Handbook for Guideline Development

Version 4

Authors:
Renée L. Mulder, Princess Máxima Center for Pediatric Oncology, Utrecht, the Netherlands;
Morven C. Brown, Newcastle University, Newcastle upon Tyne, UK;
Roderick Skinner, Great North Children's Hospital, Royal Victoria Infirmary /Newcastle University, Newcastle upon Tyne, UK;
Melissa M. Hudson, St. Jude Children's Research Hospital, Memphis, US;
Leontien C.M. Kremer, Princess Máxima Center for Pediatric Oncology, Utrecht, the Netherlands and Cochrane Childhood Cancer Group, Amsterdam, the Netherlands;
On behalf of the IGHG core group and the PanCare Guidelines Group

February 2019

This handbook has been developed by a collaborative effort of the International Late Effects of Childhood Cancer Guideline Harmonization Group (IGHG) (www.ighg.org), the PanCare Guidelines Group) (www.pancare.eu) and the Cochrane Childhood Cancer Group (CCG) (ccg.cochrane.org).

This handbook can be used by the chairs and members of the IGHG guideline panels after they have been assigned to develop a special guideline.

This Handbook should be cited as Mulder RL, Brown MC, Skinner R, Hudson MM, Kremer LCM. Handbook for guideline development; collaboration between International Guideline Harmonization Group, PanCare Guideline Group and Cochrane Childhood Cancer Group. 2019

Other contributors: Armenian SH, Bárdi E, Bhatia S, Constine SL, Frey E, Haupt R, Kühni C, Landier W, Leclercq E, Levitt G, Michel G, Oeffinger KC, Wallace WH.

Funding
Dutch Cancer Society, the Netherlands (UVA 2011-4938)
7th Framework Programme of the EU, PanCareSurFup (257505)
KiKa grant Cochrane Childhood Cancer Group KiKa 123

Contents

1 Introduction

1.1 Aim of the handbook

The principle aim of this handbook is to serve as a reference tool and provide guidance to the members of the guideline panels involved in the guideline development work of the International Late Effects of Childhood Cancer Guideline Harmonization Group (IGHG)[1] (www.ighg.org) and the PanCare Guideline Group (www.pancare.eu). The systematic approach outlined in this manual aims to improve the methodological quality of the clinical practice guidelines for the follow-up of survivors of childhood, adolescent and young adult (CAYA) cancer and positively impact on the quality of care CAYA cancer survivors receive.

1.2 Clinical practice guidelines

Clinical practice guidelines (CPGs) are defined by the Institute of Medicine as *"statements that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options*.[2]

CPGs aim to provide appropriate recommendations for practice based on a transparent process and informed by evidence. CPGs are essential to ensuring that CAYA cancer survivors receive optimum health care.[2,3] However, it is essential to ensure optimum quality of guidelines if they are to improve both the process and outcome of care.

It is critical that CPGs are developed based on the methods of evidence-based medicine (EBM). EBM is "*the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients*".[4] EBM begins with the formulation of clinically relevant questions based on the Participants, Interventions, Control group & Outcome (PICO) system, followed by a synthesis of the evidence based on an extensive literature search (e.g. systematic review or evidence tables).The data is then used to develop evidence-based clinical policy (recommendations) before applying these policies or CPGs in practice **(Figure 1)**.

EBM is an integration of best research evidence, clinical expertise and patient concerns.

