of SUDs	Single use	-
Outcomes	Clinical outcomes following use of reprocessed SUDs in humans (e.g. infection of patients, other identifiable adverse events occurring in patients, mortality, device damage or failure, and evidence of device contamination), guidelines and recommendations	Clinical outcomes following use of reprocessed SUDs in humans, guidelines and recommendations	-
(Design)	HTA/systematic review/meta-analysis; RCTs; NRS (case reports only if few other study types found); guidelines (any jurisdiction but particularly provincial)	HTA/systematic review/meta-analysis; RCTs; NRS (case reports only if few other study types found); guidelines	-

Study selection

The final CADTH report confirmed that one RCT, four non-randomized studies, and one evidence-based guideline were selected for inclusion regarding the safe and effective use of SUDs. The auditor identified these same studies plus one additional study [1]. Hailey et al described a systematic review of reuse of

medical devices marketed for single use only (this study was not listed under the additional studies provided in the substantial appendix of further information).[8]

Learnings

When the TRF provided adequate information the original reviewer's and auditor's definition of research questions (PICO(D)) and study selection were aligned. In this instance the auditor also included one additional study; the reviewer's original reason for excluding this systematic review is unknown.

[8] The appendix also listed multiple studies not found in the provided bibliography; it was assumed additional hand-searching and/or searching of the grey literature was also undertaken.

6. Results: Level 1.5 (Summary of abstracts)

- Three Level 1.5 Rapid Response reports were selected for evaluation: RB0520, RB0654, RB0721
- The evaluation of RB0520 again highlighted the importance of sufficiently defining the research question particularly when the published literature is likely to use different terminology; some differences in interpretation were more difficult to explain, particularly given the apparent inconsistency between tables and text in some of the reviewers comments
- Despite making additional assumptions during the evaluation of RB0654 the auditor was unable to exclude several studies due to ambiguity in the research question
- The evaluation of RB0721 illustrated that a well-defined the research question and useful background/contextual information led to minimal differences between reviewer and auditor

Level 1.5 Rapid Response reports provide a summary based on the abstracts of the best available evidence, including the abstracts and links to full-text documents, where available. As described in Table 2, the auditor's tasks associated Level 1.5 reports included Level 1 tasks plus an evaluation of the interpretation and summary of overall findings reported. There are specific instructions on the format and content of the "summary of overall findings" from the CADTH Summary of Abstracts Author Requirements, these are noted below.

- A Level 1.5 report will contain an overall summary of the findings that will include the relevant literature identified.
- This summary will be based on the abstracts only (ordering of full text articles is not necessary) and will only include those abstracts that fall into the categories of literature searched.
- If possible, limit the summary to one, concise paragraph.
- Summaries of studies should be written in the past tense; summaries of guidelines should be written in the present tense.
- Summarize the most pertinent outcomes followed by an overall summary statement. Specific results (numbers and percentages) and p values are not necessary.
- Include overall statement(s) regarding primary outcomes and general findings. When possible, combine study results (i.e.: "Five RCTs found XX, whereas three RCTs found YY")

6.1. Oncotype DX-Guided Treatment in Early Stage Breast Cancer: Cost-Effectiveness (RB0520)

The rationale for this Rapid Response report was to provide information to guide funding requests for Oncotype DX, a genotyping technology that is used to guide treatment for early stage breast cancer by predicting response to chemotherapy treatments by evaluating the available evidence.

Auditor Comments

Refining the research question

The TRF did not sufficiently clarify terminology around the intervention of interest. Oncotype DX is one of four genomic tests available for breast cancer.[9] Relying on the term 'Oncotype DX' or similar would lead to the exclusion of studies that do not state the branded name of the test in the title or abstract. For example, Hornberger 2012 [16] uses the phrase *"21-gene recurrence score"*, Vanderlann et al 2011 [17] use "*21-gene assay*", whilst Bacchi et al 2010 [18] use *"21-gene expression assay";* the term Oncotype-DX is not mentioned. The auditor deduced that 21-gene assay synonymous with Oncotype DX on the basis that the three other genomic tests available each assessed a different number of genes (Mammaprint tests 70 genes, Mammostrat tests five genes, Prosigna tests 58 genes). A summary of the research question as defined by the original reviewer and the auditor is presented in Table 7.

Table 7 A comparison of PICO(D) as captured by original reviewer and the auditor for RB0520

PICO(D)	Original reviewer	Auditor	Comment
Population	Patients with early stage breast cancer	Patients with early stage breast cancer	-
Intervention	Oncotype DX-guided treatment	Oncotype DX-guided treatment or 21-gene guided assay	Unclear (relevance of 21-gene assay not stated)
Comparator	Treatment guided by other predictive testing technologies Treatment guided by clinical judgement only	Treatment guided by other predictive testing technologies or clinical judgement only	-
Outcomes	Cost effectiveness	Cost effectiveness	-
(Design)	HTA/systematic review/meta-analysis; economic evaluations	HTA/systematic review/meta-analysis; economic evaluations	-

Study selection

The final CADTH report confirmed that one HTA, one systematic review, and twelve economic evaluations regarding the cost effectiveness of Oncotype DX-guided treatment in patients with early stage breast cancer were identified. The auditor identified one systematic review and nine economic evaluations. One reference included by the reviewer, was excluded by the auditor [19]; this study used the Oncotype DX test in both arms whilst the population under consideration was varied. The auditor concluded that this study did not incorporate the correct comparator (neither another testing technology nor clinical judgement) whereas the reviewer incorporated this study as an evaluation of usual care versus adapted care alongside the other test versus no test studies. The remaining three discrepant references

[9] MammaPrint, Mammostrat, and Prosigna are alternative genomic tests for breast cancer. (http://www.breastcancer.org/symptoms/testing/types/oncotype_dx) accessed 04/02/2015

referred to one HTA and two economic evaluations that did not appear on the search bibliography and were therefore attributed to grey literature searches and these differences were disregarded.

Summary and interpretation

The overall summary of findings prepared by the reviewer and the auditor are presented in Table 8. The auditor did not provide a summary table (as provided by the reviewer) as this was not specified in the auditor instructions. Accounting for differences in study selection, the reviewer and auditor summations were mostly aligned. Notable differences were level of detail in so far as the auditor had split studies by those that compared 'test versus no test' from those that compared 'test versus another test' on the basis that these represented very different evaluative scenarios. There were also differences in the treatment of two particular studies;

1. The auditor regarded Klang 2010 [20] as a study that confirmed the cost effectiveness of Oncotype DX-guided treatment whereas the reviewer reported that "Oncotype-DX was not cost effective" in the overall summary of findings but extracted the apparently contradictory conclusion "The authors concluded that Oncotype DX represented an affordable and effective approach for women with LN- and ER+ ESBC" in Table 1 of the final report.
2. The reviewer commented in the summary of findings that Retel et al [24] reported that Oncotype DX was not cost effective compared to MammaPrint (an alternative genomic test) presumably based upon the cost/QALY outcome whereas the auditor regarded these results as comparable considering the cost/LYG and based on the authors conclusions that, *"This comparison indicates that the performances of the 70-gene and the 21-gene based on reported studies are close. The 21-gene has the highest probability of being cost-effective when focusing on cost/LY, while focusing on cost/QALY, the 70-gene signature was most cost-effective"*. The reviewer presented similar comments in Table 1 of the final report.

Table 8 Overall summary of findings prepared by reviewer and auditor for RB0520

Original reviewer overall summary of findings	**Auditor overall summary of findings**
One health technology assessment [21], one systematic review [16], and twelve economic analyses [17] [18] [19] [20] [22] [23] [24] [25] [26] [27] [28] examined the cost-effectiveness of the Oncotype DX-guided treatment (also known as the 21-gene Oncotype DX recurrence score (RS) assay) in patients with early stage breast cancer (ESBC). The authors of about half of the studies concluded that Oncotype DX was cost-effective when compared to current practice [17] [18] [19] [21] [23] [26] [28]. Two studies found that Oncotype DX was not cost-effective [20] [22] and the authors in two other studies did not deem it cost-effective when compared with other gene expression profiling products [24] [25]. The abstract for the systematic review did not contain	One systematic review [16] and nine economic evaluations evaluated the cost:benefit proposition[10] for Oncotype DX-guided treatment in node positive [22] [17], node negative [24] [25] [26], [18], [20] [27] or both positive and negative ER+ EBC [23]. Seven studies compared Oncotype DX-guided treatment with standard care (a no test scenario); five studies concluded that the test was either cost effective or cost saving for node negative ER+ EBC patients [26] [17] [18] [20] [27], one study concluded that testing was cost effective for both node negative and node positive patients [23], and one study concluded that cost effectiveness was highly uncertain for the node positive population [22]. Two studies compared Oncotype DX to an alternative 70-

[10] The term cost effectiveness was deliberately avoided as a couple of the included evaluations were costing/cost minimization studies and not cost effectiveness evaluations

any detail. Details for each study are presented in Table 1.	gene test, one study reported comparable results [24] whilst one study reported that Oncotype DX was not the most cost effective [25]. The abstract for the systematic review [16] did not report any reportable detail.

Learnings

Inclusion of synonymous terms regarding the intervention of interest would have aided transparency; however once the grey literature references have been accounted for, the reviewer and auditor's selections varied around only one reference. The differences in interpretation are more difficult to explain, particularly given the apparent inconsistency in the reviewer's comments as specified above.

6.2. Mobilization of Adult Inpatients in Hospitals or Long-Term/Chronic Care Facilities: Benefits and Harms, Safety, and Guidelines (RB0654)

The requester was seeking current information on the mobilization of patients during hospital, long term, or chronic care facility residency. The population of greatest interest was frail or ill seniors but information relating to a general adult population, and particularly those who undergo acute admissions to hospitals, was also of interest.

Auditor Comments

Refining the research question

A summary of the research question as defined by the original reviewer and the auditor is presented in Table 9. The TRF lacked clarity around the intended scope of the terms 'acute care', 'chronic care', 'mobilization' and the relevance of the elderly population to the research question. The most obvious interpretation of the research question is to include studies of adult patients (of any age) admitted to any acute in-patient setting or living in any chronic-in-patient care setting; this interpretation covers a very broad spectrum of clinical scenarios and a large and heterogeneous body of published literature. The focus on elderly patients who are ill or frail also remained unqualified in terms of significance; this factor could have been used efficiently to limit the broadly defined acute/chronic setting but was not applied in this way[11].

Returning to the previous definition of acute care provided in Section 5.2 (page 14) for the purposes of study selection, the auditor assumed that acute care was defined as immediate care for short-term treatment of severe injury, severe episode of illness, or during recovery from urgent or emergency surgery. Outpatient services, rehabilitation (unless stated to be within a 'long-term/chronic' setting), palliative care, elective surgery, day surgery, and emergency services were excluded[12]. The auditor also assumed that chronic care excluded general residential care settings without specific nursing or medical support (in-patient hospital or residential nursing settings only). Mobilization was assumed to be either active or passive and to be distinct from physiotherapy or rehabilitative therapy unless it was clear that

[11] It was noted that the internal CADTH reference for this project was 'RB0654 Ambulating Elderly Patients', but unfortunately this was not specified as a criterion in the research question

[12] The Canadian CIHI define acute care as "*necessary treatment for a disease or severe episode of illness for a short period of time with the goal of early discharge*" (www.CIHI.ca. accessed 15/02/15). IHPA define rehabilitation, palliative care, geriatric evaluation and management, and psychogeriatrics as 'sub-acute' (www.ihpa.gov.au. Accessed 15/02/15)

these terms included mobilization, ambulation, strengthening or endurance training. Multi-modal therapy was excluded unless mobilization, ambulation, strengthening or endurance training was specifically defined and/or reported.

Table 9 A comparison of PICO(D) as captured by original reviewer and the auditor for RB0654

PICO(D)	Original reviewer	Auditor	Comment
Population	Adults ≥18 years of age that are inpatients (for any reason) of acute hospital admissions or long-term/chronic care facilities	Adults ≥18 years of age (seniors who are ill or frail are a subpopulation of interest) that are inpatients (for any reason) of acute hospital admissions or long-term/chronic care facilities	Unclear (TRF discusses interest in seniors who are ill and frail and final report specifically identifies evidence for this subpopulation but it was not mentioned in PICO details, no clarification of the nature of acute or chronic care)
Intervention	Mobilization (terms also used are ambulation, strengthening/ endurance training)	Active or passive mobilization (including ambulation, strengthening, endurance training)	Unclear (active or passive, relationship to physiotherapy or multi-model intervention)
Comparator	None	None specified	-
Outcomes	Patient benefits and harms Safety Guidelines	Patient benefits and harms Safety Guidelines	-
(Design)	HTA/systematic review/meta-analysis; RCTs; NRS; guidelines	HTA/systematic review/meta-analysis; RCTs; NRS; guidelines	-

NRS: non-randomized studies

Study selection

The final CADTH report confirmed that two systematic reviews, one RCT, and five non-randomized studies were identified regarding benefits and harms associated with mobilization of adult inpatients in hospitals or long-term/chronic care facilities. Of these, one systematic review and three non-randomized studies described an elderly population with the remainder referring to adult patients. The auditor identified three out of four of these studies; the auditor excluded the fourth study by Padula et al [29] (see first bullet point below). The reviewer identified one systematic review, one RCT and two non-randomized studies for the general adult population whereas the auditor identified these same four studies plus the nine studies described in Table 10 (one systematic review, two RCTs, and six non-randomized studies). Given apparent differences in study selection, the auditor also reviewed studies listed under 'Further information' (four references) and 'Additional references' (seven references) in detail; the auditor had also excluded these studies from inclusion. The vast majority of discordance arose from the

auditor including studies that the reviewer had excluded, differences regarding specific studies are highlighted below and the nine auditor-included additional studies are presented in Table 10. Other specific differences are captured below:

- Padula et al 2011 [29] was included by the reviewer but excluded by the auditor on the basis that the abstract did not report any relevant outcomes; the only information provided is that "*one assisted fall occurred on the intervention unit.*"
- A point-prevalence study of ICU mobility care practices (with 45% of patients mechanically ventilated) in Australia and New Zealand reporting low patient mobilization rates (86% were not walked) [30] was included by the reviewer; conversely, a point-prevalence study of mobilization of mechanically ventilated ICU patients in Germany concluded that 76% were not mobilized out of bed [31] was excluded by the reviewer. The auditor had included both references as both mentioned adverse events/complications associated with mobilizing this acute adult in-patient population.

Despite making additional assumptions to inform interpretation of the research question, the auditor was still unable to exclude several studies at the completion of study selection. Having considered the content of the final published CADTH report (including additional and further references) that focused almost exclusively on studies in critical ill/ICU patients except for a handful of studies addressing mobilization specifically in the elderly, and after some deliberation, it was concluded that this group of studies, whilst seemingly meeting the inclusion criteria as currently stipulated, were likely not studies that were relevant to the research question as conceptualised by the original requester and the reviewer. Several studies addressed early mobilization within an acute stroke unit setting in adult patients who had experienced stroke (two papers addressed subarachnoid haemorrhage) within the last 24 hours. It was unclear whether these studies represented an acute adult in-patient population (to be included according to the research question) or an early but sub-acute rehabilitation population (and therefore excluded). Other studies addressed mobilization of in-patients admitted for epistaxis (nose bleed) [32] and mobilization specifically to reduce pressure ulcers in a surgical critical care unit [33].

Summary and interpretation

The overall summary of findings prepared by the reviewer and the auditor are presented in Table 11. The comments are largely comparable with the reviewer providing more detail than the auditor. One study was interpreted slightly differently [30]; the reviewer noted *"observed mobilization activities included in-bed exercises, sitting over the side or out of bed, standing, and walking. Few adverse events were recorded."* whereas the auditor commented that the study *"reported low mobilization rates with 86% of patients not walked at all during the period"*. Both comments are correctly extracted from the abstract but the implications upon reading are quite different. All of the additional studies included by the reviewer were set in a high-dependency/ICU setting and compared mobilization with either standard care or no intervention; all of these studies concluded that mobilization was safe, well tolerated and led to improved functional outcomes.

Learnings

Despite making additional assumptions to inform interpretation of the research question, the auditor was unable to exclude several studies at the completion of study selection due to ambiguity in the research question.

Table 10 Studies selected by the auditor for inclusion

Reference	PICO(D)	Details	Author conclusions
Systematic review			
Li et al. Active mobilization for mechanically ventilated patients: a systematic review. Arch Phys Med Rehabil. 2013 Mar;94(3):551-61	Population	Mechanical ventilation patients	Active mobilization appears to have a positive effect on physical function and hospital outcomes in mechanical ventilation patients. Early active mobilization protocols may be initiated safely in the ICU setting and continued in post-ICU settings. However, the current available studies have great heterogeneity and limited methodologic quality. Further research is needed to provide more robust evidence to support the effectiveness and safety of active mobilization.
	Intervention	Active mobilisation	
	Comparator	Not specified	
	Outcomes	Physical function and hospital outcomes	
	(Design)	Systematic review of RCTs and NRS	
RCT			
Burtin et al. Early exercise in critically ill patients enhances short-term functional recovery. Crit Care Med. 2009 Sep;37(9):2499-505.	Population	Adult ICU patients	At intensive care unit discharge, quadriceps force and functional status were not different between groups. At hospital discharge, 6-min walking distance, isometric quadriceps force, and the subjective feeling of functional well-being (as measured with "Physical Functioning" item of the Short Form 36 Health Survey questionnaire) were significantly higher in the treatment group
	Intervention	Passive or active exercise training session for 20 mins/day plus respiratory physiotherapy and a daily standardized passive or active motion session of upper and lower limbs	
	Comparator	Respiratory physiotherapy and a daily standardized passive or active motion session of upper and lower limbs	
	Outcomes	Quadriceps force, functional status, six-minute walking distance	
	(Design)	RCT	
Schweickert et al. Early physical and occupational therapy in mechanically ventilated, critically ill patients: a randomized controlled trial. Lancet.	Population	Adult ICU patients	A strategy for whole-body rehabilitation-consisting of interruption of sedation and physical and occupational therapy in the earliest days of critical illness-was safe and well tolerated, and resulted in better functional outcomes at hospital discharge, a shorter duration of delirium, and more ventilator-free days compared with standard care
	Intervention	Early exercise and mobilisation and daily interruption of sedation	
	Comparator	Daily interruption of sedation	
	Outcomes	Independent functional status at hospital discharge	

Reference	PICO(D)	Details	Author conclusions
2009 May 30;373(9678):1874-82.		defined as the ability to perform six activities of daily living and the ability to walk independently	
	(Design)	RCT	
NRS			
Clark et al. Effectiveness of an early mobilization protocol in a trauma and burns intensive care unit: a retrospective cohort study. Phys Ther. 2013 Feb;93(2):186-96.	Population	ICU patients	Early mobilization of patients in a TBICU was safe and effective. Medical, nursing, and physical therapy staff, as well as hospital administrators, have embraced the new culture of early mobilization in the ICU
	Intervention	Early mobilization protocol	
	Comparator	Pre intervention	
	Outcomes	Complication rates, ventilator days, and ICU and hospital LOS	
	(Design)	Retrospective cohort study	
Engel et al. ICU early mobilization: from recommendation to implementation at three medical centers. Crit Care Med. 2013 Sep;41(9 Suppl 1):S69-S80.	Population	ICU patients	Instituting a planned, structured ICU early mobility quality improvement project can result in improved outcomes and reduced costs for ICU patients across healthcare systems
	Intervention	Early mobility programme	
	Comparator	Not specified	
	Outcomes	LOS, delirium and need for sedation	
	(Design)	Not specified	
Morris et al. Receiving early mobility during an intensive care unit admission is a predictor of improved outcomes in acute respiratory failure. Am J Med Sci. 2011 May;341(5):373-7.	Population	A cohort of acute respiratory failure survivors, who participated in an early intensive care unit (ICU) mobility program	Tracheostomy, female gender, higher Charlson Comorbidity Index and lack of early ICU mobility were associated with readmissions or death during the first year. Although the mechanisms of increased hospital readmission are unclear, these findings may provide further support for early ICU mobility for patients with acute respiratory failure
	Intervention	ICU mobility therapy	
	Comparator	No mobility therapy	
	Outcomes	Variables associated with readmission or death	
	(Design)	Retrospective cohort study	
Mah et al. Resource-	Population	ICU patients	A team-based, resource-efficient approach to early