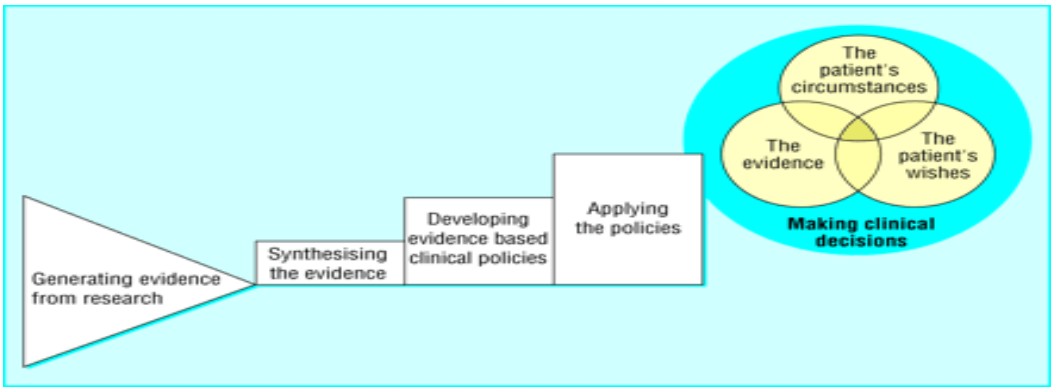


Fig. 1: The path from the generation of evidence to the application of evidence[5]

Members are encouraged to listen to the first web-training conference given by L. Kremer. This provides audio commentary and PowerPoint slides to give useful background information to evidence based clinical practice guidelines. The presentation is available at: https://connect.sunet.se/p5gqc2b67eg/

1.3 Guidelines for the long-term follow-up of CAYA cancer survivors

Several guidelines for the long-term follow-up of CAYA cancer survivors have been developed, published and widely disseminated, including those produced by the US Children's Oncology Group (COG), Dutch Childhood Oncology Group (DCOG), United Kingdom Children's Cancer Study Group (UKCCSG) and Scottish Intercollegiate Guidelines Network (SIGN). A recent survey performed by PCSF found that these guidelines were in widespread use across Europe and that some European nations have also developed additional local guidelines.

However, as existing guideline development groups worked independently, inconsistencies exist in the methodology used and also in the final recommendations of these guidelines. A major consequence of this is uncertainty among clinicians regarding which guidelines to implement.

Therefore, the goal of the IGHG and PanCare Guideline Group is to harmonise efforts and employ a systematic and rigorous methodology to produce clinical consensus in guidelines for long-term follow up of major late adverse effects in CAYA cancer survivors. We aim to promote healthy lifestyles, provide on-going monitoring of health status, facilitate early detection of late effects, and advise about timely intervention strategies to preserve health.

1.4 Structure of this handbook

In this handbook we aim to provide information that may be useful to members of the IGHG/PanCare Guideline Group guideline panels, and the working group leaders in particular as they prepare for and proceed through the guideline development work.

Specifically the handbook will:

1. Outline the key steps in the development of clinical practice guidelines.
2. Direct members to other important sources of information/documentation integral to the guideline development work.
3. Provide practical information regarding the organisation and management of the working groups.

As opposed to fully reproducing information which is documented elsewhere, this handbook provides an overview and directs members to other documents that explain the relevant issues in more detail. Links to these other documents are embedded in the text, and can be accessed by clicking on the document name. These documents include a published methodology paper, protocols from previous guideline topics that have been completed by IGHG/PCSF, and also practical examples from the published guidelines to more clearly illustrate the process.

2 Methodology utilised by IGHG/PCSF to develop evidence based CPG's

Developing a guideline encompasses three phases:

1. Preparation phase
2. Development phase
3. Finalisation phase

Members are encouraged to listen to the second web-training conference given by R Mulder. This provides useful background information to the development of clinical practice guidelines. The presentation is available at: https://connect.sunet.se/p2a8jwypnwg/

2.1 Preparation phase

The guideline panel

Convening an effective guideline panel is a crucial stage in producing a guideline. Each guideline panel will consist of a working group including:

- Chair(s): leaders in the field
- Coordinator(s): project managers administrating group activities
- Advisors: Leontien Kremer, Melissa Hudson, Renée Mulder, Rod Skinner, Sandy Constine (radiation expert), Hamish Wallace
- Working group leaders: leaders supervising literature reviews of focused clinical questions
- Working group members

Diversity is an essential feature of a guideline panel. Its exact composition should be tailored to the guideline topic and reflect the range of stakeholders involved. At a minimum the panel should comprise at least of content experts, non-expert clinicians, health care providers and methodologists. In addition, patients or their representatives may be eligible members.

Scope of the guideline

It is important to define the scope of the guideline:

- Definition of outcomes / health problem
- Age range of the population of interest:
    - Childhood, adolescent and young adult cancer survivors diagnosed with cancer up to age 30 years; depending on the health problem adaptation of the age range (e.g., 18, 21 or 25 years) may be appropriate.
- Survival time of the population of interest:
    - Childhood, adolescent and young adult cancer survivors 2-years after completion of treatment; depending on the health problem adaptation of the survival time (e.g., immediately following or 5 years post-treatment) may be appropriate.

For every guideline topic, the following key issues can be considered which are important for the final recommendations:

- Does early diagnosis result in better outcomes?
- Who needs surveillance?
- At what age or time from exposure should surveillance be initiated?
- At what frequency should surveillance be performed?
- When should surveillance be stopped?
- What surveillance modality should be used?
- What should be done if abnormalities are identified?

2.2 Development phase

In general, the guideline development process consists of five steps:

1. Evaluate concordances and discordances among recommendations in existing guidelines.
2. Formulate clinical questions.
3. Identify available evidence by systematic literature searches.
4. Summarize and grade evidence.
5. Formulate and grade recommendations.

Figure 2 outlines the main steps that IGHG & PanCare Guidelines Group will be undertaking in the development of guideline recommendations.