Reference	PICO(D)	Details	Author conclusions
efficient mobilization programs in the intensive care unit: who stands to win? Am J Surg. 2013 Oct;206(4):488-93.	Intervention	Resource-efficient mobilisation programme	mobilization is feasible and effective in the ICU.
	Comparator	No treatment	
	Outcomes	Bed-to-chair evaluation, sitting balance, ambulation	
	(Design)	NRS	
Nydahl et al. Early Mobilization of Mechanically Ventilated Patients: A 1-Day Point-Prevalence Study in Germany. Crit Care Med. 2013	Population	Mechanically ventilated ICU patients	In this 1-day point-prevalence study conducted across Germany, only 24% of all mechanically ventilated patients and only 8% of patients with an endotracheal tube were mobilized out of bed as part of routine care. Addressing modifiable barriers for mobilization, such as deep sedation, will be important to increase mobilization in German ICUs
	Intervention	Early mobilisation	
	Comparator	Not specified	
	Outcomes	Level of mobilisation and barriers	
	(Design)	Point-prevalence study	
Titsworth et al. The effect of increased mobility on morbidity in the neurointensive care unit. . J Neurosurg. 2012 Jun;116(6):1379-88	Population	Neuro ICU patients	Among neurointensive care unit patients, increased mobility can be achieved quickly and safely with associated reductions in LOS and hospital-acquired infections using the PUMP Plus program
	Intervention	Mobility initiative utilizing the Progressive Upright Mobility Protocol	
	Comparator	No intervention	
	Outcomes	Mobility, LOS, HAI, VAP, days in restraints, adverse events	
	(Design)	NRS	

Table 11 Overall summary of findings prepared by reviewer and auditor for RB0654

Original reviewer overall summary of findings	Auditor overall summary of findings
Four studies [34] [35] [36] [29] investigated mobilization strategies for elderly inpatients. One systematic review [34] examined the effectiveness of early physical rehabilitation programs for geriatric patients who were hospitalized. Patients who were involved in either multidisciplinary or exercise programs were less likely to be discharged from hospital to a nursing home than were geriatric inpatients who received usual care. One non-randomized study [35] evaluated the frequency and duration of mobilization of older patients in acute care by nurses. Standing and transferring were the most commonly observed mobilization events. Patients who were unable to move themselves were mobilized less frequently than patients who had mobility, and most instances were initiated by patients, not by the nursing staff. One non-randomized study [36] assessed the effects of a strength training program for nursing home residents with impaired mobility. After eight weeks, mobility and muscle strength in the limbs had improved. Other quality of life measures did not change. In another study [29] lower extremity strength training was added to the mobilization protocol for elderly hospitalized adults to determine the effect on falls. One assisted fall was recorded in the intervention group. Four studies [37] [38] [30] [39] investigated mobilization strategies for adult inpatients. One systematic review [37] examined the effectiveness of early mobilization of critically ill patients in the intensive care unit (ICU). Few studies were identified for inclusion in the review; however, those that were indicated that early mobilization and physical therapy were safe and could have an impact on functional outcomes. One randomized controlled trial [38] compared the effect of an early mobilization protocol with standard physical therapy on respiratory and peripheral muscles of inpatients. There were significant changes in inspiratory and peripheral muscle strength in the early mobilization group. There were no significant differences between groups in length of ICU stay or length of hospital stay. One non-randomized study [30] evaluated the mobility practices of one ICU during a 24 hour period. Observed mobilization activities included in-bed exercises, sitting over the side or out of bed, standing, and walking. Few adverse events were recorded. In one non-	One systematic review, one RCT and two non-randomized studies evaluated mobilization for adult in patients in an acute setting [37] [38] [30] [39]. All studies were set in the ICU. The systematic review [37] included 15 studies and concluded available literature supports early mobilization and physical therapy as a safe and effective intervention that can have a significant impact on functional outcomes. Two studies reported functional improvements (inspiration and peripheral muscle strength [38], saturation improvements and an exercise intensity akin to walking [39]). One study recorded mobilization activities over a 24-hr period and reported low mobilization rates with 86% of patients not walked at all during the period [30]. One systematic review and two non-randomized studies evaluated mobilisation-based interventions for elderly patients in an acute [34] [35] or chronic setting [36]. All three studies reported benefits associated with mobilisation in the elderly population. The systematic review [34] included studies of in-hospital early physical rehabilitation for patients aged 65 years+ with an outcome measure of physical functioning and concluded that patients who had participated in a multidisciplinary program or exercise program improved more on physical functional tests and were less likely to be discharged to a nursing home compared to patients receiving only usual care. The acute care NRS [35] addressed the time and motion dedicated to mobility events and concluded that limited mobilisation was an independent predictor of negative outcomes for hospitalized older patients. The only chronic care study [36] evaluated resistance training twice weekly for very elderly nursing-home residents and reported considerable improvement.

randomized study [39] early rehabilitation (including chair sitting, tilting, and walking) was provided to patients who were in the ICU for seven or more days and were mechanically ventilated for at least two days. Chair sitting was the most frequently reported intervention and was associated with a significant decline and heart rate and respiratory rate. The authors concluded that early intervention was safe and feasible for patients in the ICU.	

6.3. Public Automated External Defibrillators (AED) and Cardiopulmonary Resuscitation (CPR) Education: Clinical and Cost-Effectiveness (RB0721)

Although public provision and use of AED has been shown to be effective, cost effectiveness remains uncertain as the majority of cardiac arrests do not occur in the public setting. In addition, some research suggests that increasing and improving CPR education in various settings (i.e. high schools) can enable individuals to respond to cardiac arrests occurring in private residences until the emergency services arrive. These issues raise a question about whether public AED provision or investment in CPR education represents the best use of available resources.

Auditor Comments

Refining the research question

The TRF adequately defined the research questions and provided useful background contextual information on the rationale for the request. No comparator was stated for Q3 (not even 'none specified' or 'any'). To maximize clarity, the auditor added additional detail stipulating out-of-hospital cardiac arrest and for the intervention to be provided by a bystander to differentiate between defibrillation provision by dispatched emergency services when the origin of the AED itself was not stated. As there was more than one research question covering two different interventions, the auditor also explicitly split out the public AED research question from the public education research question for ease of understanding. A summary of the research question as defined by the original reviewer and the auditor is presented in Table 12.

Table 12 A comparison of PICO(D) as captured by original reviewer and the auditor for RB0721

PICO(D)	Original reviewer	Auditor	Comment
Research question 1: What is the clinical and cost effectiveness of public AEDs (plus or minus CPR) compared to CPR alone or no CPR for treating out-of-hospital cardiac arrest			
Population	Adults experiencing cardiac arrest	Adults experiencing out-of-hospital cardiac arrest	-
Intervention	Q1, Q2: Public automated external defibrillators (AEDs) with or without CPR Q3: Public education on CPR interventions	Public automated external defibrillators (AEDs) with or without CPR administered by bystanders	-
Comparator	Q1, Q2: None Q1, Q2: Administration of CPR alone	CPR alone administered by bystanders or no CPR	-
Outcomes	Comparative effectiveness for clinical outcomes (mortality and morbidity); cost effectiveness	Clinical outcomes (inc mortality and morbidity); cost effectiveness	-
(Design)	HTA/systematic review/meta-analysis; RCTs; NRS; economic evaluations	HTA/systematic review/meta-analysis; RCTs; NRS; economic evaluations	-

Research question 2: What is the evidence for public CPR education (e.g. in schools) for improved out-of-hospital cardiac arrest outcomes			
Population	As above	Adults experiencing out-of-hospital cardiac arrest	-
Intervention	As above	Public education on CPR interventions (e.g. high school)	-
Comparator	Not stated above	None; no education	Unclear (not stated)
Outcomes	As above	Clinical outcomes (inc mortality and morbidity)	-
(Design)	As above	HTA/systematic review/meta-analysis; RCTs; NRS; economic evaluations	-

Study selection

The final CADTH report confirmed that two non-randomized studies regarding the comparative clinical effectiveness of public AEDs (with or without CPR) versus CPR alone or no CPR in treating cardiac arrests, and three economic evaluations regarding the cost-effectiveness of using public AEDs for the treatment of out-of-hospital cardiac arrest were included. The auditor identified three out of five of these references for inclusion. One abstract (Berdowski et al 2010 [40][13]) appeared to be a reference generated by the database search (as opposed to grey literature or additional internet searches) but was not found in the provided project bibliography and was not available for selection. The second discrepancy was attributable to a cost effectiveness study by Folke et al 2009 [41]; the reviewer had included this study whereas the auditor had excluded it on the grounds that it did not provide information relating to a CPR/no CPR comparator (the abstract discussed focussed versus unfocused AED deployment).

Summary and interpretation

The overall summary of findings prepared independently by the reviewer and the auditor are presented in Table 13. Accounting for differences in study selection, the reviewer and auditor summations were aligned. Notable differences in the level of detail were apparent in so far as the reviewer provided several paragraphs with more descriptive detail whereas the auditor provided only one. The reviewer and auditor both reported the same conclusions for the common studies included.

Learnings

A well-defined research questions accompanied by useful background contextual information meant that there were minimal differences between the reviewer and the auditor in terms of study selection and summarisation.

[13] There were three Berdowski references [89] [90] [91] in the project bibliography but none matched the reference given in the published CADTH Rapid Response report [40]

Table 13 Overall summary of findings prepared by reviewer and auditor for RB0520

Original reviewer overall summary of findings	Auditor overall summary of findings
Two non-randomized studies [42] [43] regarding the comparative clinical effectiveness between the use of public AEDs (with or without CPR), versus the use of CPR alone or no CPR in treating cardiac arrests, and three economic evaluations [40] [41] [44] regarding the cost-effectiveness of using public AEDs for the treatment of out-of-hospital cardiac arrest were identified. No evidence regarding CPR education for improved out-of-hospital cardiac arrest outcomes was identified, therefore no summary on this aspect can be provided. A retrospective analysis [42] of ICU patients observed that bystander AED treatment was associated with significantly improved neurological outcomes compared to CPR only, or no treatment. A prospective cohort study [43] observed the highest rate of survival with bystander AED use, and a positive association between AED use and survival in multivariate regression models. Three studies focused on the economic aspects of public AEDs. One prospective cohort study [40] reported an increased survival rate in patients where an on-site (i.e., public) AED or dispatched AED was used compared to no AED, and that total in-hospital health care costs were lower for survivors of the on-site AED group, mainly due to reduced length of hospital stay. One economic analysis using data from Denmark [41] reported that following European Resuscitation Council (ERC) and American Heart Association (AHA) guidelines for the deployment of AEDs only in areas with a high incidence of cardiac arrest, cost incurred per quality-adjusted life year (QALY) was lower compared to unguided AED placement. The authors recommended that in order to make best use of public AEDs, they should be more widely distributed than is mandated by ERC guidelines, which could be achieved by following AHA guidelines. [41] One prospective multicentre randomized trial5 reported higher survival rates for patients who received CPR+AED compared to CPR alone, and that application of CPR+AED resulted in higher mean QALY, higher mean costs, and similar long term costs compared to CPR only [44]. In conclusion, an association between public AED use (with or without CPR) and improved survival [40] [43] [44] and neurological outcomes1 compared to CPR alone or no treatment has been observed. There is inconsistent evidence [40] [41] [44] regarding the cost-effectiveness of using public AEDs in the treatment of out-of-hospital cardiac arrest.	Two non-randomized studies and one economic evaluation relating to the effectiveness of implementation of public AEDs for out-of-hospital cardiac arrest were included. Stammet et al [42] described a retrospective ICU chart review whilst Wesifeldt et al [43] undertook a population-based cohort study of out-of-hospital cardiac arrest before emergency services arrival. Both studies concluded that, compared to CPR (and also compared to no CPR/rescue measures [42], bystander administration of an AED was associated with positive clinical outcomes in terms of survival [43] or good neurological outcomes [42] for adults experiencing out-of-hospital cardiac arrest. The trial-based economic evaluation calculated the cost effectiveness of CPR+AED versus CPR alone when applied by lay responders [44]. CPR+AED was associated with greater costs, greater quality-adjusted life years (QALYs) and similar longer-term costs when compared to CPR. The authors concluded that the incremental cost effectiveness of CPR+AED was likely similar to other common health interventions. No studies addressed public CPR education for improved clinical out-of-hospital cardiac arrest outcomes.

7. Results: Level 2 (Summary with critical appraisal)

- Four Level 2 CADTH reviews were selected for evaluation: L0161, L0227, RC0441, RC0570
- The TRF adequately defined the research question and provided useful background, there was some deviation from the presentation stipulated in the instructions to authors observed (e.g. no explicit critical appraisal, no key message summary)
- Differences observed in the review of L0227 reinforced the importance of documenting a clear and unambiguous research question has been confirmed.
- The critical learning from the evaluation of RC0441 was regarding the impact of undocumented post-hoc changes to the TRF; whilst increasing the amount of relevant information and making for a better reflection of the nature and severity of disease, the auditor was not able to match the reviewers study selection
- The research question for RC0570 was adequately defined and resulted in minor differences in study selection, summarisation and interpretation, this had minimal impact and the reviewers and auditors conclusions and implications were aligned.

Level 2 Rapid Response Reports are written summaries of the evidence from full text articles, with a critical appraisal and policy implications, and typically take at least 30 days. As described in Table 2, the auditor's tasks associated with selected Level 2 included tasks for Level 1, Level 1.5 plus a comparative evaluation of critical appraisal, data extraction, data presentation and conclusions and interpretation. Specific instructions on the format and content of key elements from the CADTH Summary with Critical Appraisal Author Requirements are noted in Appendix I.

7.1. Tinnitus Retraining Therapy: A Review of the Clinical Effectiveness (L0161)

This report was requested due to the multiple regional funding requests received for tinnitus retraining therapy. The requestor raised concerns that favourable findings were a consequence of the duration of therapy (18 months) and time-related spontaneous improvement in symptoms. It was suspected also that most evidence was of poor quality and funded by manufacturers of tinnitus masking devices.

Auditor Comments

Refining the research question

The TRF adequately defined the research question and provided useful background contextual information on the rationale for the request. A summary of the research question as defined by the original reviewer and the auditor is presented in Table 14.

Table 14 A comparison of PICO(D) as captured by original reviewer and the auditor for L0161

PICO(D)	Original reviewer	Auditor	Comment
Population	Any patient with tinnitus	Patients with tinnitus	-
Intervention	Tinnitus retraining therapy (TRT) - adaptation therapy using counselling and a tinnitus control instrument (TCI) or sound generators	Tinnitus retraining therapy (TRT) – defined as counselling plus a tinnitus control instrument or sound generator	-
Comparator	None	Not specified	-
Outcomes	Clinical effectiveness - adaptation, improvement of symptoms, relief	Clinical effectiveness - adaptation, improvement of symptoms, relief	-
(Design)	HTA/systematic review/meta-analysis; RCTs; CCTs; observational studies	HTA/systematic review/meta-analysis; RCTs; CCTs; observational studies (if few other study types found)	-

Study selection

The final CADTH report confirmed that two systematic reviews, three RCTs, and two CCTs regarding TRT were included in the final review. The auditor identified two RCTs, two CCTs and seven observational studies. The reviewer and auditor included the same two RCTs and CCTs. Two references (both systematic reviews) included by the reviewer were attributed to additional internet searches and did not appear in the project bibliography reviewed by the auditor [45] [46]. The seven observational studies identified for inclusion by the auditor were listed only in the appendix by the reviewer [47] [48] [49] [50] [51] [52] [53]. It was somewhat unclear whether these were to be regarded as 'included' studies or just studies of general interest. The reviewer stated, *"since there were a number of systematic reviews, RCTs, and controlled clinical trials identified in the search, observational studies of TRT were not summarized and have been listed in Appendix I"* and this was confirmed by the consistency between the reviewers and auditors selection of these studies. However, the Instructions to Authors for Level 2 reviews suggests that only studies which do not meet all the selection criteria but that may be of interest to the requestor should be listed as an appendix.

Other specific differences in study selection:

- The RCT published by Hiller et al 2005 [54] was included by the original reviewer but excluded by the auditor on the basis that the study assessed sound generators as an additive to CBT (the intervention was defined as CBT plus low level white noise generator, the comparator as CBT alone); in the auditor's view this did not match the inclusion criterion for TRT as counselling and CBT represent distinct psychotherapeutic interventions

- Henry et al 2009 [55] was listed in the appendix by the reviewer but was entirely excluded by the auditor as it used tinnitus retraining counselling alone (without any type of sound generation device)
- Several studies listed in the appendix were not found in the project bibliography and were assumed to be attributable to additional searches [56] [57] [58] [59]

Summary of study characteristics

The reviewer did not tabulate the summary of study characteristics; information was provided only in narrative format. The auditor tabulated relevant study summary information. Having completed this tabulation independently, the auditor pasted the reviewer's comments under relevant table headings to facilitate comparison between the reviewer's and the auditor's extraction of information. It should be noted that the reviewer's text was not intended (and therefore not written or formatted) to be presented in this fashion and there maybe differences purely due to this factor.