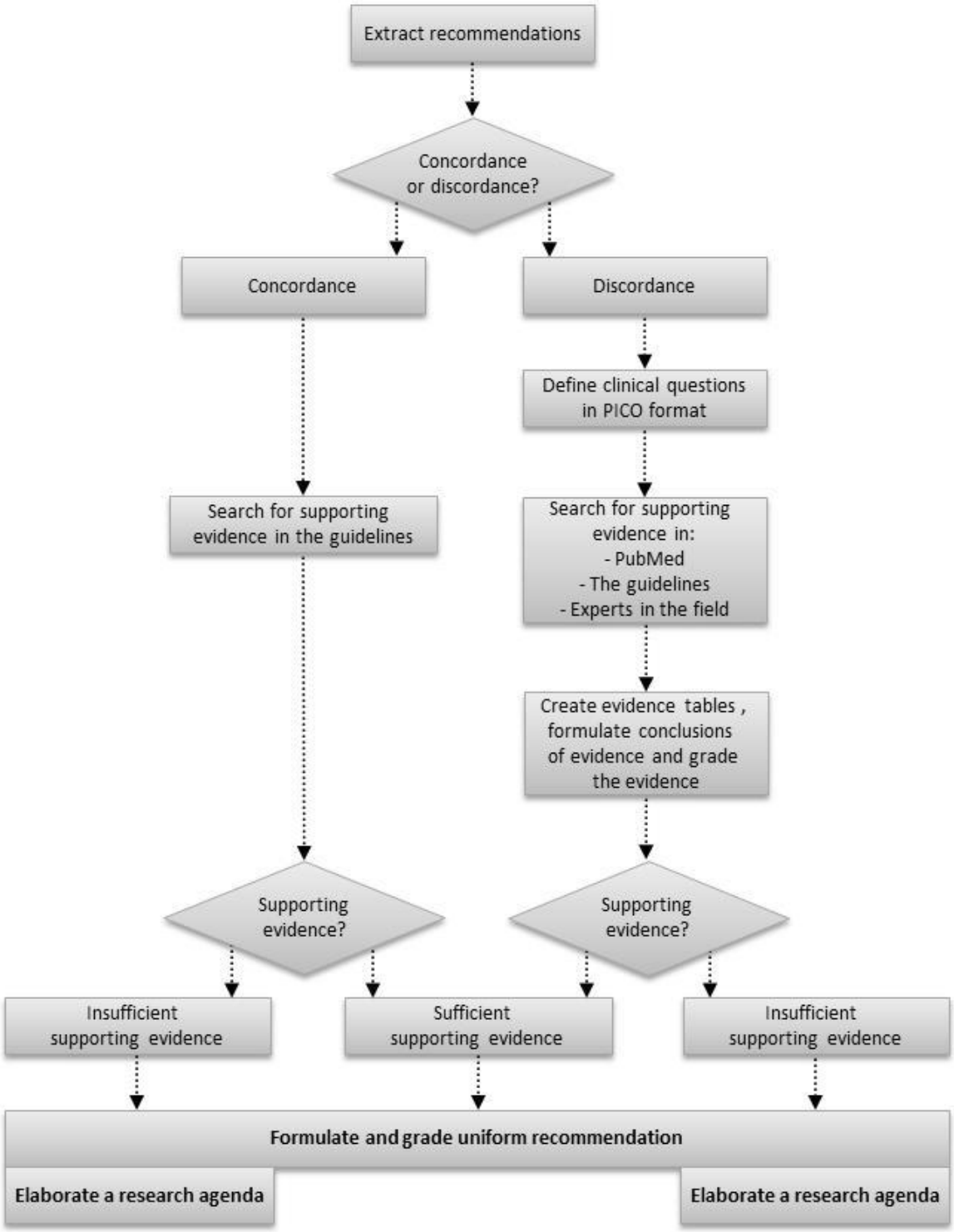


Fig. 2: Key stages in the development of recommendations

Step 1: Evaluate concordances and discordances of current recommendations

The first step is to extract the recommendations for the topic from the existing and more widely disseminated guideline groups (e.g., COG, DCOG, UKCCLG and SIGN guidelines). The level of discordance/concordance between these recommendations is then evaluated.

If recommendations are concordant, the quality of the supporting evidence will be reviewed to determine if it is sufficient or insufficient. Extensive evidence summaries will not be developed for concordant recommendations.

Discordant recommendations will form the basis for the formulation of clinical questions. These clinical questions will clearly state what the evidence aims to answer.

Below is an example of the evaluation of the concordance and discordance regarding the surveillance of breast cancer in survivors (Table 1).

Table 1: Concordance and discordance 'Who needs breast cancer surveillance?'

	COG	DCOG	UKCCLG	Concordant/ discordant
Who needs breast cancer surveillance?				
At risk **Chest radiation**	Yes	Yes	Yes	**Concordant**
± Alkylating agents	Not specified	Not specified	Yes	**Discordant**
High risk	Not specified	≥7-20 Gy chest radiation (excl. TBI) ≥14-40 Gy abdominal radiation	Not specified	**Discordant**
Highest risk	≥20 Gy chest radiation	≥20 Gy chest radiation ≥40 Gy abdominal radiation TBI	Not specified	**Discordant**

Step 2: Formulate clinical questions

Effective and efficient guideline development involves asking and answering key clinical questions. These questions should be clear, focused and closely define the boundaries of the topic. They will serve as a starting point for the systematic literature search that aims to identify all the available evidence. These questions also form the basis of the development of recommendations.

The **PICO** (Participants, Interventions, Control group & Outcome) framework is helpful to identify the main elements of the clinical question. It breaks the question down into four key elements:

- Who are the **P**articipants you want to study? (e.g., gender, age, disease)
- What is the **I**ntervention you want to examine? (etiologic/risk factor; e.g., type of treatment)
- What do you want to **C**ompare against your intervention of interest? (e.g., alternative interventions - this is not always necessary or relevant)
- What are the **O**utcomes you want to measure? (e.g., improved quality of life, late effects)

Examples of the formulation of a clinical question is shown in Table 2.

Table 2: Example clinical questions derived from the PICO structure

Does early diagnosis result in better outcomes?

P	I	C	O	Clinical question
Childhood, adolescent and young adult cancer survivors with a CNS neoplasm	Tumor size, asymptomatic or symptomatic stage	Not applicable	Mortality, recurrence, survival, adverse events, quality of life	Does the detection of a meningioma in a smaller size or asymptomatic stage contribute to a reduced mortality rate in CAYA cancer survivors?

Who needs surveillance?

P	I	C	O	Clinical question
Female childhood, adolescent and young adult cancer survivors	Low dose chest radiation	Childhood cancer survivors treated without chest radiation	Breast cancer risk	What is the risk of breast cancer in female CAYA cancer survivors treated with 1-9 Gy chest radiation compared to survivors treated without chest radiation?

At what age or time from exposure should surveillance be initiated?

P	I	C	O	Clinical question
Female childhood, adolescent and young adult cancer survivors	Chest radiation	N/A	Latency time breast cancer	What is the latency time (time of onset) to develop breast cancer in CAYA cancer survivors treated with chest radiation?

At what frequency should surveillance be performed?

P	I	C	O	Clinical question
Female childhood, adolescent and young adult cancer survivors	Chest radiation	N/A	Breast cancer risk over time	Does the breast cancer risk change over time (improve, deteriorate, plateau) in female CAYA cancer survivors treated with chest radiation? What is the timing of such change?

When should surveillance be stopped?

P	I	C	O	Clinical question
Female childhood, adolescent and young adult cancer survivors	Chest radiation	N/A	Breast cancer risk in CAYA cancer survivors aged >50 years	What is the risk of breast cancer in CAYA cancer survivors treated with chest radiation aged >50 years?

What surveillance modality should be used?

P	I	C	O	Clinical question
Female childhood, adolescent and young adult cancer survivors	MRI	Mammography	Diagnostic value to detect breast cancer	What is the diagnostic value (sensitivity, specificity, predictive value) of a MRI compared to a mammography to detect breast cancer in female CAYA cancer survivors?

What should be done if abnormalities are identified?

P	I	C	O	Clinical question
Childhood, adolescent and young adult cancer survivors	Physical activity training	No physical activity training	Pulmonary outcomes	What are the positive and adverse effects of physical activity on pulmonary outcomes in CAYA cancer survivors?

Step 3: Identify and select the evidence

It is important that the literature search is thorough, objective and rigorous. An inefficient or biased literature search can compromise the validity of the recommendations and the guidelines. The aim is to identify as many relevant studies as possible (within the limits of resources and time). It is also essential that the literature search is transparent, well documented and reproducible.

Where adequate published systematic reviews exist, additional literature searches may be limited to updating, covering the time period since the review was conducted.

Carrying out a literature search to identify and select relevant studies will involve:

1. Designing search strategies
2. Defining in- and exclusion criteria
3. Selecting studies for evidence summaries

1. Design search strategies

Where to search? Searches are carried out in bibliographic databases. There are several that can be searched but Medline and Embase are two of the key international health databases. Although there is significant overlap in these databases, differences do exist. The Cochrane Central Library of Controlled Trials is also a database for systematic reviews that can be searched.

In addition to searching bibliographical databases, papers should also be identified through references in the existing guidelines, as well as important reviews and key papers known to the group members.

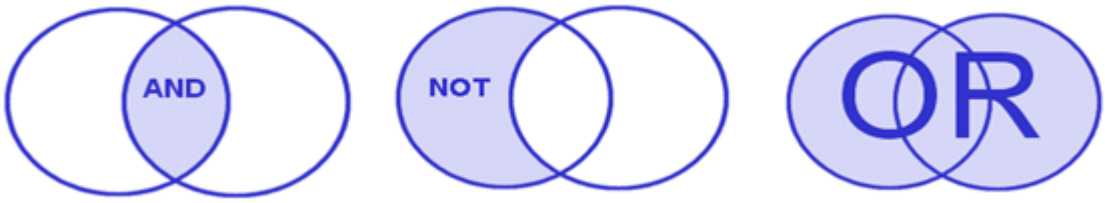
What to search? In order to search for and identify relevant studies, a search strategy must be developed. The search strategy is based on the main concepts in the clinical question identified through the PICO framework e.g. population, intervention, comparison and outcome. The clinical questions should be translated into key words and/or search terms. The Cochrane Childhood Cancer Group (ccg.cochrane.org) will develop the search strategies. However, members of the panel will be asked to suggest appropriate search terms and to check if the final search strategy is comprehensive.

Searches can be conducted in databases using either controlled vocabulary based on Medical Subject Headings (MeSH) or by using free-text/keywords. MeSH headings are useful as they index all articles that use different spellings/words to describe the same concept (e.g. cancer, lymphoma, leukemia, Ewing's sarcoma) under the same subject heading (e.g. Neoplasms). This precludes the need to search for a large list of synonyms. To identify keywords, however, look for the exact word you are searching for within the title and/or abstract of the articles within the database.