Table 15 Table of characteristics of included studies as extracted by the auditor and the reviewer

	First Author, Publication Year, Country	Study Design, Inclusion criteria, Length of Follow-up	Patient Characteristics	Intervention (n)	Comparator(s) (n)	Clinical Outcomes
	RCTs					
Auditor	Caffier 2006 Germany [60]	RCT Tinnitus duration of more than 6 months, age between 18 and 80 yr, available linguistic and intellectual skills to fill out the questionnaires, and exclusion of Meniere's disease and vestibular schwannoma. All severity levels were considered. 2 years	Mean age 51 years, 54% male, 48 patients were recruited but eight were excluded due to lack of compliance, withdrawal, or incomplete records. Mean duration of tinnitus at baseline: 6.8 yrs (6 mo. to 28 yrs)	TRT (Counselling, auditory training, progressive muscle relaxation, psychosomatic/ psychotherapeutic care if necessary; mandatory binaural provision of TCIs) N=20	Waiting list control N=20 [Control group offered TRT after 12 months]	Tinnitus questionnaire VAS loudness, annoyance, and awareness Severity questionnaire anxiety, sleep and concentration disturbances, and psychosocial stress
Reviewer		Adults with tinnitus for longer than six months duration were recruited from a German tinnitus center.	A total of 48 patients (mean age 51 years, 54% male) were enrolled, however, eight (17%) were excluded from the analysis due to lack of compliance, withdrawal, or incomplete records	Modified TRT program that included counselling, a sound generator, relaxation, and psychotherapeutic care if needed	Wait-list control After 12 months, the control group was offered TRT	See Table 17

Auditor	Zachriat 2004 Germany [61]	RCT. Tinnitus duration of more than 3 months, absence of treatable organic causes of tinnitus, absence of Meniere's disease, hearing capacity sufficient for communication within groups, tinnitus disability score ≥25, no ongoing psychotherapy or masker treatment. 21 months	Mean age 51-56 years, 66-74% male, nine were excluded due to withdrawal. Mean duration of tinnitus at baseline varied between 65-90 months (4 to 324 months).	Habitution-based training (TRT) (counselling of 5 sessions over 6 months, plus sound generator use for minimum 6 hrs per day) N=31	TCT (11 weekly group sessions of relaxation training, CBT-based counselling, coping strategies) N=29 EDU (single group session on physiology and psychology of tinnitus) N=23	Tinnitus diary Tinnitus questionnaire Tinnitus coping questionnaire, Catastrophizing of cognitions questionnaire Dysfunctional cognitions Jastreboff questionnaire Subjective success questionnaire Symptom checklist

Reviewer		RCT Patients were recruited from the community and were enrolled if they had tinnitus of greater than three months duration and a tinnitus disability score ≥25 on the TQ.	The mean age per group varied from 52 years to 56 years and 59% to 74% of participants were male.	A total of five group habituation-based treatment sessions of 90 to 120 minutes in duration were spaced over a period of six months. Patients in the habituation-based treatment group also received bilateral wide band noise generators and were instructed to wear the devices for at least six hours per day…followed for 18 or 21 months 31 in the habituation-based treatment	Coping training consisted of 11 weekly group sessions of 90 to 120 minutes duration… followed for 18 or 21 months 29 in the coping training Patients in the control group received a single educational session which provided similar information as was given in the first session of the other two programs. Patients in the control group were followed for 14 weeks 23 in the control group	See Table 17
	Controlled clinical trials					
Auditor	Davis 2008, Australia [62]	CCT (parallel group, repeated measures) No inclusion criteria stated by authors 12 months	Mean age 49.8 years, approx. 50% male, 38 patients excluded prior to treatment allocation, 5 were lost to follow-up prior to study completion Mean duration of disturbing tinnitus at baseline 3.6 years (range 0.2 to 23), mean pretreatment TRQ score	Customized acoustic stimulus with instructions to use for at least 2 hrs per day at a volume that masked tinnitus plus rehabilitation programme N=13 Customized acoustic stimulus with instructions to use for at least 2 hrs per day half the time at a volume	Counselling only (counselling and support, self-help book) N=13	• Tinnitus reaction questionnaire • VAS for loudness, severity and relaxation • Perceived benefit survey • Objective audiologic measurements

			for the 50 patients was 39.3 (range: 17 to 91).	that masked tinnitus plus rehabilitation programme N=9 Noise generator with instructions to use for at least 2 hrs per day + counselling (counselling and support, self-help book) N=15		
Reviewer		Eight-eight patients were enrolled (selection process not reported) and allocated by alternation to each of the four treatment groups.	From the patients enrolled, 38 (43%) were excluded from the analysis for various reasons. The patients had a mean age of 50 years, 52% were male, and had moderate to severe tinnitus. The two Neuromonics groups were combined since patients did not adhere to the prescribed volume settings and analysis of results showed no difference between groups.	Two groups received Neuromonics customized acoustic stimulation, but at different volume levels.	The third group received broadband noise generator for the sound therapy portion of treatment. All groups received counselling.	Tinnitus related distress was measured using the Tinnitus Reaction Questionnaire (TRQ; score range 0 to 104) with a change of 40% required to mark a clinically important improvement from baseline. See Table 17

Auditor	Henry 2006, country [63]	CCT with alternate allocation to treatment groups Military veterans with clinically significant tinnitus (i.e., a tinnitus condition warranting 18 months of individualized treatment) 18 months	Mean age 59-61years, approx. 95% male, large numbers excluded during screening, 48 patients excluded prior to treatment allocation, 5 patients were excluded from analyses for incomplete data on predictive variables. Not reported	TRT (structured education counselling, sound therapy using TRT 'approved' devices used for a minimum of 8 hrs per day) N=64	TM (unstructured counselling + sound therapy using any device) N=59	• Tinnitus Handicap Inventory • Tinnitus Handicap Questionnaire • Tinnitus Severity Index • Percentage ratings of awareness and annoyance
Reviewer		Veterans with clinically significant tinnitus from the community and from an audiology clinic to participate in this 18 month clinical trial.	A total of 123 patients were alternately assigned to either tinnitus masking or TRT treatment groups (95% male, mean age 60 years).	Patients in the TRT group received structured counselling according to TRT methods. Both groups also received sound therapy using a sound generator, hearing aid, or other ear level devices to use for at least eight hours per day	Those in the masking group received informal counseling. Both groups also received sound therapy using a sound generator, hearing aid, or other ear level devices to use as needed	Patients were assessed using three validated tools (Tinnitus Severity Index [TSI], Tinnitus Handicap Questionnaire [THQ], Tinnitus Handicap Inventory [THI]), and a visual analog scale assessing tinnitus awareness and annoyance. The authors analyzed data using a multilevel regression model appropriate for this type of data. See Table 17

Critical appraisal

The results of the reviewer's critical appraisal were not tabulated or explicitly defined, with the available text for each study being largely descriptive in nature. For the purposes of comparison, the key points mentioned in the reviewer's 'Limitations' section are listed as bullet points below. Key points from the auditor's critical appraisal for each included study are summarized in Table 16.

- Patient selection criteria and randomisation methods used were unclear
- Poor reporting of baseline characteristics
- Blinding not mentioned
- Unclear if analyses were undertaken on an ITT basis with substantial numbers lost to follow-up and/or excluded from analyses and/or with missing data
- Small sample size, unevenly sized treatment groups
- Contact time with investigators varied by treatment
- Selective outcomes reported (at 6 months for one outcome but 12 months for another)
- Patients included represented a range of tinnitus severity

The reviewer concluded that the studies were of low methodological quality and a high risk of bias. The auditor was in agreement with this conclusion. Although presented differently, the general comments raised by the reviewer matched the specific comments listed by the auditor.

Table 16 Summary of auditor's critical appraisal of included studies

Reference	Strengths	Limitations
RCT		
Caffier 2006 [60]	• Patients were randomly assigned to treatment groups (method not described) • Inclusion criteria were clearly stated • Withdrawals and dropouts were reported	• No power calculations were undertaken • Patients and observers were not blinded to treatment allocation • Small patient numbers were recruited • All severity levels were considered eligible for participation • Limited description of baseline characteristics • Limited reporting of results according to randomized treatment allocation
Zachriat 2004 [61]	• Hypotheses were clearly stated • Inclusion criteria were clearly stated • Patients were randomly assigned to treatment groups (by throwing dice) • Withdrawals and dropouts were Baseline characteristics were reported • Analyses appear to include all randomized patients who received treatment • All outcomes are accounted for	• Randomisation and treatment allocation possibly affected by "practical reasons" leading to imbalanced treatment groups • No power calculations were undertaken • Patients and observers were not blinded to treatment allocation • Small patient numbers were recruited • Unequal duration of follow-up
Davis 2008 [62]	• Interventions clearly explained • Withdrawals and dropouts were	• No inclusion criteria specified • No information about patient recruitment

Reference	Strengths	Limitations
	clearly explained • Clinician contact time was balanced between treatment groups • Authors considered the sample to be representative of a typical clinic	reported • Patients were allocated by alternating them between groups in a random order • A large number of patients were excluded from analyses (38/88) • Small patient numbers were recruited • Limited description of baseline characteristics • Lack of patient compliance meant that two interventions became indistinguishable and were combined for analysis
Henry 2006 [63]	• Interventions clearly explained • Outcomes clearly defined and reported • Mostly balanced average clinician contact time	• Limited description of baseline characteristics • Treatment allocation by alternate groups • Over-representation of men (likely due to patient selection from military pool) • Post-hoc amendments to study protocol (additional outcome measures were added) • Incomplete outcome data for majority of patients (only complete for 46/118)

Summary of findings

Again the reviewer's summary of study findings was not tabulated, the auditors study findings are presented below in Table 17. Having completed this tabulation independently, the auditor pasted the reviewer's comments under relevant table headings to facilitate comparison between the reviewer's and the auditor's summarising of information.

Conclusions and implications

The reviewer did not provide a statement of key messages or findings but concluded *"Due to the low methodological quality of the studies available it is difficult to draw any firm conclusions on the effectiveness of TRT in the management of tinnitus. Compared to baseline measures, most TRT treatments showed improvement in outcomes over time. However, most studies had substantial losses to follow-up which may have biased results in favour of the treatment if patients with poor outcomes were excluded from the analysis. The potential for bias in the selection of patients, lack of blinding of outcome assessors, and use of non-equivalent comparator groups was also a concern and may be a consideration for decision-making about TRT for patients with tinnitus"*. The auditor was in agreement regarding the methodological limitations of these studies.

Learnings

There was some deviation from the presentation stipulated in the instructions to authors and what was reported in the final review (e.g. no explicit critical appraisal, no key message summary), however it was suspected that this may be due to changes in the instructions to authors that may have since taken place.

Table 17 Table of study findings and authors conclusions as extracted by the reviewer and the auditor

	Reference	Study findings	Author's conclusions
	RCT		
Auditor	Caffier 2006 [60]	• TQ score: after 12 months, the control group did not show any significant changes, mean TQ scores improved by 16 points on average $p < 0.001$ compared to baseline following TRT • VAS: Not reported by randomized treatment • Severity: Not reported by randomized treatment	The outpatient interdisciplinary TCT, consisting of cognitive tinnitus desensitization, TCI provision, and psychosomatic support if required, represents a successful treatment strategy for both mild to severe tinnitus
Reviewer		Initial TQ scores were similar between groups and remained unchanged for the control group. The mean TQ score decreased 16 points in the TRT group ($p<0.001$) after 12 months.	The authors concluded that the modified TRT program represented a successful treatment strategy for patients with mild to severe tinnitus
Auditor	Zachriat 2004 [61]	• Tinnitus coping questionnaire: Reductions for coping and habituation-based treatment groups were maintained to the end of follow up • Catastrophizing of cognitions questionnaire: no differences between groups, significant reduction from baseline with TCT ($p<0.05$) and HT ($p<0.05$) • Dysfunctional cognitions • Jastreboff questionnaire: TCT was successful in 44% of the patients, HT in 40% and in EDU only 15% of the patients at T4, percentages increase to 50% (TCT) and 44.6% (TRT) at T7 • Subjective success questionnaire: no significant differences between groups although change in TCT and HT groups greater than in EDU • Symptom checklist	Findings reveal highly significant improvements in both tinnitus coping training and habituation-based treatment in comparison with the control group. While tinnitus coping training and habituation-based treatment do not differ significantly in reduction of tinnitus disability, improvement in general well-being and adaptive behaviour is greater in tinnitus coping training than habituation-based treatment. The decrease in disability remains stable throughout the last follow-up in both treatment conditions.

Reviewer		At 14 weeks, tinnitus coping therapy and habituation-based treatment were more efficacious in reducing tinnitus-related disability than the control group ($p<0.05$) but did not differ from one another. Reductions in TQ scores in the coping and habituation-based treatment groups were maintained to the end of follow up, however no information was available to compare these differences to the control group. Tinnitus perception was not statistically significantly different between groups. At 18 months, 23% of patients in the habituation-based treatment group were still using the noise generators.	Based on the results of the trial the authors recommended that patients with chronic tinnitus be offered an educational session first, and then those with continuing complaints be offered further treatment with either habituation-based treatment or coping training.
	CCT		
Auditor	Davis 2008 [62]	• TRQ score: Only the neuromonics group reported statistically significant improvements from baseline in TRQ scores at 3,6, and 12 months ($p < 0.001$); statistically significant between group differences were between Neuromonics and Noise+Counseling and Neuromonics and Counseling-Only groups • VAS: At 12 months a significant improvement from baseline with Neuromonics only ($t = -12.86$, $p < 0.001$ severity; $t = -11.23$, $p < 0.001$ relaxation; $t = -4.86$, $p < 0.001$ tolerance); statistically significant between group differences between Neuromonics and Noise+Counseling and Neuromonics and Counseling-Only • Perceived benefit survey: 63-86% (depending on criterion) of neuromonics patients reported benefit, other groups reported much less (0-40%) • Objective audiologic measurements: Only the neuromonics group reported statistically significant improvements from baseline in TRQ scores at 6 and 12 months ($p < 0.001$); no significant between group differences reported	In our study, patients who received the customized stimulus (Neuromonics group) reported significantly greater and more consistent alleviation of tinnitus symptoms than did patients who participated in a counseling and support program with and without delivery of a broadband noise stimulus (Noise + Counseling group and Counseling-Only group, respectively).

Reviewer		TRQ scores in the Neuromonics group were statistically significantly lower at 3, 6, and 12 months, compared to baseline. No statistically significant differences in TRQ scores over time were detected for the counselling and the broadband noise group. TRQ scores were statistically significantly lower in the Neuromonics group compared to the other groups at 12 months.	The authors concluded that the Neuromonics therapy was superior to counselling or noise plus counselling
Auditor	Henry 2006 [63]	• Tinnitus Handicap Inventory/Tinnitus Handicap Questionnaire/ Tinnitus Severity Index: both groups resulted in declines in tinnitus handicap and severity but the decline in TRT patients was considerably greater than the decline in TM patients. The greater declines occurred most strongly in patients who began treatment with a “very big” tinnitus problem, for these patients, the rate of improvement was considerably faster in TRT compared to TM. • Percentage ratings of awareness and annoyance: for awareness TM showed a big decline (about 14 points every six months or 42 points over 18 months), with TRT adding an additional four-point decline every six months. Annoyance was reduced in TM patients by 26 percentage points, but it was reduced in TRT patients by an additional 14 percentage points	Both groups showed significant declines (improvements) on these measures, with the TRT decline being significantly greater than for TM. The greater declines in TRT compared to TM occurred most strongly in patients who began treatment with a “very big” tinnitus problem. When patients began treatment with a “moderate” tinnitus problem, the benefits of TRT compared to TM were more modest.
Reviewer		Analysis was limited by missing data with only 37% of patients having complete data for all outcomes at each assessment period over the 18 month follow-up. Data for TSI, awareness, and annoyance were the most complete (>70% of patients had data for all outcome periods). The authors reported that both treatment groups showed improvement in the outcome measures over time, with the TRT showing a greater rate of improvement (statistical significance unclear).	The authors concluded that those patients whose tinnitus had the greatest impact in their lives showed the strongest benefit to TRT therapy (statistical significance unclear)

7.2. Zoledronic Acid Intravenous Infusion: A Review of the Clinical Effectiveness and Guidelines (L0227)

At the time of writing, once yearly zoledronic acid was not available on drug formulary. There was a need to establish what the administration costs were and whether the once yearly infusion was as effective as oral options. Existing oral therapies (such as Fosamax) have been associated with pathologic fractures, there is evidence that once yearly zoledronic acid was effective as a second line therapy for these patients.

Auditor Comments

Refining the research question

The TRF did not adequately define the research question and could be improved to provide greater clarity to a third party. The population of interest was specified as 'outpatients with osteoporosis or who are at risk for hip, spinal and bone fractures', the use of 'or' could be taken to imply that studies of patients without a confirmed diagnosis of osteoporosis but at risk from fracture would be of interest (e.g. osteopenic patients and patients undergoing hormonal therapy for cancer or for other serious comorbidity). Furthermore, the list of comparators for Q1 is not definitive and for Q2 is not stated. To facilitate study selection, the auditor assumed that patients must have a confirmed diagnosis of osteoporosis and/or the study authors themselves must use the term 'osteoporosis' (there was no mention of osteopenia or sub-clinical BMD loss made in the TRF) and that patients must not have a serious co-morbidity (e.g. cancer). It was assumed that treatment naïve patients would also be fracture naïve (although a mixed population would also be accepted). The auditor also listed available oral osteoporosis agents as comparators and assumed that any active comparator was relevant to the second line patient population (Q2). A summary of the research question as defined by the original reviewer and the auditor is presented in Table 18.

Table 18 A comparison of PICO(D) as captured by original reviewer and the auditor for L0227

PICO(D)	Original reviewer	Auditor	Comment
Population	Q1,Q3 Outpatients with osteoporosis or who are at risk for hip, spinal and bone fractures Q2 Outpatients with osteoporosis who are taking an oral osteoporosis agent (such as fosomax) and who have pathologic fractures	Adult osteoporosis outpatients without other serious comorbidity who are at risk of pathologic fracture (and have not experienced fracture) or who are already taking an oral osteoporosis agent and experienced pathologic fracture	Query whether 'or' is meant to be 'and' as all patients must have osteoporosis (auditor interpreted it to be 'and'), unclear whether treatment naïve patients can have a history of fracture, unclear on status of serious comorbidity
Intervention	Aka zeldoronate, zometa, zomera, aclasta, reclast	IV zoledronic acid administered once yearly (zeldoronate, zometa, zomera, aclasta, reclast)	-

Comparator	Q1 Oral osteoporosis agents (for example- Fosomax, risendronate, ibandronate)	Oral bisphosphonate osteoporosis agents (fosomax or alendronate, risendronate or actonel or atelvia, ibandronate or Boniva, etidronate or didronel, pamidronate or ardeia, tiludronate or skelid)	Query whether the reviewer's list is definitive. No comparators are specified for Q2, unclear whether this means placebo is also relevant
Outcomes	Q1 Comparative clinical effectiveness, safety of IV drugs Q2 Effectiveness and safety of IV drug as a second-line treatment Q3 Guidelines and recommendations surrounding the yearly infusion- for example how long should they be at the clinic	Comparative clinical effectiveness and safety; guidelines regarding administration of yearly infusion	-
(Design)	HTA/systematic review/meta-analysis; RCTs; NRS; guidelines	HTA/systematic review/meta-analysis; RCTs; NRS; guidelines	-

Auditor Comments

Study selection

The final CADTH report confirmed that three RCTs comparing the clinical effectiveness of once-yearly zoledronic acid infusion with oral bisphosphonates were identified; no clinical evidence addressing the use of annual zoledronic acid intravenous infusion as a second-line therapy for patients with pathologic fractures who were taking oral osteoporosis therapies was identified; and no guidelines on the use of once-yearly zoledronic acid intravenous infusion in outpatient settings were identified. The auditor identified two of the three RCTs included by the reviewer reporting the clinical effectiveness versus oral bisphosphonates and concluded that that there were no studies relevant to Q3. For Q2, the auditor's view was less clear-cut keeping in mind that the reviewer had included the study by McClung et al [64]. The auditor had excluded McClung et al 2007 [64] on the basis that it reported a second line treatment population (patients who had already undergone treatment with alendronate and so met the first part of the inclusion criteria regarding relevant population for Q2) but there was no mention or requirement for patients to have already experienced fracture so the population did not fully meet the criteria for Q2. The auditor also had a question-mark regarding Reid et al 2009 [65] and the nature of patient comorbidities given that the population was entirely composed of patients with glucocorticoid-induced osteoporosis attributable to rheumatoid arthritis (approximately 40%), polymyalgia rheumatic, systemic lupus erythematosus and asthma; however, the auditor ultimately decided to include this study.

Summary of study characteristics

The auditor and reviewer's summary of study characteristics are provided in Table 19. Again it should be noted that the reviewer's summary is a very comprehensive narrative and not intended for tabulation, where possible, relevant information has been extracted and placed alongside the auditors for ease of direct comparison. The information extracted by the reviewer and the auditor was largely aligned.

Table 19 Table of characteristics of included studies

	First Author, Year, Country	Study Design, Length of Follow-up	Patient Characteristics, Sample Size (n)	Intervention (n)	Comparator(s) (n)	Clinical Outcomes
Auditor	Orwoll, 2010 [66], North America	Multicentre, double-blind, double-dummy, active-controlled RCT Restrictions around prior bisphosphonate usage, only 3 men had received a bisphosphonate at any time previously 2 years	Mean age approx. 64 years, 100% male, well matched for baseline T-score at hip and femoral neck, history of fracture in 63-70% of patients, 41 patients withdrew	Once-yearly IV infusion of zoledronic acid 5mg Daily calcium (1000mg) and vitamin D (800 to 1000 IU) was also provided N=154	Weekly oral alendronate 70mg capsule Daily calcium (1000mg) and vitamin D (800 to 1000 IU) was also provided N=148	• Percentage change in lumbar spine BMD at 24 months • Percentage change in lumbar spine BMD at 6 and 12 months • Percentage change in hip, femoral neck, lumbar spine, trochanter and total body at 6, 12 and 24 months • Biochemical markers of bone formation • Adverse events • Patient preference questionnaire

Reviewer			The study population included 302 men aged 25 to 85 years old with a BMD *T*-score of -2.0 SD at the femoral neck and -1.0 SD at the lumber spine, or -1.0 at the femoral neck with a prior low trauma vertebral or non vertebral fracture or with a radiographic vertebral fracture identified during screening period.	Once-yearly 5 mg zoledronic acid intravenous infusion at day 1 and day 365, plus weekly oral placebo capsule or weekly (n = 154)	Weekly oral alendronate capsule (70 mg) plus yearly placebo intravenous infusion (n = 148)	The primary efficacy outcome was expressed in terms of percentage change in BMD of the lumbar spine from baseline to month 24, with an intention to examine whether zoledronic acid was not inferior to alendronate. Secondary efficacy outcome measures included percentage change in BMD at month 6 and 12 months from the baseline; percentage change in BMD at the total hip, femoral neck, lumber spine, trochanter, and total body at month 6, 12, and 24 months from the baseline, and changes in the levels of markers of bone resorption (β-CTx and urine NTx) and formation (serum P1NP and serum BSAP) during the course of the treatment. Adverse events (AEs) along with bone safety, serum chemistry and renal safety were monitored and recorded. Also, patients' treatment preference was examined using a questionnaire.

Auditor	Reid 2009 [65], 54 centres in 12 European countries, Australia, Hong Kong, Israel, and the USA	Multicentre, randomized, double-blind, double-dummy, non-inferiority study Patients taking glucocorticoids within 3 months were included in the prevention subgroup, longer than 3 months were included in the treatment subgroup Patients with previous treatment with bisphosphonates or other drugs that affect the skeleton were excluded unless specific washout undertaken 1-year	Mean age 53-56 years, 68% female of whom 66% were menopausal, 14% patients in the treatment subgroup had fractures at baseline, 62 patients withdrew with reasons fully explained	5 mg intravenous infusion of zoledronic acid administered once yearly plus daily oral placebo plus 400–1200 IU per day vitamin D and 1 g per day calcium starting up to 28 days before the infusion and continuing throughout Treatment n=272 Prevention n=144	5 mg daily dose of risedronate plus one placebo intravenous infusion on day 1 plus 400–1200 IU per day vitamin D and 1 g per day calcium starting up to 28 days before the infusion and continuing throughout Treatment n=273 Prevention n=144	• Percentage change from baseline in lumbar spine BMD at 12 months • Percentage change from baseline in BMD for total hip, femoral neck, trochanter, and distal radius • Occurrence of thoracic and lumbar vertebral fractures at 12 months • Changes in bone turnover biomarker concentrations • Adverse events • Renal function • HRQoL (EQ-5D)

Reviewer		Patients who had been taking glucocorticoids within 3 months prior to trial were included in the prevention subgroup and those who had been taking glucocorticoids for longer than 3 months were included in the treatment subgroup.	The study involved 833 women and men aged 18 to 85 years (68% of patients were women) from 54 centers located in 12 European countries, Australia, USA, Hong Kong and Israel.	For both subgroups (treatment and prevention) patients were randomized to receive 5 mg zoledronic acid intravenous infusion on day 1 plus daily oral placebo or daily 5 mg oral risedronate and one placebo intravenous infusion on day 1. In the treatment subgroup, 272 patients randomly received annual zoledronic acid infusion and 273 received oral risedronate. In the prevention subgroup 144 patients randomly received zoledronic acid infusion and 144 received oral risedronate. The duration of the trial was one year	The primary efficacy measure was the percentage change in BMD of the lumbar spine (L1-L4) at 12 months relative to baseline…Percentage changes from baseline in BMD of total hip, trochanter, distal radius, and femoral neck were secondary outcome measures. Other endpoints included changes in the concentration of bone turnover biomarker from baseline and assessment of AEs renal impairment and health-related quality of life (HRQL) measured using visual analogue and utility score techniques.

Critical appraisal

Although some information relevant to critical appraisal was included in the summary of findings for each study as captured in Table 21 (for example information about power calculations, sample size and use of double-dummy procedures), the reviewer did not provide an explicit critical appraisal of the included studies (tabulated or otherwise). In the 'Limitations' section, the reviewer provided the following comment, *"The studies differ in terms of population, comparators, and types of osteoporosis, which makes it difficult to compare the studies. For example, Reid et al 2008 studied men and women (18 to 85 years old) whereas Orwoll et al 2015 examined men (25 to 85 years old) and McClung et al 2007 studies postmenopausal women (45 to 79 years old). Reid et al 2008 included only glucocorticoid-induced osteoporosis patients and therefore the generalizability of this study to all osteoporosis patients could be limited."* A summary of the auditor's critical appraisal is characterised in Table 20.

Table 20 Summary of auditor's critical appraisal of included studies

	Reference	Strengths	Limitations
Auditor	Orwoll et al 2010 [66]	• Interventions clearly described • Inclusion criteria, efficacy variables and results described • Randomized and blinded using double-double technique • Statistical power calculations described • Baseline characteristics described • Withdrawals accounted for, ITT population and safety population defined for analyses • Missing data appropriately adjusted (LOCF)	• Method of randomisation not described
Auditor	Reid et al 2009 [65]	• Interventions clearly described • Randomisation via interactive voice response system • Blinded using double-double technique • Withdrawals accounted for, ITT population and safety population defined for analyses • Statistical power calculations described	

Summary of findings

The auditor and reviewer's summary of study characteristics are provided in Table 21. Again it should be noted that the reviewer's summary is a very comprehensive narrative and not intended for tabulation, where possible, relevant information has been extracted and placed alongside the auditors for ease of comparison. Similar information was extracted by the reviewer and the auditor.

Table 21 Table of study findings and authors conclusions

	Reference	Study findings	Author's conclusions
Auditor	Orwoll et al 2010 [66]	• Percentage change in lumbar spine BMD at 24 months (LOCF): LSM estimate of percentage increase were 6.1% with zolendronate and 6.2 with alendronate (LSM difference 0.13%, 95% CI 1.12 to 0.85, p=0.79) • Percentage change in lumbar spine BMD at 6 and 12 months: LSM difference 0.46% (95%CI 0.62 to 0.55, p=0.27) at 6 months; LSM difference 0.77% (95%CI 1.62 to 0.08, p=0.07) at 12 months • Percentage change in hip, femoral neck, trochanter and total body at 6, 12 and 24 months: Hip 6 months 0.03% (95% CI 0.61 to 0.55%), 12 months 0.26% (95% CI 0.9 to 0.37), 24 months 0.57% (95%CI 1.29 to 0.30); femoral neck 6 months 0.97% (95% CI 0.79 to 2.73%), 12 months 0.62% (95% CI 0.79 to 2.74), 24 months 0.57% (95%CI 1.0 to 2.16); trochanter 6 months 0.09% (95% CI 0.66 to 0.84%), 12 months 0.23% (95% CI 1.10 to 0.63), 24 months 0.60% (95%CI 1.68 to 0.47) • Biochemical markers of bone formation: At 12 and 24 months both treatment resulted in reductions in levels of bone resorption and formation • Adverse events: overall incidence similar in both groups (93.5% v 93.2%) with myalgia, chills, fatigue, malaise, backache, pain and influenza-like illness frequently observed; other data reported • Patient preference questionnaire: 74.2% preferred once-yearly infusion, 15.3% preferred weekly oral alendronate and 10.5% expressed no preference	In men with osteoporosis, an annual IV infusion of zoledronic acid 5mg was similarly effective in increasing BMD and suppressing bone turnover markers to weekly alendronate …patient adherence to therapy for 12 months after infusion, coupled with patient preference, the zolendronic acid regimen may offer and attractive therapeutic option.

Reviewer		The result showed that patients in both treatment groups experienced increased BMD at lumbar spine, total hip, femoral neck, and trochanter over the 24-month treatment period. Mean percentage increase in lumbar spine BMD for zoledronic acid group and alendronate group was 6.1% and 6.2%, respectively. Zoledronic acid was neither superior nor inferior to alendronate, demonstrated by the lower bounds of 95% CI of absolute difference in percentage change, which were higher than -1.5%. In both treatment groups levels of markers of bone resorption and formation were reduced at 12 and 24 months, and more pronounced results occurred in the zoledronic acid group at 3, 6, 15 and 18 months than in the alendronate group. At the end of the 24-month treatment period, new fracture incidence was reported in 4 patients (2.4%) taking zoledronic acid and in 6 patients (4%) taking alendronate. The authors found “no differences between the treatment groups with respect to the proportion of participants who responded to therapy.” Of 275 patients who responded to the questionnaire on treatment preference, 204 patients (74.2%) preferred annual zoledronic acid intravenous infusion, whereas 42 patients (15.3%) preferred oral alendronate 70 mg once a week, and 29 patients (10.5%) were indifferent. Incidence of AEs (Table 1) was reported in 93.5% of patients taking zoledronic acid and in 93.2% in patient taking alendronate, with myalgia, chills, fatigue, malaise, backache, pain and influenza-like illness frequently observed (>5%). These AEs were observed within 3 days of treatment in at least 5% of patients receiving zoledronic acid and after 3 days in the alendronate group. Three days after treatment, 135 (88.2%) and 134 (90.5%) patients experienced AEs in the zoledronic acid and alendronate group, respectively. Similar frequency of severe adverse events (SAEs) was observed for the zoledronic acid group and alendronate group, and it was reported that the SAEs in the two groups were not meaningfully different.	Authors concluded that once-yearly intravenous infusion of zoledronic acid 5 mg was equally effective in increasing BMD and reducing bone turnover markers to weekly oral alendronate 70 mg in men with osteoporosis.

Auditor	Reid et al 2009 [65]	• Percentage change from baseline in lumbar spine BMD at 12 months: by 12 months zoledronic acid had increased lumbar spine bone mineral density more than had risedronate in both the treatment (LSM 4·06% [SE 0·28] vs 2·71% [SE 0·28], mean difference 1·36% [95% CI 0·67 to 2·05]) and prevention subgroups (2·60% [0·45] vs 0·64% [0·46], 1·96% [1·04 to 2·88]) • Percentage change from baseline in BMD for total hip, femoral neck, trochanter, and distal radius: zoledronic acid significantly increased bone mineral density at the femoral neck compared with risedronate, in both the treatment (1·45% [0·31, n=247 patients] vs 0·39% [0·30, n=239], 1·06% [0·32 to 1·79]) and prevention (1·30% [0·45, n=126] vs –0·03% [0·46, n=135], 1·33% [0·41 to 2·25]) subgroups. Similar findings were reported for the trochanter and total hip • Occurrence of thoracic and lumbar vertebral fractures at 12 months: With the treatment and prevention subgroups combined, the frequency of new vertebral fractures was very low for patients receiving both zoledronic acid (n=5) and risedronate (n=3), with no significant difference between drug groups • Changes in bone turnover biomarker concentrations: reductions in biomarkers at 12 months were significantly greater in patients on zoledronic acid than in those on risedronate in both the treatment and prevention subgroups • Adverse events: overall occurrence of adverse events was significantly higher in the zoledronic acid group in both the treatment and prevention subgroups, mainly caused by a higher frequency of symptoms (eg, influenza like illness, pyrexia) that were reported within 3 days of starting the drug. After 3 days, the occurrence of adverse events was similar in the two drug groups. The frequency of serious adverse events recorded by the investigators was similar between drug groups • Renal function: Confirmed adjudicated clinically significant renal events occurred in nine patients given zoledronic acid and six given risedronate, all but one of which were reversible. • HRQoL (EQ-5D): no significant differences between drug groups apart from utility score at 3 months in the prevention subgroup	A single 5 mg intravenous infusion of zoledronic acid is non-inferior, possibly more effective, and more acceptable to patients than is 5 mg of oral risedronate daily for prevention and treatment of bone loss that is associated with glucocorticoid use. One IV infusion of zoledronic acid provides greater increases in BMD and more rapid and substantial decreases in bone turnover than daily risedronate.

Reviewer		The results showed that both zoledronic acid and risedronate increased lumbar spine BMD in treatment and prevention subgroups. Once-yearly zoledronic acid intravenous infusion was not inferior to oral daily risedronate. Both treatment arms reduced levels of bone turnover markers; however, the reductions were significantly greater in the zoledronic acid group than the risedronate group. The study also found that "overall occurrence of adverse events was significantly higher in the zoledronic acid group in both the treatment and prevention subgroups, mainly caused by a higher frequency of symptoms (e.g., influenza-like illness, pyrexia) that were reported within 3 days of starting the drug." Table 3 provides summary of adverse events between drug groups and Table 4 shows types and frequency of serious AEs.	The study concluded that "a single 5 mg intravenous infusion of zoledronic acid is non-inferior, possibly more effective, and more acceptable to patients than is 5 mg of oral risedronate daily for prevention and treatment of bone loss that is associated with glucocorticoid use."

Conclusions and implications

The reviewer's section on conclusions and implications for decision making for decision or policy summarized the studies found and also provided additional information regarding a safety warning affecting zoledronic acid IV infusion. The reviewer's key messages were as follows *"Evidence suggest that once-yearly 5 mg zoledronic acid intravenous infusion is not inferior to oral bisphosphonates in the treatment of osteoporosis; however, zoledronic acid infusion may be associated with higher incidence of serious adverse events, renal dysfunction and impairment in particular."* The auditor did not disagree but had used a slightly different approach, the auditor considered that once-yearly 5mg zoledronic acid administered as an IV infusion may be as effective (non-inferior) for the prevention and treatment of bone loss by producing greater increase in BMD and decrease in bone turnover compared to risedronate and alendronate in patients who have not received previous treatment. Furthermore adverse events associated with zoledronic acid were typically higher during the first three days following administration but were comparatively similar overall.

Learnings

The importance of documenting a clear and unambiguous research question has been confirmed.

7.3. Rasburicase for Adults with Acute Tumor Lysis Syndrome (TLS): A Review of Clinical and Cost Effectiveness and Safety (RC0441)

Rasburicase was already funded for paediatric patients with acute TLS but not for adults who are typically treated with allopurinol or hydration. It was suggested that as rasburicase has a faster onset of effect than either allopurinol or hydration to prevent renal failure, this treatment could also be more beneficial for adult patients.

Auditor Comments

Refining the research question

A summary of the research question as defined by the original reviewer and the auditor is presented in Table 22. The research question as it was presented in the TRF was clear. The auditor added additional detail around the handling of mixed adult/paediatric and mixed TLS prophylaxis/TLS treatment populations.

Table 22 A comparison of PICO(D) as captured by original reviewer and the auditor for RC0441

PICO(D)	Original reviewer	Auditor	Comment
Population	Adult patients with acute TLS	Adult patients with a confirmed diagnosis of acute TLS, mixed populations were included provided results for adult patients were reported separately from paediatric (and treatment from prophylaxis)	Additional assumptions added for clarity regarding mixed populations

PICO(D)	Original reviewer	Auditor	Comment
Intervention	Rasburicase; also called urate oxidase and Fasturtec	Rasburicase (urate oxidase, Fasturtec)	-
Comparator	Q1 and 2: Allopurinol, hydration (likely intravenous), placebo, or no comparator Q3: Allopurinol or hydration	Q1 and 2: Allopurinol, hydration (likely intravenous), placebo, or no comparator Q3: allopurinol or hydration	-
Outcomes	Reduction or elimination of need for hemodialysis Safety: adverse events Cost-effectiveness	Reduction or elimination of need for hemodialysis Safety Cost-effectiveness	-
(Design)	HTA/systematic review/meta-analysis; RCTs; NRS; economic evaluations	HTA/systematic review/meta-analysis; RCTs; NRS; economic evaluations	-

Study selection

The final published CADTH review confirmed that one RCT, four prospective studies, and two retrospective studies met inclusion criteria. The reviewer did not identify any economic evaluations. In contrast, and despite the apparent clarity of the research question as described above, the auditor selected only one study. The reason for this difference in study selection was an apparent post-hoc (or at least post-TRF) change in the project protocol. Whereas the TRF and the published research questions specified adult patients with acute TLS, the purpose of the review as defined in the published Context and Issues section and the PICO selection criteria also included patients "at risk" of TLS. Similarly, the reviewer seemed to have included studies of the treatment or prophylaxis of hyperuricemia[14] as well as TLS[15]. The research questions (as found on the TRF and as published in CADTH report) do not make any mention of hyperuricemia nor the "at risk" population, the research questions are stated below.

Specific differences in study selection are summarized below:

- Based on abstracts, five studies included by the reviewer were excluded by the auditor as they included patients at risk of TLS but seemingly without a current diagnosis of TLS [67] [68] [69] [70] [71]
- Jeha et al 2005[16] was included by the reviewer but excluded by the auditor as there was no mention of TLS in the title or abstract [72]

[14] Hyperuriceamia is an excess of uric acid in the blood; TLS refers to the constellation of metabolic disturbances that may follow the initiation of cancer treatment and is characterized by rapid development of hyperuricemia, along with hyperkalemia, hyperphosphatemia, hypocalcemia, and acute renal failure. Accessed 03/03/15 http://emedicine.medscape.com/article/282171-overview.

[15] *"In all studies, rasburicase was administered for the treatment or prophylaxis of hyperuricemia (Bosly et al 2003; Jeha et al 2005; Trifilio et al 2011) or TLS (Cortes et al 2010; Wang et al 2006, Coiffer et al 2003, Hummel et al 2008)"* page 4 of published CADTH review.

[16] Full paper not checked as reference not provided.

Summary of study characteristics, Critical appraisal, Main study findings and Author conclusions

Information extracted by the reviewer is presented alongside the information extracted by the auditor in Table 23 for the one study that was commonly selected (Hummel et al 2008 [73]). The reviewer provided additional remarks within a well-written narrative commentary of all of the reviewer's included studies. Where possible, these comments have been captured below. The reviewer also provided tabulated information. Data for Hummel et al 2008 [73] as presented by the auditor and the reviewer is summarized in Table 23 (Summary of study characteristics), Table 24 (Critical appraisal), and Table 25 (Main study findings and Author conclusions). As can be seen from these tables, overall the auditor and reviewer extracted very similar information and made similar comments regarding critical appraisal, there were some minor differences in level of detail observed but this did not affect the interpretation of the information presented.

Conclusions and implications

As the reviewer and auditor had included a very different selection of studies, it was difficult to directly compare conclusions and implications. Both the auditor and the reviewer noted the same results and conclusions from the one common study (namely that rasburicase for the treatment of TLS was effective at potentially lower doses than conventionally used).

Perhaps the most important question is whether, based on independent study selection, the reviewer and auditor ultimately drew similar conclusions and implications for decision-making. The reviewer's section on conclusions and implications for decision making for decision or policy was necessarily different due to the inclusion of more studies alongside and comments regarding reducing hyperuricemia and the risk of TLS (which was not addressed by the auditor). The reviewer's key findings were as follows *"Rasburicase was found to lower uric acid levels in adult cancer patients who had, or were at risk for developing, hyperuricemia or tumour lysis syndrome (TLS). Rasburicase was found to be well tolerated in adult cancer patients, with a low rate of adverse events, renal failure, or need for hemodialysis. No evidence on the cost effectiveness of rasburicase compared with allopurinol or hydration for adults with acute TLS was identified."* The auditor agreed that the available evidence (albeit now from only a very small single study of limited methodological quality and potentially at a higher risk of bias) suggested that rasburicase was an effective treatment for TLS (by lowering uric acid levels) in adult cancer patients and that the treatment was well-tolerated with no patient requiring renal replacement therapy. In the auditor's view there was no evidence available for rasburicase at a standard dose as the only included study had evaluated lower doses. The auditor agreed that no evidence on cost effectiveness versus allopurinol or hydration was identified. Although reducing the quantity (and the quality) of information available, the key findings (whilst much less robust) are not remarkably different in content.

Learnings

The critical learning from this evaluation was regarding the impact of undocumented post-hoc changes to the protocol (or TRF). Whilst undoubtedly increasing the amount of relevant information and making for a better reflection of the nature and severity of disease (as a severe life-threatening condition, prophylaxis for those at risk is preferable to treatment), the post-TRF inclusion of 'at risk' patients meant that the auditor was not able to match the reviewers study selection.