It is important to consider and include all of the related terms, variations in spellings and synonyms for each concept included in your search. A combination of subject headings and keywords is usually recommended to ensure that as many relevant records as possible are identified.

In **Appendix 1** standard search strategies as used by the IGHG and PanCare Guideline Group are shown. In addition, an example of a full search strategy taken from the male gonadal dysfunction guidelines protocol can be found in **Appendix 2**.

How to search? In the example in Appendix 2 many of the terms relating to the PICO framework are combined by 'OR'. This is a Boolean operator. Other Boolean operators are 'AND' and 'NOT'. Boolean operators make it possible to combine the results from two or more different searches using controlled vocabulary or keywords.



- **AND** – retrieves only those articles in which all of the terms appear
- **NOT** – used to exclude a term from your search
- **OR** – retrieves those articles in which either of the terms appear

For an explanation of search strategies and Boolean operators please see Lundh et al (2007). Development of a search strategy.[6]

2. Defining in- and exclusion criteria

It is important to define clear inclusion and exclusion criteria for the selection of studies, based on the **PICO**s. The following criteria should be considered:

- **Study population:**
  - Childhood, adolescent and young adult cancer survivors
  - At least 75% diagnosed with cancer prior to the defined age range (i.e.,18 / 21 / 25 / 30 years)
  - At least 50% survived the defined survival time (i.e., immediate end of treatment / ≥ 2 years post-treatment / ≥ 5 years post-treatment)
- **Outcomes:**
  - Outcome definition of specific late effect
  - Studies investigating one of the following outcomes depending on the clinical question:
    - For 'Who needs surveillance?': risks and risk factors.
    - For 'At what age or time from exposure should surveillance be initiated?: latency time / time of onset after exposure.
    - For 'At what frequency should surveillance be performed?': risks over time.
    - For 'What surveillance modality should be used?': diagnostic value / sensitivity, specificity, positive predictive value, negative predictive value of diagnostic tests.
    - For 'What should be done if abnormalities are identified?': effectiveness of interventions that may result in better outcomes.
- **Types of studies:**
  - Include all study designs except case reports and case series (systematic reviews provide the highest level of evidence followed by randomized controlled trials, observational studies)
    - For 'At what frequency surveillance should be performed?' longitudinal studies with more than one measurement per patient should be included.
    - For 'What surveillance modality should be used?' diagnostic studies should be included.
    - Regarding reviews: During screening of abstracts include all reviews (both systematic and narrative reviews). In cases of systematic reviews, include and use conclusions for generating evidence tables. In cases of narrative reviews, exclude, but screen reference lists in order to check for missing relevant papers.
  - Define minimum sample size, for example at least N=20 depending on the clinical problem and availability of evidence.
  - Prioritize, when available, studies that controlled for important confounding factors:
    - Cohort study: multivariable / multiple regression analysis;
    - Case-control study: matching or risk stratification.
  - Limit search to English language publications.
  - Define dates of search parameters, e.g., published from a specific date onwards (i.e. 1990).

3. Identify and select studies

Once the literature search of the electronic databases is complete, the following steps should be taken for selecting the studies:

- Two reviewers will assess if publications meet inclusion criteria based on the titles and abstracts of the studies.
- Every abstract will be assessed regarding the appropriateness of study inclusion (i.e., should be included, should be excluded, or inclusion uncertain).
- The results of reviewers' assessments will be compared and discrepancies discussed and resolved.
- The coordinator will obtain all "included" and "uncertain" abstracts in full text and send it to the two reviewers to determine if the inclusion criteria are met.
- Each full text paper will be reviewed and assessed regarding the appropriateness of inclusion of the study (i.e., should be included or excluded). Reason for study exclusion should be noted.
- Identify the clinical question for which the study should possibly be included.
- Discuss discrepancies with companion reviewer to reach consensus.

Besides to the PubMed search additional studies will be identified by:

- References in reviews
- References supporting the existing long-term follow-up guidelines
- Experts in the field

When evidence is lacking for childhood and young adult cancer survivors, we will carefully extrapolate evidence from other populations.