Table 23 Table of characteristics of included studies

	First Author, Publication Year, Country	Study Design, Length of Follow-up	Patient Characteristics, Sample Size	Intervention (n)	Comparator(s) (n)	Clinical Outcomes
Auditor	Hummel 2008 [73], Germany	Retrospective NRS January 2002 and 31 July 2006.	Median age 67 yr (range 16–88), 21 patients were female, 46 had hematologic malignancy, TLS was already present in 29 patients, TLS induced in 13 patients, 8 patients were treated for prophylaxis TLS categorised according to Cairo and Bishop as laboratory TLS (n=8) and clinical TLS (n=34)	Rasburicase (dose at physician discretion; median total dose 0.049 mg⁄kg) + IV hydration +/- sodium bicarbonate if urine pH was low	No comparator	• Uric acid level within 24 hrs after first dose • Creatinine (renal function) • Renal replacement therapy • Adverse events • Drug costs
Reviewer		Retrospective study Single-center Jan 2002-July 2006	50 consecutive cancer patients (ALL, AML, CLL, CMPD, NHL) treated with rasburicase for clinical or laboratory TLS (n=42) or for the prophylaxis of TLS (n=8) – median age 67 years (range 16-88), 58% male	Rasburicase, various doses (median overall dose 3 mg) + IV hydration		Uric acid level within 24 hours after first dose, creatinine levels

ALL=acute lymphoblastic leukemia; AML=acute myeloid leukemia; CLL=chronic lymphocytic leukemia; CMPD=chronic myeloproliferative disease; NHL=non-Hodgkin's lymphoma; LDH=lactate dehydrogenase; RCT=randomized controlled trial; TLS=tumour lysis syndrome

Table 24 Summary of auditor's critical appraisal of included studies

	Reference	Strengths	Limitations
Auditor	Hummel 2008 [73]	• Clearly defined diagnosed TLS (laboratory, clinical) and prophylactic patient groups	• Small (n=50) retrospective uncontrolled analysis • Lower than recommended dosages used • Outcomes not clearly listed and poorly defined
Reviewer		• Study reflected clinical practice, patients had high uric acid levels and impairment of renal function	• Retrospective analysis • Small sample size (N=50) • Varying doses of rasburicase administrated • Rasburicase was administered with hydration • Outcomes not clearly defined (normal uric acid range, normal creatinine range)

Table 25 Table of main study findings and authors conclusions

		Main study findings	Author's conclusions
Auditor	Hummel 2008 [73]	• Uric acid level within 24 hrs after first dose: in the majority of patients the UA level was below 475.8 lmol⁄L after the first dose, those receiving between1.5 and 4.5 mg as a first dose were divided by UA level below 475.8 lmol⁄L and above after the first dose; in those above this threshold, the UA was significantly at baseline • Renal replacement therapy: No patient required renal replacement therapy • Adverse events: was well tolerated by all patients, no adverse events occurred • Drug costs: Treatment costs were reduced by 96.8%	Our data support the use of rasburicase at lower doses than recommended by the manufacturer for prophylaxis and treatment of TLS. Applying low doses of rasburicase with close monitoring of UA levels and repeating further doses as required, allows efficient and cost-effective treatment and prophylaxis of TLS.

		Main study findings	Author's conclusions
Reviewer		Dose of rasburicase, mg (mg/kg) Prophylaxis group (n=8): 3.75 (0.056) Treatment group (n=42): 3 (0.044) This study used a uric acid level of 475.8 µmol/L (5.4 mg/dL) as a divide. Patients who had a uric acid level of >475.8 µmol/L after rasburicase treatment had a statistically significantly higher uric acid level before treatment than patients who achieved a uric acid level of <475.8 µmol/L (P=0.0270). Baseline creatinine was elevated in 42 patients. In this subgroup, median creatinine levels decreased from 206 to 118.5 µmol/L within 7 days after rasburicase administration. No patients in this subgroup were on dialysis at the time of rasburicase administration. No patients required renal replacement therapy. No adverse events occurred in this study.	"This is the first study to demonstrate the efficacy of low doses of rasburicase for prophylaxis and treatment of TLS in a patient cohort with markedly elevated serum uric acid levels and a large proportion of patients with impaired renal function. Our data support the use of rasburicase at lower doses than recommended by the manufacturer for prophylaxis and treatment of TLS." (p. 335)

7.4. Endovascular Thermal Ablation Technologies for Treatment of Varicose Veins: A Review of Clinical Effectiveness, Safety, Cost-Effectiveness and Guidelines – An Update (RC0570)

Endovascular thermal ablation technologies (EVLT) had previously been assessed and the requestor desired an update of the previous guidance in order to inform regional provincial policy on these procedures.

Auditor Comments

Refining the research question

A summary of the research question as defined by the original reviewer and the auditor is presented in Table 26. The research question as it was presented in the TRF was clear and the auditor made no additional assumptions.

Table 26 A comparison of PICO(D) as captured by original reviewer and the auditor for RC0570

PICO(D)	Original reviewer	Auditor	Comment
Population	Patients with varicose veins Subpopulation: working age patients with varicose veins	Patients with varicose veins	_

Intervention	Endovascular thermal ablation (EVTA) – includes endovascular laser therapy (EVLT) and radio frequency ablation (RFA)	Endovascular thermal ablation (EVTA) including endovascular laser therapy (EVLT/EVLA) and radio frequency ablation (RFA)	_
Comparator	EVLT and RFA versus standard treatment (surgery and sclerotherapy); EVLT versus RFA	Standard treatment (surgery and sclerotherapy); EVLT versus RFA	_
Outcomes	Clinical benefits, clinical harms, cost-effectiveness, guidelines and recommendations	Clinical benefits, clinical harms, cost-effectiveness, guidelines and recommendations	_
(Design)	HTA/systematic review/meta-analysis; RCTs; NRS; economic evaluations; guidelines	HTA/systematic review/meta-analysis; RCTs; NRS; economic evaluations; guidelines	_

Study selection

The published CADTH report stated eleven studies were included (one HTA, two systematic reviews, four RCTs, three clinical practice guidelines, and one recommendation report). Given the number of systematic reviews that have been published and the fact that a previous CADTH report was already available, identifying studies that had been not already been included in either of these sources was a time-consuming task. Multiple RCTs, one clinical practice guideline, one non-randomized study and one systematic review were already included in systematic reviews and/or noted in the previous version of the CADTH report. The auditor identified ten studies (one HTA, four systematic reviews, three RCTs, one clinical practice guideline and one recommendation report) from the project bibliography. The reviewer excluded non-randomized studies post-hoc without explanation, presumably based on the volume of higher quality evidence retrieved (although the previous version of the CADTH report included non-randomized studies); the auditor had identified three potentially relevant non-randomized studies but on the basis of the reviewer's post-hoc exclusion, these studies are not discussed further.

Other specific differences are described below:

1. Based on abstracts the auditor identified two additional systematic reviews [74] [75] that appeared to be wholly or partially relevant to the research question on the basis of published abstracts, these studies are described in Table 27
2. Two clinical practice guidelines included by the reviewer were attributed to internet searches as they did not appear in the project bibliography and these are not discussed further [76] [77]
3. One RCT [78] was included by the reviewer but excluded by the auditor on the basis that it was included in a Cochrane review that was also already included [79] [17]

[17] Two studies were excluded in the same Cochrane but were included by both the reviewer and the auditor [83] [84]

Table 27 Two systematic reviews excluded by the reviewer but included by the auditor

Reference	PICO(D)	Details	Author conclusions
Burihan MC. Endovenous ablation (radiofrequency and laser) and foam sclerotherapy versus conventional surgery for great saphenous vein varices. Sao Paulo Med J. 2014;132(1):69.	Population	Great saphenous varicose veins	*Currently available clinical trial evidence suggests RFA and EVLT are at least as effective as surgery in the treatment of great saphenous varicose veins. There are insufficient data to comment on USGFS. Further randomized trials are needed. We should aim to report and analyze results in a congruent manner to facilitate future meta-analysis*
	Intervention	EVLT, RFA, USGFS	
	Comparator	Conventional surgery	
	Outcomes	Recurrent varicosities, recanalization, neo-vascularization, technical procedure failure or need for re-intervention, patient quality of life (QoL) scores and associated complications.	
	(Design)	SR of RCTs	
Siribumrungwong B, Noorit P, Wilasrusmee C, Attia J, Thakkinstian A. A systematic review and meta-analysis of randomised controlled trials comparing endovenous ablation and surgical intervention in patients with varicose vein. Eur J Vasc Endovasc Surg. 2012 Aug;44(2):214-23.	Population	Great saphenous varicose veins	*The primary failure and recurrence in EVLA and RFA were non-significantly different compared with surgery. However, they had lower haematoma, less wound infection, less pain and quicker return to normal activities*
	Intervention	EVLA, RFA, UGFS, surgery	
	Comparator	Any listed intervention	
	Outcomes	Rate of primary failure, clinical recurrence, infection	
	(Design)	SR of RCTs	

Summary of study characteristics

Information extracted by the reviewer is presented alongside the information extracted by the auditor in Table 28 for the studies that were commonly selected. The reviewer provided additional remarks within a well-written narrative commentary of all of the reviewer's included studies across the three research questions. Accounting for differences in study selection, the reviewer and auditor's summary of study characteristics were largely aligned with some differences in level of detail and style of extraction as shown below.

Critical appraisal

Information extracted by the reviewer is presented alongside the information extracted by the auditor in Table 29 for the studies that were commonly selected. The reviewer provided additional remarks within an impressive narrative commentary of all of the reviewer's included studies across the three research

questions. Accounting for differences in study selection, the reviewer and auditor's critical appraisal were largely aligned with some differences in level of detail and style of extraction.

Main study findings

Information extracted by the reviewer is presented alongside the information extracted by the auditor in Table 30 for the studies that were commonly selected. The reviewer provided additional remarks within an impressive narrative commentary of all of the reviewer's included studies across the three research questions. Accounting for differences in study selection, the reviewer and auditor's summary of study findings were largely aligned with some differences in level of detail and style of extraction as shown below.

Conclusions and implications

The reviewer's conclusions and implications for decision or policy making are presented below. As this is a comprehensive summary the auditor's comments are annotated in square brackets for transparency.

"Most of the studies included in this review compared non-invasive procedures with surgery. Few reports have addressed the comparison between EVLT and RFA. [Agreed] *However, those were reports of high quality level like HTA, SRs and exhaustive CPGs* [One SR was of questionable quality [80]]. *But given the nature of the interventions, RCTs included in the present review and other reviews were open-labeled studies, limiting the strength of the available evidence.* [Agreed that lack of blinded evidence is not ideal]

Most of the available evidence showed similar or slight differences in clinical effectiveness between EVLT, RFA, UGFS and surgery although some studies found effectiveness benefits with non-invasive procedures [Auditor concluded there was no evidence to suggest non-invasive techniques resulted in significantly improved clinical effectiveness]. *Surgery was associated with more pain compared to RFA and longer convalescence, higher risks of infection, or sensory problems when compared with non-invasive treatments.* [Agreed] *The decrease of clinical severity and the increase of quality of life observed after treatment were comparable with all the reviewed procedures.* [Agreed] *Patient satisfaction was also similar.* [Agreed]

Cost-effectiveness advantages over surgery had been attributed to EVTA and UGFS, respectively, in two different good quality reports from UK. However, discrepancies between these two studies in terms of cost-effectiveness at a threshold of £20,000 were observed for EVTA procedures. Therefore, cost-effectiveness of EVTA over surgery is not clear. [Disagreed, but a consequence of one of these studies not being included by the auditor, auditor concluded that EVLT and RFA unlikely to be cost-effective at an acceptable threshold based on this single study] *Taken together, these economic studies highlight the cost-effectiveness sensitivity to local costs input and assumptions as well as their questionable applicability to the Canadian context.* [Agreed – limited relevance to Canada]

…In conclusion, non-invasive procedures, like EVLT, RFA and UGFS, are not inferior to surgery with potential benefits in terms of pain, time to return to normal activity, complications and cost-effectiveness [Agreed apart from cost effectiveness]. *Hence, our findings are in accordance with Ontario Health Technology Advisory Committee recommending implementation with guidance on their clinical eligibility. Little or no clinical effectiveness or safety differences between non-invasive procedures have been observed. Cost is more likely to determine cost-effectiveness between them* [Agreed]*."*

Based on abstracts the two systematic reviews that had been additionally included by the reviewer also concluded that EVLT and RFA were as least as effective as surgery, one study also concluded that there may be additional benefits such as lower rates of infection, less pain and a more rapid return to activities of daily living. With the exception of the differences regarding economic cost effectiveness whereby the inclusion by the auditor of only one out of two of the studies included by the reviewer generated a slightly different view, the reviewers and auditors conclusions and implications are aligned.

Table 28 Table of characteristics of included studies

		Study Design, Length of Follow-up	Patient Characteristics, Sample Size (n)	Intervention	Comparator(s)	Clinical Outcomes
Auditor	Carroll 2013 [81], UK	Clinical: systematic review and meta-analyses with searches across multiple databases for RCTs only, extraction performed by one reviewer ad checked by a second, critical appraisal based on published tool. Date of search: July 2011 Economic: same methods as clinical with no limitation on study design, studies had to report economic outcomes in terms of cost-effectiveness, cost–utility or cost–benefit. Primary economic analyses conducted alongside clinical trials were assessed using the checklist by Drummond et al, modelling studies using a checklist modified from Eddy Date of search: September 2012	Adults aged ≥16 years being treated specifically for varicose veins. No minimum duration of follow-up. Clinical: 34 different studies included a total of 3873 participants, aged from 33 to 54 years, the percentage of female participants ranged from 54% to 95%. All participants were required to have varicose veins diagnosed by duplex scanning and categorised according to the CEAP score, the vast majority scored C2. Approximately half of the included studies reported inadequate randomisation, allocation concealment, between-group comparability or intention-to-treat analyses. Economic: two RCT-based economic analyses and two modelling analyses were included.	Endovenous laser ablation, RFA, FS and TIPP.	Any form of varicose veins management	• Failure of the procedure • Second or further procedures • Technical recurrence • Second or further procedures • Symptomatic recurrence • Clinical symptoms, as measured by the VCSS • Pain • Time to return to work or normal activity. • Post-operative complications (adverse events) • Cost effectiveness

Reviewer		Clinical: included literature up to July 2011. Economic literature: included CEA, CUA, CBA, up to Sept 2012, an economic model has been developed.	Included English language RCTs, patients 16 years of age and older. No minimal duration of follow-up.	EVLT, RFA, foam sclerotherapy, transilluminated-powered phlebectomy	Any form of varicose veins management	Clinical: • Failure of the procedure • Recurrence • Clinical symptoms measured by the VCSS • Pain • Time to return to work or normal activity Safety: • Post-operative complications (adverse events) Cost-effectiveness
Systematic reviews						
Auditor	Nesbitt 2014 [79], UK	Update of previous review, systematic review and meta-analyses with searches across multiple databases for RCTs only, extraction independently extracted by two reviewers, critical appraisal based on Cochrane tool. Date of search: January 2014	Any age with varicose veins affecting the GSV system, confirmed on duplex ultrasound imaging, who were suitable for any of the treatment options. 13 trials prospective RCTs including 3081 patients were included, patients were aged 18 to 79 years and predominantly female	• Foam sclerotherapy • Laser endovenous ablation (EVLT) • Radiofrequency endovenous ablation (RFA) • Saphenofemoral junction ligation and stripping of the great • saphenous vein (GSV)	Not specified	• Recurrence or recanalisation • Ultrasound evidence of neovascularisation • Technical failure • Patient satisfaction determined by quality of life (QoL) • Post-operative complications • Length of the procedures • Hospital stay • Procedural costs

Reviewer		SR/MA of RCTs on the treatment of GSV varices. Update of 2011 Cochrane review, included literature up to January 2014.	13 studies, 3081 patients. 3 studies compared UGFS vs surgery, 8 EVLT vs surgery, 5 RFA vs surgery. Sample size range from 28 to 390 patients. Mean age range: 33 to 56 years. Female % range: 50 to 93.	EVLT, RFA, UGFS	Surgery (HLS)	• Recurrent varicosities (clinical and symptomatic) • Recanalisation • Neovascularisation • Technical failure • QoL scores • Complications
Auditor	Tellings 2011 [80], Netherlands	Systematic review with meta-analysis with searches across multiple databases for RCTs and NRS (including case series and qualitative opinions) Date of search: not reported	All surgical and endovenous techniques were included provided they use ultrasound examination to qualify outcome, studies published in English, German, Spanish, French and Dutch were eligible for inclusion. 17 studies were included, five reported on different types of surgery (3 stripping, 1 stripping versus ligation alone and 1 ligation with stab avulsion), 10 on EVLA and two on UGFS.	All treatments	All treatments	• Outcome • Complications • Patient satisfaction • Cost–time effectiveness
Reviewer		SR of all studies on the treatment of SSV insufficiency	17 reports: 5 surgery, 10 EVLT, 2 ultrasound-guided foam sclerotherapy	All treatments	All treatments	• Clinical effectiveness • Patient satisfaction • Complications • Cost-time effectiveness
RCT						

Auditor	Mozafar 2014 [82], Iran	Randomized, controlled single centre study For each patient, only one limb was included in the study. For patients with bilateral disease, the leg with the most severe symptoms was included. 18 months	Mean age was 38.9±9.31, approx. 70% female, patient demographics in both groups did not show any significant difference prior to intervention N=65	EVLT N=30	High ligation of saphenous vein (HLS) N=35	• Recurrence • CEAP score • AVVSS score • Patient satisfaction • Adverse events
Reviewer		RCT, open-label, Length of follow-up: up to 18 months	EVLT n = 30 patients, 73% females Surgery n = 35 patients, 71 % females Mean age: 39 years, patients had GSV or saphenofemoral joint (SFJ) insufficiency with reflux and symptoms or chronic venous insufficiency	EVLT	Surgery (High ligation of saphenous vein)	• Clinical recurrence • Severity (CEAP staging, VCSS score) • Patient satisfaction
Auditor	Lattimer 2013 [83], UK	Randomized, controlled single centre study Median 15 months (interim report only)	Mean age 47-50, 54-61% female, 191 assessed for eligibility, 110 randomised, only 44 EVLA and 46 UGFS patients had completed follow-up at the interim report	EVLA + phlebectomy N=44	UGFS N=46	• GSV occlusion • VCSS score • AVVQ score • Saphenous treatment score • Recurrent reflux • Adverse events
Reviewer		RCT, open-label, Length of follow-up: 15 months (preliminary results)	EVLT + phlebectomy (n = 44), mean age : 47, 61% women. UGFS (n = 46), mean age: 50, 54% women. Patients had GSV venous reflux.	EVLT + phlebectomy	Ultrasound- guided foam sclerotherapy	• GSV occlusion • Severity (VCSS, STS) • QoL (AVVQ)

Auditor	Samuel 2013 [84], UK	Randomized, controlled single centre study Patients with primary, symptomatic, unilateral varicose veins, with isolated SPJ incompetence, causing reflux into the SSV 1 year	Mean age 47.5-.8 years, 64-75.5% female, demographics were comparable between groups. 767 assessed for eligibility and 106 randomised.	EVLT N=53	Surgery (ligation and stripping) N=53	• Abolition of SSV reflux at 6 weeks • Safety • Pain • Patient satisfaction • CEAP score • VCSS score • HRQoL
Reviewer		RCT, open-label, Length of follow-up: up to 1 year	EVLT (n =53), 64% women. Surgery (n = 53) 76% women. Mean age: 48 Patients had unilateral	EVLT	Surgery (ligation and stripping)	• Abolition of SSV reflux • Pain scores • Recovery time • Complication rates • Severity (VCSS)
Clinical guidelines						
Auditor	Pavlovic 2014 [85], International	Clinical consensus guidelines	Guidelines created during consensus meeting under the auspices of the International Union of Phlebology (IUP).	Endovenous Thermal Ablation	Endovenous Thermal Ablation	• Efficacy • Safety • Tolerability • Cosmetic outcome • Patient satisfaction/ preference
Reviewer		CPG	Guideline drafted during consensus conference in collaboration with the International Union of Phlebology, based on a systematic review.	EVTA procedures	EVTA procedures	• Efficacy • Safety • Tolerability • Patient satisfaction/ preference • Cosmetic outcome
Recommendations						

Auditor	Ontario Health Advisory Committee 2013 [86], Canada	Clinical recommendations	Based on two published HTAs that compared nonsurgical endovascular ablation techniques with surgery by examining clinical effectiveness, safety, costs, and budgetary implications	EVLT, RFA	EVLT, RFA, surgery	• Safety profile • Impact on health-related quality of life (HRQOL) • Durability • Patient satisfaction • Effectiveness compared with surgical ligation and vein stripping
Reviewer		Recommendations	Based on 2 HTAs from Health Quality Ontario (2011 and 2010).	EVLT, RFA	EVLT, RFA, surgery (vein ligation + stripping)	• Effectiveness • Durability • Health-related quality of life • Patient satisfaction • Safety

AVVQ = Aberdeen varicose veins questionnaire; CBA = cost-benefits analysis; CEA = cost-effectiveness analysis; CEAP = Clinical-Etiology-Anatomy-Pathophysiology; CPG = clinical practice guidelines; CUA = cost-utility analysis; EVLT = endovenous laser therapy; EVTA = endovenous thermal ablation; GSV = great saphenous vein; HLS = high ligation and stripping; HTA = health technology assessment; MA = meta-analysis; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; QoL = quality of life; RCT = randomized controlled trial; RFA = radio frequency ablation; SFJ = saphenofemoral junction; SR = systematic review; SSV = small saphenous vein; STS = saphenous treatment score; UGFS = ultrasound-guided foam sclerotherapy; USA = United States of America; VCSS = Vascular Clinical Severity Score; w/o = without.

Critical appraisal

Table 29 Table of critical appraisal of included studies

		Strengths	Limitations
Auditor	Carroll 2013 [60], UK	• Comprehensive scoping and planning for scope and methods • Clinical o Clearly stated research questions, methods, critical appraisal, data extraction and recommendations • Economic o Clearly stated research questions, critical appraisal, treatment options and details such as perspective, drug cost, time horizon, discounting and sensitivity analyses	UK specific
Reviewer		• Review (clinical and economic): Clear description of a priori design, literature search, duplicate study selection, selection criteria, list of all studies with their characteristics and appraisal. Conclusions reflected the quality of studies. Homogeneity of included studies has been addressed. • Economic model: Study had a well-defined question, description of the competing treatments and established effectiveness of the therapies. Perspective, time horizon, discounting were stated. Costs with their references were disclosed and appropriate. Sensitivity analyses were performed	• Review (clinical and economic): Publication bias has not been assessed. No declaration of conflict of interest or sources of funding. • Economic model: Applicability of costs from United Kingdom to Canada remains uncertain.
Systematic reviews			
Auditor	Nesbitt 2014 [79], UK	• Cochrane review • A priori design with complete description of methods • Details of excluded studies provided • Detailed consideration of potential sources of bias • Comprehensive presentation of study level data • Comprehensive presentation of meta-analysis	NA

Reviewer		A priori-designed SR with MAs. Clear description of literature search, duplicate study selection, list of included & excluded studies, their characteristics, their critical appraisal. Homogeneity and possibility of publication bias or conflict of interest have been assessed.	
Auditor	Tellings 2011 [80], Netherlands	• Inclusion and exclusion criteria clearly stated • Adequate description of searching methods • Outcomes transformed • Clear description of results • Limitations acknowledged	• No search dates provided • No details of study exclusions • No 'PRISMA' type flowchart • No description of role of reviewers or how data extraction handled • No critical appraisal of included studies (critical point given that non-RCTs were included) • Differences between studies acknowledged but data are still combined to yield overall estimate
Reviewer		Clear description of included studies and its characteristics. Homogeneity has been assessed.	A review protocol has not been mentioned. Years of literature search and mentioned. Inclusion of grey literature is unclear. List of excluded studies is not shown. Individual quality of studies was not described. Publication bias was not assessed. Conflicts of interest were not assessed. Studies included for cost-time effectiveness assessment were or poor quality.
RCTs			
Auditor	Mozafar 2014 [82], Iran	• Inclusion and exclusion criteria stated • Baseline characteristics described • No loss to follow-up	• Method of randomisation not described • Outcomes not clearly pre-specified and assessors were not blinded • No power calculations documented • Small sample size

Reviewer		Inclusion and exclusion criteria, subject characteristics and interventions were described. Measurement of reflux by duplex ultrasound deemed to be accurate. No loss to follow up.	Study subjects and people measuring study outcomes were not blinded. No statistical test for main outcome. No power calculation, very small samples size. % of enrolment not mentioned. No description of randomization procedures or whether it was concealed until recruitment.
Auditor	Latinmer 2013 [83], UK	• Clear description of interventions • Baseline demographics described • Power calculations reported • Outcomes specified	• Method of randomisation not described • Outcome assessors were not blinded • Interim report only
Reviewer		Clear description of subjects, outcomes, interventions, findings, actual P values. Measurement of reflux by ultrasound deemed to be accurate. Detailed description of different venous outcomes. Sample size calculation.	Study subjects and people measuring study outcomes were not blinded. EVLT group had more mild cases (C2), not taken into account for analysis. No mention of losses to follow-up (number, reasons, analysis). Source population and hospital settings are unclear. Randomization was disclosed in a previous publication.
Auditor	Samuel 2013 [84], UK	• Inclusion and exclusion criteria clearly stated • Outcomes clearly described and reported • Power calculations described • Baseline demographics described	• Outcome assessors not blinded to treatment • Methods for randomisation described • Large number assessed but not randomised, and loss to follow-up not explained
Reviewer		Clear description of subjects, randomization procedure, outcomes, interventions, findings, actual P values. Measurement of reflux by duplex ultrasound deemed to be accurate. Sample size calculation for main outcome.	Study subjects and people measuring study outcomes were not blinded. Patients lost to follow-up not described.
Guidelines			
Auditor	Pavlovic 2014 [85], International	• Recommendations graded by strength of evidence using published scale • Recommendations are comprehensive and clearly described	• Recommendations based on systematic review but methodology is not described adequately (no details of search, selection, critical appraisal, extraction etc.)
Reviewer		Recommendations were well described. Recommendations were graded depending of strength of available evidence. Studies were cited.	Methodology, health questions covered, composition of the development group, target users were poorly described. Assessment of bias in the covered literature has not been

		Patient preferences and side-effects have been considered. Authors declared no conflict of interest.	mentioned. Applicability has not been addressed. Not externally reviewed. No mention of an updating process.
Recommendations			
Auditor	Ontario Health Advisory Committee 2013 [86], Canada	• Recommendations based on published HTAs	• No new searches undertaken to update previous publications
Reviewer		Recommendations based on previous HTAs providing evidence on EVLT and RFA.	No update of literature or evidence

EVLT = endovenous laser therapy; HTA = health technology assessment; MA = meta-analysis; NHS = National Health Service; P = probability value; RCT = randomized controlled trial; RFA = radio frequency ablation; SR = systematic review

Table 30 Summary of study findings

		Main study findings	Author's conclusions
Auditor	Carroll 2013 [81], UK	• Failure of the procedure: reported proportion of initial failures was very small for all techniques. Second or further procedures: Where reported, retreatment consisted of stripping and ligation for RFA, or further sessions of sclerotherapy for FS or stripping • Second or further procedures: Very few studies reported reoperation rates beyond 1-month follow-up. • Technical recurrence: The risk of experiencing a technical recurrence of varicose veins over time was lower for EVLA (HR: 6 months 0.70; 1 year 0.77; 2 years 0.84) and RFA (HR: 6 months 0.92; 1 year 0.93; 2 years 0.94) than for ligation and stripping. The risk of experiencing a technical recurrence of varicose veins over time was initially higher for FS (HR: 6 months 1.12; 1 year 1.02) than for ligation and stripping, but lower after 2 years (HR 0.92). Symptomatic recurrence: Very few studies reported symptomatic recurrence beyond 1-month follow-up. • Clinical symptoms, as measured by the VCSS: Meta-analysis found lower post-intervention VCSS for both FS and EVLA than for stripping, but a slightly higher score for RFA than for stripping • Pain: lower post-operative pain for RFA than for stripping, as well as reduced pain for FS and a slightly increased level of pain for EVLA than for stripping • Time to return to work or normal activity: significantly quicker return to work or normal activity was reported by all relevant studies for both FS and RFA than for stripping, studies comparing EVLA and stripping reported either no difference or more rapid return to work for participants with EVLA • Post-operative complications (adverse events): There were no consistent or statistically significant differences between any of the interventions in terms of complications or adverse events. • Cost effectiveness: FS costs are £530 less than stripping, and it is marginally • More effective (+ 0.0015 QALYs), with a probability of being the most cost-effective treatment above 90% for willingness-to-pay thresholds in the range £20,000–50,000. EVLT and RFA both cost more than stripping (total costs +£1302 and + £1617, respectively) with little difference in QALYs (+ 0.0025 and + 0.0012) compared with stripping, with ICERs of £518,000 and £1,353,000, respectively, they cannot be considered cost-effective (robust to parameter variation and model time horizon)	This assessment of the currently available evidence suggests there is little to choose between the minimally invasive techniques in terms of efficacy or cost, and each offers a viable, clinically effective alternative to stripping. FS might offer the most cost-effective alternative to stripping, within certain time parameters. High-quality RCT evidence is needed. Future trials should aim to measure and report outcomes in a standardised manner, which would permit more efficient pooling of their results.

Reviewer		34 RCTs comprising 3,873 patients (range: 28 to 710 patients/trial) with VV; mean age range of 33 to 54 years; predominantly female (54 to 95%, depending on trial); majority of patients were C2 on CEAP score. 14 trials evaluated EVLT (8 vs surgery, 6 vs RFA, 1 vs UGFS); 13 trials evaluated RFA (6 vs surgery, 6 vs EVLT, 1 vs UGFS); 13 trials evaluated UGFS (10 vs surgery, 1 vs EVLT, 1 vs RFA) • Clinical effectiveness and safety: o Failure of procedure: EVLT: 5/467 (1%); RFA: 16/431 (4%); UGFS: 21/295 (7%); HLS: 20/681 (3%) o Risk of technical recurrence [HR (95% CrI)]: ▪ EVLT vs stripping: 6 mo: 0.70 (0.27 to 1.45); 1 y: 0.77 (0.37 to 1.54); 2 y: 0.84 (0.44 to 1.81) ▪ RFA vs stripping: 6 mo: 0.92 (0.39 to 2.11); 1 y: 0.93 (0.42 to 2.22); 2 y: 0.94 (0.42 to 2.51) ▪ UGFS vs stripping: 6 mo: 1.12 (0.53 to 2.27); 1 y: 1.02 (0.49 to 1.84); 2 y: 0.92 (0.43 to 1.60) o Symptomatic recurrence: Small number of reported events; no difference between groups o VCSS: UGFS vs stripping: -1.63 (-2.90 to -0.42), no difference between other groups. o Time to return to work/normal activity: 5 out of 7 studies favored RFA or UGFS vs surgery. o Pain: EVLT vs stripping: No difference between groups; RFA vs stripping: RFA favoured (median: –1.26 (95% CrI, –1.95 to –0.61); UGFS vs stripping: no difference. o Post-operative complications: Hematoma, paresthesia, infection, phlebitis were commonly reported, but overall event numbers were small. DVT and PE were rare. • Cost-effectiveness: o From SR: 4 economic studies identified (2 prospective analyses, 2 modeling analyses) ▪ Expected net benefits from different treatment approaches were similar, but sensitive to assumptions, creating uncertainty about relative CE. o From economic model: ▪ EVLT and RFA were more costly, while UGFS was less costly, than surgery with little difference in QALYs. ▪ Neither EVLT nor RFA were considered cost-effective compared with surgery at a threshold of £20,000 to £30,000. Robust model. ▪ UGFS was the most cost-effective with a probability of 90% at a threshold of £20,000-£50,000. Sensitive to time horizon. ▪ Between-treatment cost differentials were expected to vary by setting and time.	"This assessment of the currently available evidence suggests that there is little to choose between the minimally invasive techniques in terms of efficacy, and each offers a viable, clinical alternative to stripping. Based on data reviewed, only foam sclerotherapy offers a cost-effective alternative to stripping. Training and experience in the minimally invasive techniques might be required before more substantial, relative clinical benefits are apparent." (p. 69)

Systematic review			
Auditor	Nesbitt 2014 [79], UK	• UGFS versus surgery o Recurrence: Clinician noted: no difference, (OR 1.74, CI 0.97 to 3.12, P =0.06), symptomatic: no difference, OR 1.28, CI 0.66 to 2.49 o Recanalization: < 4 months: OR 0.66 (CI 0.20 to 2.12) > 4 months: OR 5.05 (CI 1.67 to 15.28). o Neovascularisation:OR 0.05 (CI 0.00 to 0.94). o Technical failure: no difference, OR 0.44 (CI 0.12 to 1.57) • EVLT versus surgery o Recurrence: Clinician noted: no difference, OR 0.72 (CI 0.43 to 1.22), symptomatic recurrence: no difference, OR 0.87 (CI 0.47 to 1.62) o Recanalisation: no difference, early: OR 1.05 (CI 0.09 to 12.77), late: OR 4.14 (CI 0.76 to 22.65, P =0.10) o Neovascularisation: reduced in EVLT with OR 0.05 (CI 0.01 to 0.22, P < 0.0001) o Technical failure: reduced in EVLT with OR 0.29 (CI 0.14 to 0.60, P = 0.0009) • RFA versus surgery o Recurrence: Clinician noted: no difference, OR 0.82 (CI 0.49 to 1.39), symptomatic recurrence (single study): no difference, OR 2.00 (CI 0.30 to 13.26) o Recanalisation: no difference, early: OR 0.68 (CI 0.01 to 81.18), late: OR 1.09 (CI 0.39 to 3.04) o Neovascularisation: no difference, OR 0.31 (CI 0.06 to 1.65) o Technical failure: no difference, OR 0.82 (CI 0.07 to 10.10) • Post-operative complications: generally low, especially major complications • Pain: reporting varied greatly between the studies but in general pain was similar between the treatment groups • HRQoL: quality of life generally increased similarly in all treatment groups • Length of the procedures: lack of congruity with the presented results prevented any meaningful meta-analysis • Hospital stay: the majority of patients were operated on as day cases. • Procedural costs: the costs involved for each study varied	Currently available clinical trial evidence suggests that UGFS, EVLT and RFA are at least as effective as surgery in the treatment of great saphenous varicose veins. Due to large incompatibilities between trials and different time point measurements for outcomes, the evidence is lacking in robustness. Further randomised trials are needed, which should aim to report and analyse results in a congruent manner to facilitate future meta-analysis.

Reviewer		• UGFS vs surgery: o Clinician noted recurrence: no difference, (OR 1.74, CI 0.97 to 3.12, P =0.06) o Symptomatic recurrence: no difference, (OR 1.28, CI 0.66 to 2.49) o Recanalisation (single study): o < 4 months: OR 0.66 (CI 0.20 to 2.12) > 4 months: OR 5.05 (CI 1.67 to 15.28). o Neovascularisation (single study): o OR 0.05 (CI 0.00 to 0.94). o Technical failure: no difference, OR 0.44 (CI 0.12 to 1.57). • EVLT vs surgery; o Clinician noted recurrence: no difference, OR 0.72 (CI 0.43 to 1.22). o Symptomatic recurrence: no difference, OR 0.87 (CI 0.47 to 1.62). o Recanalisation: no difference, early: OR 1.05 (CI 0.09 to 12.77), late: OR 4.14 (CI 0.76 to 22.65, P =0.10). o Neovascularization: reduced in EVLT with OR 0.05 (CI 0.01 to 0.22, P < 0.0001). o Technical failure: reduced in EVLT with OR 0.29 (CI 0.14 to 0.60, P = 0.0009). o Long-term (5 years) outcomes: (single study) similar findings between interventions. • RFA vs surgery: o Clinician noted recurrence: no difference, OR 0.82 (CI 0.49 to 1.39). o Symptomatic recurrence (single study): no difference, OR 2.00 (CI 0.30 to 13.26) o Recanalisation: no difference, early: OR 0.68 (CI 0.01 to 81.18), late: OR 1.09 (CI 0.39 to 3.04). o Neovascularisation: no difference, OR 0.31 (CI 0.06 to 1.65). o Technical failure: no difference, OR 0.82 (CI 0.07 to 10.10). • QoL scores, complications and pain: o Similar between groups.	"Currently available clinical trial evidence suggests that UGFS, EVLT and RFA are at least as effective as surgery in the treatment of great saphenous varicose veins. Due to large incompatibilities between trials and different time point measurements for outcomes, the evidence is lacking in robustness. Further randomised trials are needed, which should aim to report and analyse results in a congruent manner to facilitate future meta-analysis." (p. 4)

Auditor	Tellings 2011 [80], Netherlands	• Outcome o Surgery: five papers with success rates from 24% to 100%; one article showed no significant difference between strip and ligation alone (45% versus 35%) o EVLT: ten papers with success rates varied from 91% to 100% o UGFS: two papers success rates were 100% and 82%; SSV's with a diameter ≤5 mm showed higher occlusion rates (86% versus 77%) o Combined analysis: The difference in success is significant; surgery 47.8%, EVLA/UGFS 94.9%. • Complications o Surgery: Two studies show no major complications, three others report DVT in 1.8%, 2% and 3.5%, sural nerve damage is reported in one article (2.1%), paresthesia was found in a range of 1.7–34% o EVLT: DVT reported in 1.3-5.7% of patients, paresthesia in 1.3–11% of the o treated legs, phlebitis in 0–8% of the cases o UGFS: no major complications reported, thrombophlebitis and pigmentation o were common minor complications that resolved over time • Patient satisfaction: Six studies describe patient satisfaction after treatment • (4 on EVLA and 2 on surgery) and report a decline in symptoms after treatment • Cost–time effectiveness: One study compared EVLA with RFA and concluded that the choice depends on the cost of equipment, disposables and procedure time, two articles state, that cost differences are self-evident (UGFS versus surgery)	In the absence of large, comparative randomized clinical trials, minimally invasive techniques appear to have a tendency towards better results than surgery, in the treatment of the insufficient SSV. We demonstrated significant differences in outcomes, in favour of the minimally invasive techniques. However, it should be noted that the articles are very heterogeneous and therefore hard to compare.
Reviewer		• 17 studies (RCT, non-RCT) in SSV VV including 10 EVLT, 5 surgery (stripping and/or ligation) and 2 UGFS studies EVLT studies comprised a range of 37 to 390 legs and follow-up of 0.5 month to 3 years. Surgery studies included 52 to 204 legs with follow-up of 1.5 months to 5 years. UGFS studies included 23 and 141 legs and follow-up of 1.5 months and 11 months. o Success rates: Surgery ranged from 24% to 100%; EVLT ranged from 91% to 100%; UGFS ranged from 82% to 100%. Difference in success rate between surgery (47.8%) and EVLT/UGFS (94.9%), P < 0.05. o Major complications: Surgery: DVT (1.8% to 3.5%), sural nerve damage (2.1%); EVLT: DVT (1.3% to 5.7%); UGFS: none o o Paresthesia: Surgery: 1.7% to 34%; EVLT: 1.3% to 11%	"…lack of [published evidence]… specifically on the treatment of SSV insufficiency… (p. 183) "…the results in the articles published do not allow us to draw definite conclusions on the ideal treatment for SSV insufficiency." (p.183)
RCT			