Step 4: Summarize and appraise quality of evidence

1. Evidence tables

The evidence found in the literature should be summarized in evidence tables **(see Appendix 3)**. The evidence tables provide information about study and patient characteristics, primary study outcomes, risk of bias, including selection bias, attrition bias, detection bias and confounding **(see Appendix 4)**, and additional remarks, such as other factors that may bias results.

For an explanation of the different types of bias, please see van Dalen et al (2007) Quality of studies included in a systematic review and associated risk of bias[7] and the Cochrane Bias Methods Group.

2. Summary of findings tables of the body of evidence

a. Description of studies

For each clinical question a summary of findings table of the body of evidence will be completed. A summary of findings table provides key information of every single study about the main patient characteristics, the magnitude of effects for the defined outcomes and determinants, and the quality of that study **(see Appendix 5)**.

b. Grading the quality of the body of evidence

The Grades of Recommendation, Assessment, Development and Evaluation Working Group (GRADE Working Group) has developed a system for grading the quality of a body of evidence,[8-11] The quality of a body of evidence is defined as the extent to which one can be confident that an identified effect or association is true. Assessing the quality of a body of evidence involves several considerations. There are five reasons for downgrading the quality of a body of evidence:

- Study limitations (risk of bias, **see Appendix 4**): the confidence in the body of evidence decreases when studies have major limitations that may bias the risk estimates. Every study addressing a particular outcome will differ, to some degree, in the risk of bias. The reviewers must make an overall judgement on whether the quality of evidence for an outcome warrants downgrading on the basis of study limitations.
- Inconsistency of the results (heterogeneity): the confidence in the body of evidence decreases when there is a degree of inconsistency of effect between or within studies (when studies yield widely differing estimates of effect). When heterogeneity exists and affects the interpretation of results, but reviewers fail to identify a plausible explanation, the quality of the evidence decreases.
- Indirectness of the study population, intervention and outcomes: the confidence in the body of evidence decreases when the study population and outcomes from the studies are not generalizable to the population and outcome of interest. Reviewers should make judgements transparent when they believe downgrading is justified, based on differences in anticipated effects in the group of primary interest.
- Imprecision of the effect estimates: the confidence in the body of evidence decreases when the effect estimates are imprecise. This is the case if studies include relatively few patients and few

events and thus have wide confidence intervals. Another criteria to consider is the clinical decision threshold. This is the threshold of the effect size that would change the decision whether or not to adopt a clinical action.

- Risk of publications bias: the confidence in the body of evidence decreases when investigators fail to report studies or outcomes on the basis of results, typically those studies that show no effect.

There are also circumstances in which the quality of the body of evidence can be upgraded:

1. Large magnitude of effect: the confidence in the body of evidence increases when methodologically well-performed observational studies yield large, consistent and precise estimates of the magnitude of effect.
2. Dose response gradient: the confidence in the body of evidence increases when there is evidence for a dose response across or within studies, or when inconsistency across studies is explained by a dose response.
3. Plausible confounding: the confidence in the body of evidence increases when adjustment for confounding factors would have increased the effect size.

In **Appendix 6** the criteria for grading the quality of the body of evidence is described. The rational for grading the quality of the body of evidence should be described in the summary of findings table **(see Appendix 5)**.

The evidence is graded according to four levels:

- ⨁⨁⨁⨁ High: further research is unlikely to change the confidence in the estimate of effect.
- ⨁⨁⨁⊖ Moderate: further research is likely to have an important impact on the confidence in the estimate of effect and may change the estimate.
- ⨁⨁⊖⊖ Low: further research is very likely to have an important impact on the confidence in the estimate of effect and is likely to change the estimate.
- ⨁⊖⊖⊖ Very low: any estimate of effect is very uncertain.

c. Formulation of the conclusions of evidence

Based on the quality of the body of evidence an overall conclusion will be formulated for that specific clinical question **(see Appendix 5 and 7)**.

Grading the evidence gives an impression of the quality of the included studies. It is not related to the importance of the recommendation but to the strength of the supporting evidence.

For an explanation of the different types of factors that may decrease or increase the quality of a body of evidence, please see the in-depth publications on the GRADE website.