Auditor	Mozafar 2014 [82], Iran	• Recurrence: At 18 months recurrence rate in the EVLT and HLS groups was 6.7 and • 11.7 %, respectively • CEAP score: similar improvements at 18 months • AVVSS score: significantly less in the EVLT group at 12 (p=0.019) and 18 months (p=0.08) of follow up • Patient satisfaction: Similar in both groups at 6, 12, and 18 months of follow-up • Adverse events: No significant difference was observed in major adverse events. A single case of infection was reported in the HLS group.	The results of our study further establish the efficacy of EVLT as an alternative to conventional treatment and expand those findings to a broader population base to include people of Middle Eastern decent. Furthermore, they show that EVLT offers better long-term symptom relief when compared to conventional surgical treatment. It also indicates that the two methods are not significantly different in other aspects. Yet, due to its more favorable cosmetic outcome, and less invasive nature, most patients are likely to choose EVLT for treatment
Reviewer		• 65 patients (EVLT: 30; HLS: 35) with GSV VV; mean age: 39 years; majority female (72%); 78% were C2 or C3 on CEAP score. • After 12 months: o Recurrence rate: EVLT: 6.7%; HLS: 11.7% o AVVSS score: Lower in EVLT group (P = 0.019) • After 18 months: o AVVQ score: Lower in EVLT group (P = 0.008) o CEAP score: Similar improvements in both groups after 1 week and sustained to 18 months. o No DVT reported in either group o Similar frequency of dysesthesia between groups (EVLT: 8.6%; HLS: 6.7%) • Patient satisfaction was similar in both groups.	"The results of our study further establish the efficacy of EVLT as an alternative to conventional treatment and expand these findings to a broader population base to include people of Middle Eastern decent." (p.770)
Auditor	Latinmer 2013 [83], UK	• GSV occlusion: 42/44 (95.5%) for EVLA and 31/46 (67.4%) for UGFS achieved success p=0.001 • VCSS score: Statistical reduction (P < 0.0005) but no significant difference in improvement between groups • AVVQ score: Statistical reduction (P < 0.0005) but no significant difference in improvement between groups, at 3 months, significantly favoured EVLA (p=0.19) • Saphenous treatment score: Statistical reduction (P < 0.0005) but no significant difference in improvement between groups	EVLA and UGFS are equally effective at abolishing global venous reflux with overall success of 41% and 43%, respectively. The high reflux rate was not related to deterioration in quality of life indicating that this reflux was largely asymptomatic

		• Adjuvant sclerotherapy: 4.7 more frequent with UGFS group than EVLT (47 vs 10) • Recurrent reflux: 18/44 (41%) EVLA and 20/46 (43%) UGFS avoided recurrent reflux • Adverse events: not reported by treatment	
Reviewer		• After 15 months: o Occlusion of GSV: EVLT 42/44 (95.5%) more effective vs UGFS 31/46 (67.4%). o Abolishment of global reflux: both equally effective. o Number of legs (n = 100, EVLT vs UGFS) with: total reflux abolition (18 vs 20), above-knee (6 vs 8), below-knee (12 vs 11), combined reflux (8 vs 7), loss to follow-up (6 vs 4). o Statistical reduction of VCSS, AVVQ and STS (P < 0.0005). o No difference between groups. • Overall need for adjuvant therapy (sclerotherapy): 4.7 more frequent in the UGFS group vs EVLT group (47 vs 10).	"EVLA and UGFS are equally effective at abolishing gobal venous reflux with overall success of 41% and 43%, respectively. The high reflux rate was not related to deterioration in quality of life indicating that this reflux was largely asymptomatic." (p. 394)
Auditor	Samuel 2013 [84], UK	• Abolition of SSV reflux at 6 weeks (technical success): significantly higher for 51 (96.2%) EVLA patients than 38 (71.7%) patients in the surgery group (P < 0.001); RR for EVLA early success 1·34 (1·11–1·44), risk difference of 0·24 (0·09–0·30; NNT for EVLA rather than surgery to avoid a residual refluxing SSV post-procedure was 4.0 (3.2–10.9). • Safety: sensory disturbance was significantly higher in the surgical group at 6-week follow up 14 (26.4%) v 4 (7.5%) patients in the EVLA group, P = 0.009, persistent sensory disturbance affected 5 (9.4%) surgery patients and 2 (3.7%) EVLA patients P = 0.434 at 1 year. A single major complication of DVT in the PV was recorded during the 1 week DUS evaluation post-surgery • Pain: Between days 4 and 7, pain scores were significantly lower in the EVLA group than in the surgical group (Day 4, P = 0.025; Day 5, P = 0.008; Day 6, P = 0.033; Day 7, P = 0.042) • Patient satisfaction: Patient satisfaction was equally high with either treatment • VCSS score: significant improvement over the follow-up period, from a baseline median (IQR) of 3 (2–4) to 0 (0–1) at the end of 12 months (P <0.001), no significant difference between the groups	EVLA produced the same clinical benefits as conventional surgery but was more effective in addressing the underlying pathophysiology and was associated with less peri-procedural morbidity allowing a faster recovery.

		• HRQoL: Both treatments produced a similar durable improvement in disease-specific and generic QOL scores over the study period (P < 0·001)	
Reviewer		• 106 patients/legs (EVLT: 53; surgery: 53) with SSV VV; mean age: 48 years; majority female (70%); 81% were C2 on CEAP score. • Pain and return to normal functioning: o Pain scores lower (P < 0.05) in the EVLT group vs surgery group from day 4 to day 7. o Patients returned to normal functioning more quickly after EVLT than surgery (P < 0.001). • After 6 weeks: o Abolition of SSV reflux: EVLT favored over surgery (96.2% vs 71.7%, P < 0.001) ▪ RR of early success with EVLT vs surgery: 1.34 (95% CI, 1.11 to 1.44); RD: 0.24 (95% CI, 0.09 to 0.30) ▪ NNT with EVLT to avoid one residual SSV post-procedure: 4.0 (95% CI, 3.2 to 10.9) o Sensory disturbance (especially sural nerve): More frequent with surgery than EVLT (26.4% vs 7.5%, P = 0.009) Most cases resolved after 1 year (P = 0.434). o Low frequency (EVLT vs surgery) of phlebitis (5.7% vs 1.9%), infection (0 vs 1.9%), hematoma (0 vs 3.8%), DVT (0 vs 1.9%). • After 1 year: o Clinical recurrence: Similar in surgery vs EVLT (16.9% vs 9.4%, P = 0.390) o VCSS: Similar improvement between groups. o QoL: ▪ AVVQ: Similar improvement between groups. ▪ SF-36 V1, EQ-5D: Similar improvement between groups.	"The immediate postoperative benefits and short-term technical outcomes of EVLT would support the future consideration of this procedure as the standard treatment of small saphenous insufficiency, provided the long-term results are no worse than following surgery." (p. 425)
Guidelines			
Auditor	Pavlovic 2014 [85], International	• Efficacy • Safety • Tolerability • Cosmetic outcome • Patient satisfaction/ preference	

Reviewer		• Only clinical evidence considered; no health economic guidance issued. • Veins indicated for EVTA (all GRADE I recommendations): o GSV o SSV o Accessory SV (intrafascial part) o Giacomini vein and cranial extension of SSV o Other superficial veins in subcutaneous tissue o Insufficient perforating veins o Residual intrafascial veins post-treatment o Venous malformations • While RFA has some specific requirements for vein segment length, EVLT does not. • To enable catheter advancement, EVTA requires that veins be free of synechiae or membrane webs or tortuosity. • Calculations are recommended for determining the appropriate energy for treatment by EVTA (GRADE IA). o RFA: energy delivery will vary by system employed (e.g., Closure FAST™, Celon™ system). o EVLT: Appropriate energy density is the main driver of success. • Major complications to consider in EVTA (GRADE IC): o DVT/PE (though reported post-procedure incidence low: 0-2%) o Damage to arteries (e.g., arterial fistulas – very rare) o Severe nerve damage (very rare) o Skin burns (especially when treated without tumescence) o Infection o Intra-procedural fiber breakage o Stroke (based on single case report) • Minor complications to consider in EVTA (GRADE IC): o Pain o Bruising o Erythema o Hematoma o Hyperpigmentation o Paresthesias o Tender or non-tender palpable treated vessel (especially thigh GSV)	

		o Infection o Telangiectatic matting • EVTA is often performed with adjunctive phlebectomy or UGFS. • During one procedure, EVTA may be carried out on ≥ 2 incompetent veins.	
Recommendations			
Auditor	Ontario Health Advisory Committee 2013 [86], Canada	• ELT and RFA are less invasive, safe, and cost-effective alternatives to surgical vein stripping that should be made available to people with symptomatic varicose veins (VV) and saphenous venous reflux demonstrated on a full duplex ultrasound investigation and, when feasible, following a failed trial of conservative management • There is an absolute medical necessity for a surgical approach including RFA or ELT treatment of VV associated with venous ulcer, thrombophlebitis, or bleeding. However, the decision to recommend a similar treatment approach based on other symptoms attributed to chronic venous reflux should be made on an individual basis and guided by validated disease severity scales such as the Venous Clinical Severity Score • Any intervention for VV for cosmetic indications should not be provided as an insured service • Mechanisms to ensure quality assurance for both the physicians performing endovascular • treatments and the facility where the treatments are being performed should be considered as part of any implementation plan	See Recommendations
Reviewer		• EVLT and RFA are less invasive, safe and cost-effective alternatives to surgery for treatment of symptomatic VV with saphenous reflux. • Should be made available when bleeding, thrombophlebitis, venous ulcer. Chronic venous reflux also included if based on severity scale like VCSS. • Cosmetic intervention should not be publicly funded. • Quality assurance mechanism should be implemented.	

AVVQ = Aberdeen Varicose Vein Questionnaire; CEAP = clinical status, etiology, anatomy, pathophysiology scale; CI = confidence interval; CrI = credible interval; DVT = deep vein thrombosis; EVLT = endovenous laser therapy; EVTA = endovenous thermal ablation (includes EVLT and RFA); GSV = great saphenous vein; HLS = high ligation and stripping; HR = hazard ratio; ICER = Incremental cost-effectiveness ratio; mo = month; NNT = number needed to treat; P = probability value; QALY = quality-adjusted life year; HRQoL = quality of life; RCT = randomized controlled trial; RD = risk difference; RFA = radiofrequency ablation; RR = relative risk; SR = systematic review; SSV = small saphenous vein; STS = saphenous treatment score; SV=saphenous veins; UGFS = ultrasound-guided foam sclerotherapy; U/S = ultrasound; VCSS = Venous Clinical Severity Score; vs = versus

8. Discussion

The goal of this evaluative series was to challenge the reproducibility of key elements of published CADTH Rapid Response reports in terms of methodological quality and transparency. The auditor was not necessarily 'right' any more than the reviewer was necessarily 'wrong' (and vice versa). Auditor-reviewer differences can be regarded as a signpost that more detail or additional information is needed (indeed within a systematic review these differences would be part of the process and may even involve a third reviewer as an adjudicator). In some instances minor differences could be regarded as artifactual and simply an inevitable consequence of two different reviewers acting independently.

As summarized for each CADTH product level evaluated in Table 31, Table 32, and Table 33, differences did arise and typically were the result of lack of definition in the original research question.

Table 31 Summary of objectives and results for the three Level 1 Rapid Response reports appraised

Element for assessment	Criteria for assessment	RA0611 Medical marijuana	RA0627 Robotic IV preparation	RA0677 Reprocessing SUDs
Research questions	Clear, unambiguous, comprehensive?	Insufficient	Insufficient	Yes
Study selection	Reviewer and auditor in agreement?	No (+3)	Yes (0)	Mostly aligned (+1)

Research question categories applied -Yes: PICO did not require any additional assumptions; acceptable: one PICO element required additional auditor assumptions; insufficient: two or three PICO elements required additional auditor assumptions; Not clear: four or all PICO elements required additional author assumptions. Numbers in brackets indicate the number of studies where differences were noted. Study selection and interpretation categories applied – Yes: 0 differences; mostly aligned: differences in 1-2 studies; No - differences in 3 or more studies

Table 32 Summary of objectives and results for the three Level 1.5 Rapid Response reports appraised

Element for assessment	Criteria for assessment	RB0520 Oncotype DX testing	RB0654 Patient mobilisation	RB0721 Automated external defibrilators
Research question(s)	Clear, unambiguous, comprehensive?	Acceptable	Insufficient	Yes
Study selection	Reviewer and auditor in agreement?	Mostly aligned (-1)	Could not be completed	Mostly aligned (-1)
Interpretation of common studies	Reviewer and auditor in agreement?	Mostly aligned (2)	Yes	Yes
Summary of common studies	Reviewer and auditor in agreement?	Mostly aligned (2)	Mostly aligned (1)	Yes

Research question categories applied - Yes: PICO did not require any additional assumptions; acceptable: one PICO element required additional auditor assumptions; insufficient: two or three PICO elements required additional auditor assumptions; Not clear: four or all PICO elements required additional author assumptions. Numbers in brackets indicate the number of studies where differences were noted. Study selection and interpretation categories applied – Yes: 0 differences; mostly aligned: differences in 1-2 studies; No - differences in 3 or more studies

Table 33 Summary of objectives and results for the four Level 2 Rapid Response reports appraised

Element for assessment	Criteria for assessment	L0161 Tinnitus retraining	L0227 IV zoledronic acid	RC0441 Rasburicase in adults	RC0570 EVLT for varicose veins
Research question(s)	Clear, unambiguous, comprehensive?	Yes	Insufficient	Acceptable	Yes
Study selection	Reviewer and auditor in agreement?	Mostly aligned (-1)	Mostly aligned (-1)	No (-6)	Mostly aligned (-1/+2)
Critical appraisal of common studies	Reviewer and auditor in agreement?	Yes	No[18]	Yes	Yes
Data extraction of common studies	Reviewer and auditor in agreement?	Yes	Yes	Yes	Yes
Synthesis of evidence	Data clear, missing or over-represented?	NA	NA	NA	NA
Conclusions and implications for decision making	Valid and appropriate conclusions and interpretation?	Yes	Yes	Mostly aligned	Yes

Yes: PICO did not require any additional assumptions; acceptable: one PICO element required additional auditor assumptions; insufficient: two or three PICO elements required additional auditor assumptions; Not clear: four or all PICO elements required additional author assumptions. Numbers in brackets indicate the number of studies where differences were noted. Study selection and interpretation categories applied – Yes: 0 differences; mostly aligned: differences in 1-2 studies; No - differences in 3 or more studies. NA = not applicable

Other potential explanatory factors in considering the differences that were documented include:

- References attributable to the grey literature or hand searching (these elements were not considered by the auditor and although every effort was made to account for these studies and negate their impact as these do not reflect true differences)
- Amongst the Level 2 reports - that the auditor was undertaking study selection based upon title and abstract only whereas the reviewer had presumably access the full publications, it is feasible that the reviewer had considered the additional studies included by the auditor but discarded them for reasons not apparent from the abstract alone
- Some differences may be simply artifactual and a consequence of the reviewer and auditor acting independently; unless resource-intensive steps are taken (double-data extraction, use of a third reviewer for example) two independent reviewers acting in isolation from one another will inevitably lead to differences; a third reviewer would no doubt make decisions that differ from the auditor should the exercise be repeated

[18] Although information relevant to critical appraisal was included in the summary of findings for each study as captured in Table 21 (for example information about power calculations, sample size and use of double-dummy procedures), the reviewer did not provide an explicit critical appraisal of the included studies (tabulated or otherwise).

- Undocumented post-hoc changes to the project scope; this was not observed frequently and most often referred to the post-hoc exclusion of non-randomized studies but when this did arise in review RC0441 it had a dramatic impact

The detailed instructions to authors provided by CADTH meant that when studies were commonly considered there were minimal differences between reviewer and auditor in specific tasks (critical appraisal, data extraction and presentation, conclusions and interpretation[19]). As previously explained, whether or not the same studies were considered was a direct consequence of the level of detail provided in the research question. The need to sufficiently refine the research question is a common goal for systematic reviewers and rapid reviewers alike; however, as frequently observed herein, when combined with limited background information, what may appear sufficient and even detailed can become inadequate when actually applied to a list of potentially relevant references.

The original reviewer was likely involved in query and discussion with at least the topic requestor and information scientist making it likely the reviewer's knowledge of the background context to the project exceeded that captured on the TRF. This becomes problematic when external auditor queries arise, as this knowledge cannot be used to infer missing information in the same way. Requirements for detailed and complete terminology (including synonyms) should be anticipated particularly when the published literature is likely to be inconsistent. Additional assumptions or caveats around key terms should also be anticipated (for example, how to treat a mixed adult/paediatric population, a blurred diagnostic definition, or a composite outcome).

8.1. Recommendations for program improvements

In terms of recommendations for program improvements and communications about the limitations of different types of review, one element stands out clearly regarding the project TRF. In its current form, the TRF is undoubtedly a crucial document and the requirements are (mostly) easy to understand, interpret and fulfill. It is interesting to consider whether this document is intended for use in the same way as a traditional protocol wherein pre-hoc methods and decisions are captured and intended to remain fixed and unchanged in order to reduce the risk of bias. At initial review, each and every TRF considered was detailed and unambiguous; however, once the auditor began to consider the therapy area in more detail and/or use the research question to select studies problems began to appear. Regardless of current use, it is recommended that the TRF (either in its entirety or a dedicated section) become a living document with an expanded background section that actively captures any iterative features of the review process. Adding to the document to capture refinements and decisions would enhance reproducibility and create a record of the knowledge and understanding of the reviewer at the time of the review.

[19] Some differences in level of detail and style of presentation were observed but this is to be expected

9. Conclusion

It is important to consider how high we should seek to set the reproducibility threshold for rapid reviews, as the line between ultimate quality and coverage (a traditional systematic review) and a rapid review (with some methodological trade-offs made to improve timeliness) is a very fine one. As described in the Background to this report, there is much variability in the methods of rapid review and although increasing demand has fuelled expansion, this situation has yet to improve. Within this context CADTH are to be commended on the clarity, consistency, and transparency associated with each of the products offered and also for a continued focus on quality and improvement.

This evaluation has shown that, all things considered, the current process is working and the results are largely reproducible. Where the process and results were not reproducible, the reasons are mostly identifiable and understandable. Whilst a marker of a robust and objective process, reproducibility may not the ultimate goal for rapid reviews; indeed, it would be interesting to see how many published systematic reviews held up against such an evaluation.

10.Bibliography

1. Hailey D, J.: Reuse of single use medical devices in Canada: clinical and economic outcomes, legal and ethical issues, and current hospital practice.. Int J Technol Assess Health Care. 24(4), 430-6 (2008)
2. Gough D, T.: Clarifying differences between review designs and methods. Syst Rev. 1(28) (2012)
3. Khangura S, K.: Evidence summaries: the evolution of a rapid review approach. Syst Revs 1(10) (2012)
4. Khangura S, P.: Rapid review: an emerging approach to evidence synthesis in health technology assessment. Int J Technol Assess Health Care. 30(1), 20-7 (Jan 2014)
5. Shea BJ, G.: Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews.. BMC Med Res Methodol 7(10) (2007)
6. Downs SH, B.: The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health 52(6), 377-84 (1998)
7. Drummond MF, S.: Methods for the economic evaluation of health care programmes. Oxford University Press, Oxford
8. Lynch ME, C.: Cannabinoids for treatment of chronic non-cancer pain; a systematic review of randomized trials. Br J Clin Pharmacol. 72(5), 735-44 (Nov 2011)
9. Curtis A, C.: Cannabinoids for Tourette's Syndrome.. Cochrane Database Syst Rev., CD006565 (April 2009)
10. Martin-Sanchez E, F.: Systematic review and meta-analysis of cannabis treatment for chronic pain. Pain Med. 10(8), 1353-68 (Nov 2009)
11. ang T, C.: Adverse effects of medical cannabinoids: a systematic review. CMAJ 178(13), 1669-78 (Jun 2008)
12. Richards BL, W.: Neuromodulators for pain management in rheumatoid arthritis.. Cochrane Database Syst Rev. , CD008921 (January 2012)
13. Phillips TJ, C.: Pharmacological treatment of painful HIV-associated sensory neuropathy: a systematic review and meta-analysis of randomised controlled trials. PLoS ONE. 5(12), e14433. (2010)
14. Seger AC, C.: Impact of robotic antineoplastic preparation on safety, workflow, and costs. J Oncol Pract. 8(6), 344-9 (Nov 2012)
15. Dehmel C, B.: Do centrally pre-prepared solutions achieve more reliable drug concentrations than solutions prepared on the ward? Intensive Care Med. 37(8), 1311-6 (Aug 2011)
16. Hornberger J, A.: Clinical validity/utility, change in practice patterns, and economic implications of risk stratifiers to predict outcomes for early-stage breast cancer: a systematic review.. J Natl Cancer Inst. 104(14), 1068-79 (Jul 2012)
17. Vanderlaan BF, B.: Cost-effectiveness of 21-gene assay in node-positive, early-stage breast cancer. Am J Manag Care. 17(7), 55-64 (2011)

18. Bacchi CE, P.: Potential economic impact of the 21-gene expression assay on the treatment of breast cancer in Brazil. Rev Assoc Med Bras. 56(2), 186-91 (Mar 2010)

19. Kondo M, H.: Economic evaluation of the 21-gene signature (Oncotype DX) in lymph node-negative/positive, hormone receptor-positive early-stage breast cancer based on Japanese validation study (JBCRG-TR03). Breast Cancer Res Treat. 127(3), 739-49 (Jun 2011)

20. Klang SH, H.: Economic implications of 21-gene breast cancer risk assay from the perspective of an Israeli-managed health-care organization.. Value Health. 13(4), 381-7 (Jun 2010)

21. Ont Health Technol Assess Ser: Gene expression profiling for guiding adjuvant chemotherapy decisions in women with early breast cancer: an evidence-based and economic analysis. Available at: http://www.health.gov.on.ca/english/providers/program/mas/tech/reviews/pdf/gep_20101213.pdf

22. Hall PS, M.: Economic evaluation of genomic test-directed chemotherapy for early-stage lymph node-positive breast cancer.. J Natl Cancer Inst. 104(1), 56-66 (Jan 2012)

23. Lamond NW, S.: Cost-utility of the 21-gene recurrence score assay in node-negative and node-positive breast cancer.. Breast Cancer Res Treat. 133(3), 1115-23 (Jun 2012)

24. Retel VP, J.: Head-to-head comparison of the 70-gene signature versus the 21-gene assay: cost-effectiveness and the effect of compliance. Breast Cancer Res Treat 131(2), 627-36 (Jan 2012)

25. Yang M, R.: Cost effectiveness of gene expression profiling for early stage breast cancer: A decision-analytic model.. Cancer (Feb 2012)

26. Hornberger J, C.: US insurance program's experience with a multigene assay for early-stage breast cancer.. J Oncol Pract 7(3), e38s-e45s (May 2011)

27. Tsoi DT, I.: Cost-effectiveness analysis of recurrence score-guided treatment using a 21-gene assay in early breast cancer.. Oncologist 15(5), 457-65 (2010)

28. Dublin (Ireland): National Centre for Pharmacoeconomics; : Cost-effectiveness of Oncotype DX® to target chemotherapy use in lymph-node-negative, oestrogen-receptor-positive, early-stage breast cancer in Ireland.

29. Padula CA, D.: Impact of lower extremity strengthening exercises and mobility on fall rates in hospitalized adults.. J Nurs Care Qual. 26(3), 279-85 (Jul 2011)

30. Berney SC, H.: Intensive care unit mobility practices in Australia and New Zealand: a point prevalence study.. Crit Care Resusc. 15(4), 260-5 (Dec 2013)

31. Nydahl P, R.: Early Mobilization of Mechanically Ventilated Patients: A 1-Day Point-Prevalence Study in Germany. Crit Care Med. 17 (Dec 2013)

32. Kristensen VG, N.: Mobilisation of epistaxis patients--a prospective, randomised study documenting a safe patient care regime. J Clin Nurs. 20(11-12), 1598-605 (Jun 2011)

33. Dickinson S, T.: Can the use of an early mobility program reduce the incidence of pressure ulcers in a surgical critical care unit?. Crit Care Nurs Q. 36(1), 127-40 (Jan 2013)

34. Kosse NM, D.: Effectiveness and feasibility of early physical rehabilitation programs for geriatric hospitalized patients: a systematic review. BMC Geriatr 13(107) (2013)

35. Doherty-King B, Y.: Frequency and duration of nursing care related to older patient mobility.. J

Nurs Scholarsh 46(1), 20-7 (2014)

36. Krist L, D.: Can progressive resistance training twice a week improve mobility, muscle strength, and quality of life in very elderly nursing-home residents with impaired mobility? A pilot study. (Accessed 2013) Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3639017 PubMed: PM23637524

37. Adler J, M.: Early mobilization in the intensive care unit: a systematic review.. Cardiopulm Phys Ther J 23(1), 5-13 (Mar 2012)

38. Dantas CM, S.: Influence of early mobilization on respiratory and peripheral muscle strength in critically ill patients.. Rev Bras Ter Intensiva. 24(2), 173-8 (Jun 2012)

39. Bourdin G, B.: The feasibility of early physical activity in intensive care unit patients: a prospective observational one-center study.. Respir Care. 55(4), 400-7 (Apr 2010)

40. Berdowski J, K.: Survival and health care costs until hospital discharge of patients treated with onsite, dispatched or without automated external defibrillator.. Resuscitation 81(8), 962-7 (Aug 2010)

41. Folke F, L.: Location of cardiac arrest in a city center: strategic placement of automated external defibrillators in public locations. Circulation 120(6), 510-7 (Aug 2009)

42. Stammet P, C.: Impact of initial intervention on long-term neurological recovey after cardiac arrest: data from the Luxembourg "North Pole" cohort.. Bull Soc Sci Med Grand Duche Luxemb 1, 60-70 (2012)

43. Weisfeldt ML, S.: Survival after application of automatic external defibrillators before arrival of the emergency medical system: evaluation in the resuscitation outcomes consortium population of 21 million.. J Am Coll Cardiol. 55(16), 1713-20 (Apr 2010)

44. Nichol G, H.: Cost-effectiveness of lay responder defibrillation for out-of-hospital cardiac arrest. Ann Emerg Med. 54(2), 226-35 (Aug 2009)

45. Martin CW, W.: Tinnitus retraining therapy. (Accessed 2004) Available at: http://www.worksafebc.ca/health_care_providers/Assets/PDF/tinnitus_retraining_therapy.pdf

46. Savage J, C.: Tinnitus. (Accessed 2009) Available at: http://clinicalevidence.bmj.com/

47. Forti S, C.: Are results of tinnitus retraining therapy maintained over time? 18-month follow-up after completion of therapy. Audiol Neurootol. 14(5), 286-9 (2009)

48. Bessman P, H.: The tinnitus intensive therapy habituation program: a 2-year follow-up pilot study on subjective tinnitus. Rehabil Psychol 54(2), 133-7 (May 2009)

49. Hanley PJ, D.: Treatment of tinnitus with a customized, dynamic acoustic neural stimulus: clinical outcomes in general private practice. Ann Otol Rhinol Laryngol. 117(11), 791-9 (Nov 2008)

50. Hatanaka A, A.: Pros and cons of tinnitus retraining therapy. Acta Otolaryngol (Stockh) 128(4), 365-8 (Apr 2008)

51. Herraiz C, H.: Long-term clinical trial of tinnitus retraining therapy. Otolaryngol Head Neck Surg. 133(5), 774-9 (Nov 2005)

52. Herraiz C, H.: Tinnitus retraining therapy: prognosis factors. Am J Otolaryngol. 28(4), 225-9

(Jul 2007)

53. Mazurek B, F.: A modified version of tinnitus retraining therapy: observing long-term outcome and predictors. Audiol Neurootol. 11(5), 276-86 (2006)

54. Hiller W, H.: Does sound stimulation have additive effects on cognitivebehavioral treatment of chronic tinnitus? Behav Res Ther 43(5), 595-612 (May 2005)

55. Henry JA, L.: Randomized clinical trial: group counseling based on tinnitus retraining therapy. J Rehabil Res Dev. 44(1), 21-32 (2007)

56. Seydel C, H.: Long-term improvement in tinnitus after modified tinnitus retraining therapy enhanced by a variety of psychological approaches. Audiol Neurootol. 15(2), 69-80 (2010)

57. Gudex C, S.: Effectiveness of a tinnitus management programme: a 2-year follow-up study. BMC Ear Nose Throat Disord 26(9), 6 (Jun 2009)

58. Ito M, S.: Association between tinnitus retraining therapy and a tinnitus control instrument. Auris Nasus Larynx 36(5), 536-40 (Oct 2009)

59. Aazh H, M.: Simplified form of tinnitus retraining therapy in adults: a retrospective study. BMC Ear Nose Throat Disord 8(7) (Nov 2008)

60. Caffier PP, H.: Outcomes of long-term outpatient tinnituscoping therapy: psychometric changes and value of tinnitus-control instruments. Ear Hear 27(6), 619-27 (Dec 2006)

61. Zachriat C, K.-H.: Treating chronic tinnitus: comparison of cognitivebehavioural and habituation-based treatments. Cognitive Behav Ther 33(4), 187-98 (2004)

62. Davis PB, W.: Treatment of tinnitus with a customized acoustic neural stimulus: a controlled clinical study. Ear Nose Throat J 87(6), 330-9 (Jun 2008)

63. Henry JA, S.: Outcomes of clinical trial: tinnitus masking versus tinnitus retraining therapy. J Am Acad Audiol 17(2), 104-32 (Feb 2006)

64. McClung M, R.: Intravenous zoledronic acid 5 mg in the treatment of postmenopausal women with low bone density previously treated with alendronate. Bone 41(1), 122-8 (Jul 2007)

65. Reid DM, D.: Zoledronic acid and risedronate in the prevention and treatment of glucocorticoid-induced osteoporosis (HORIZON): a multicentre, double-blind, double-dummy, randomised controlled tria. Lancet 373(9671), 1253-63 (Apr 2009)

66. Orwoll ES, M.: Efficacy and safety of a once-yearly i.v. Infusion of zoledronic acid 5mg versus a once-weekly 70-mg oral alendronate in the treatment of male osteoporosis: A randomized, multicenter, double-blind, active-controlled study. J Bone Miner Res 25(10), 2239-50 (Oct 2010)

67. Cortes J, M.: Control of plasma uric acid in adults at risk for tumor lysis syndrome: efficacy and safety of rasburicase alone and rasburicase followed by allopurinol compared with allopurinol alone--results of a multicenter phase III study. J Clin Oncol 28(27), 4207-13 (Sep 2010)

68. Wang LY, S.: Recombinant urate oxidase (rasburicase) for the prevention and treatment of tumor lysis syndrome in patients with hematologic malignancies. Acta Haematol. 115(1-2), 35-8 (2006)

69. Coiffier B, M.: Efficacy and safety of rasburicase (recombinant urate oxidase) for the prevention and treatment of hyperuricemia during induction chemotherapy of aggressive non-

Hodgkin's lymphoma: results of the GRAAL1 (Groupe d'Etude des Lymphomes de l'Adulte Trial on Rasburicase Activity in Adult Lymphoma) study. Clin Oncol 21(23), 4402-6 (Dec 2003)

70. Bosly A, S.: Rasburicase (recombinant urate oxidase) for the management of hyperuricemia in patients with cancer: report of an international compassionate use study. Cancer 98(5), 1048-54 (Sep 2003)

71. Trifilio SM, P.: Effectiveness of a single 3-mg rasburicase dose for the management of hyperuricemia in patients with hematological malignancies. Bone Marrow Transplant 46(6), 800-5 (2011)

72. Jeha S, K.: Efficacy and safety of rasburicase, a recombinant urate oxidase (Elitek), in the management of malignancy-associated hyperuricemia in pediatric and adult patients: final results of a multicenter compassionate use trial. Leukemia 19(1), 34-8 (2005)

73. Hummel M, R.: Effective treatment and prophylaxis of hyperuricemia and impaired renal function in tumor lysis syndrome with low doses of rasburicase.. Eur J Haematol. 80(4), 331-6 (Apr 2008)

74. Siribumrungwong B, N.: A systematic review and meta-analysis of randomised controlled trials comparing endovenous ablation and surgical intervention in patients with varicose vein.. Eur J Vasc Endovasc Surg. 44(2), 214-23 (Aug 2012)

75. MC., B.: Endovenous ablation (radiofrequency and laser) and foam sclerotherapy versus conventional surgery for great saphenous vein varices.. Sao Paulo Med J. 132(1), 69 (2014)

76. Expert Panel on Interventional Radiology. In: Radiologic management of lower-extremity venous insufficiency. (Accessed 2012) Available at: http://www.acr.org/~/media/ACR/Documents/AppCriteria/Interventional/RadiologicManagementLowerExtremityVenousInsufficiency.pdf

77. National Clinical Guideline Centre. : Varicose veins in the legs: the diagnosis and management of varicose veins (Clinical guideline CG168). (Accessed 2013) Available at: http://www.nice.org.uk/guidance/cg168/resources/cg168-varicose-veins-in-the-legs-full-guideline3

78. Biemans AA, K.: Comparing endovenous laser ablation, foam sclerotherapy, and conventional surgery for great saphenous varicose veins.. J Vasc Surg. 58(3), 727-34 (Sep 2013)

79. Nesbitt C, E.: Endovenous ablation (radiofrequency and laser) and foam sclerotherapy versus conventional surgery for great saphenous vein varices.. Cochrane Database Syst Rev. 10(CD005624) (2011)

80. Tellings SS, C.: Surgery and endovenous techniques for the treatment of small saphenous varicose veins: a review of the literature.. Phlebology 26(5), 179-84 (Aug 2011)

81. Carroll C, H.: Clinical effectiveness and cost-effectiveness of minimally invasive techniques to manage varicose veins: a systematic review and economic evaluation.. Health Technol Assess. 17(48), i-141 (Oct 2013)

82. Mozafar M, A.: Endovenous laser ablation of the great saphenous vein versus high ligation: long-term results.. Lasers Med Sci. 29(2), 765-71 (Mar 2014)

83. Lattimer CR, K.: Interim results on abolishing reflux alongside a randomized clinical trial on

laser ablation with phlebectomies versus foam sclerotherapy. Int Angiol. 32(4), 394-403 (Aug 2013)

84. Samuel N, C.: Randomized clinical trial of endovenous laser ablation versus conventional surgery for small saphenous varicose veins.. Ann Surg. 257(3), 419-26 (Mar 2013)

85. Pavlovic MD, S.-P.: Guidelines of the First International Consensus Conference on Endovenous Thermal Ablation for Varicose Vein Disease - ETAV Consensus Meeting 2012. Phlebology (Feb 2014)

86. Ontario Health Technology Advisory Committee: OHTAC recommendation: endovascular ablation for varicose veins. (Accessed 2013) Available at: http://www.hqontario.ca/Portals/0/Documents/eds/recommendation-endovascular-ablation-varicose-veins-1307-en.pdf

11.Appendix

Critical Appraisal of Individual Studies	Summary of Study Characteristics	Summary of critical appraisal	Summary of findings
Describe the method used to critically appraise the selected articles in the report. •The methods used to assess the risk of bias in individual studies must be described here. For instance, to assess the validity of RCTs, the adequacy of randomization, allocation concealment, blinding of patients, clinicians or health care providers, data collectors and outcome assessors, loss to follow-up, early stopping of trial, and description of intention-to-treat analysis must be discussed. •If an appraisal instrument was used, it should be described (for example, SIGN50, QUADAS, AGREE, etc.). The critical appraisal will vary by study design (for example, randomized controlled trial, economic evaluation, etc.). A scale or checklist that numerically summarizes the criteria into a single number can be misleading and must be avoided. •Defined criteria for each study design used to critically appraise the studies must be presented to ensure consistency. •A reference list and templates of critical	•A summary of study characteristics, such as PICOS elements, years of publications, and countries of origin of included studies, must be summarized in the text to provide an overview of the selected studies. •Study-level characteristics must be tabulated to ensure consistency of information across all studies and to facilitate study comparisons by the reader. As well, it will indicate clearly to the reader if any data are missing from or are unclear in the study. •Table 2 and 3 are example tables showing how study characteristics of included clinical and economic studies could be presented. Some of the column headings may vary depending on the research questions and decisions regarding PICOS elements. If the tables are longer than one page, include them in the appendix. •Separate tables should be created by study type. For instance, one table would present the study characteristics of included systematic reviews and meta-analyses and another table would present the characteristics of included	Describe the results of the critical appraisal of the included studies. •A narrative summary of the critical appraisal must be presented in the text. •An assessment of each study must be presented in a table using a standard approach with the criteria defined in the methods section. •Large tables (i.e., 1 page or more) should be presented in the appendix. •Tables should be presented according to the hierarchy of evidence. As such, health technology assessments, systematic reviews and meta-analyses are presented first. These are followed by RCTs and non-randomized studies. It may be possible to present the results of the critical appraisal of RCTs and non-randomized studies in one table. The results of the critical appraisal of economic evaluations and evidence-based guidelines will be presented in separate sections according to the research questions. •If a critical appraisal is conducted for multiple studies, the strengths and limitations of each study should be described in a tabular format and placed	The study findings of the selected studies must be presented and summarized for each study type included in the report. •The main study findings must be summarized in the text. The generalizability of these findings, key limitations of the study and its overall conclusions must be highlighted. •The complete study outcomes, including measures of association (for example, relative risk, mean differences) and, 95% confidence intervals (CIs) and conclusions must be presented in a table. •Table 5 presents an example of the main study findings and authors' conclusions. •Tables should be presented according to the hierarchy of evidence. As such, health technology assessments, systematic reviews and meta-analyses are presented first. These are followed by RCTs and non-randomized studies. It may be possible to present the main study findings and authors' conclusions of RCTs and non-randomized studies in one table. The main study findings and authors' conclusions of economic

Critical Appraisal of Individual Studies	Summary of Study Characteristics	Summary of critical appraisal	Summary of findings
appraisal techniques and tools: Critical Appraisal Reference List •A formal quality assessment of non-comparator studies or case reports should not be conducted since these study designs are considered to be inferior quality compared with systematic reviews, randomized control trials, or prospective, cohort studies.	RCTs. •Tables should be presented according to the hierarchy of evidence. As such, health technology assessments, systematic reviews and meta-analyses are presented first. These are followed by RCTs and non-randomized studies. It may be possible to present the study characteristics of RCTs and non-randomized studies in one table. The characteristics of economic evaluations and evidence-based guidelines will be presented in separate sections according to the research questions. •All abbreviations and acronyms used in a table MUST BE spelled out at the bottom of the table. •If there are one or two studies selected, a table may not be necessary and study characteristics may be described in a narrative summary.	in the appendix. •If a critical appraisal is conducted for few studies, the strengths and limitations of study may be presented in the main body of the report. •All abbreviations and acronyms used in a table MUST BE spelled out at the bottom of the table •Table 4 presents an example table of quality assessment of included studies.	evaluations and evidence-based guidelines will be presented in separate sections according to the research questions. •If reported, the 95% confidence interval (CI) should be presented to indicate the precision of the study results. The p-value should be presented to indicate whether the results were statistically significant only if the 95% CIs are not reported in the study. •If the author's conclusions or guidelines recommendations are copied directly from the study, quotation marks and page numbers must be included for each conclusion or recommendation. •Separate tables may be created by study type. For instance, one table would present the study findings of included systematic reviews and RCTs and another table would present the characteristics of included economic evaluations. Tables should be presented according to the hierarchy of evidence. As such, health technology assessments, systematic reviews and meta-analyses are presented first. These are followed by RCTs and non-randomized studies. It may be possible to present the study findings and

Critical Appraisal of Individual Studies	Summary of Study Characteristics	Summary of critical appraisal	Summary of findings
			conclusions of several study designs (for example, systematic reviews, RCTs and non-randomized studies) in one table. The findings and conclusions of economic evaluations and recommendations of evidence-based guidelines will be presented in separate sections according to the research questions. •If there are one or two studies selected, a table may not be necessary and study findings may be described in a narrative summary •For an economic evaluation, the study perspective, cost data and outcomes (for example, cost per QALY, cost per clinical outcome, cost per patient adverse event avoided) must be presented. If the economic evaluation was based on an economic model, the type of model (for example, decision analytic tree, Markov model, etc.) and the assumptions used must be described. •All abbreviations and acronyms used in a table MUST BE spelled out at the bottom of the table. •Large tables (i.e., 1 page or longer) should be placed in the appendix.