Step 5: Formulate recommendations

Once the selection and summary of the evidence is complete, the available evidence must be combined and translated into recommendations. For this purpose we will use the GRADE Evidence to Decision (EtD) framework **(Appendix 8)**.[12] The EtD framework ensures that all important criteria for making a decision is considered and informs the guideline panel about the relative pros and cons of the interventions or options being considered. It makes the decision making process structured and transparent.

The following criteria should be considered when formulating clinical recommendations:

1. Problem: Is the problem a priority?
2. Desirable effects: Are the desirable anticipated effects large?
3. Undesirable effects: Are the undesirable anticipated effects small?
4. Certainty of the evidence: What is the overall certainty of the evidence of effects?
5. Values: Is there important uncertainty about or variability in how much people value the main outcomes?
6. Balance of effects: Are the desirable effects large relative to the undesirable effects?
7. Resources required: Are the resource required small?
8. Certainty of evidence of required resources: What is the overall certainty of the evidence of the resources require?
9. Cost effectiveness: Is the incremental cost small relative to the net benefits?
10. Equity: What would be the impact on health equity?
11. Acceptability: Is the intervention/option acceptable to key stakeholders?
12. Feasibility: Is the intervention feasible to implement?

In addition, it is important to consider the need to maintain flexibility of application across health care systems.

Based on all the consideration a balance of consequences will be made for all different interventions/options:

- Undesirable consequences *clearly outweigh* desirable consequences in most settings
- Undesirable consequences *probably outweigh* desirable consequences in most settings
- The balance between desirable and undesirable consequences *is closely balanced or uncertain*
- Desirable consequences *probably outweigh* undesirable consequences in most settings
- Desirable consequences *clearly outweigh* undesirable consequences in most settings

For an explanation of the Evidence to Decision framework, please see the tutorials on GRADEpro website.

How important each of the considerations are for a recommendation can vary. To make a recommendation, the guideline panel must consider the implication and importance of each of the above judgments. In many cases, this will be straightforward and not require detailed consideration. However, when there is uncertainty or disagreement, it can help to explicitly consider this for each criterion. Based on the overall assessment across criteria the guideline panel must reach a conclusion about the direction and the strength of the recommendations.

Recommendations will be classified into three categories: strong recommendation to do (green); moderate recommendation to do (yellow); and recommendation not to do (red) **(see Appendix 9)**.[13] The guideline panel should provide a justification for the recommendations, based on the criteria used in their assessment.

The recommendations should be a stand-alone text written in a complete sentence. The wording should be unambiguous, clearly defined, easy to translate into clinical practice, and agreed by the complete guideline panel.

The recommendations can include the following items:

- Who needs surveillance?
- At what age or time from exposure should surveillance be initiated?
- At what frequency should surveillance be performed?
- When should surveillance be stopped?
- What surveillance modality should be used?
- What should be done if abnormalities are identified?

A first draft of the recommendations will be prepared by a smaller group (i.e. chairs, advisors, and working group leaders). Next, the recommendations will be discussed and further formulated by the total guideline panel. Additional experts and patients/survivors in the field should be invited to participate in this final discussion.

Below is an example of the recommendations from the breast cancer surveillance guideline **(Table 3)**.

Table 3: Example recommendations

Who needs breast cancer surveillance?
Breast cancer surveillance is recommended for female childhood, adolescent and young adult cancer survivors treated with ≥10 Gy chest radiation (level A evidence, strong recommendation).
Breast cancer surveillance is reasonable for female childhood, adolescent and young adult cancer survivors treated with high abdominal field radiation. The surveillance decision should be an individual one, taking into account additional risk factors1 and personal values regarding the harms and benefits of surveillance (see survivor information form) (level B evidence, moderate recommendation).

2.3 Finalisation phase

1. Writing the guideline

All guideline topics will be summarized in a manuscript appropriate for publication in a peer-reviewed journal. The guideline should include the following items:

- Background
- Methods: clinical questions, search strategy, selection of literature
- Results: description of evidence, overall conclusions, quality of the evidence
- Considerations: translation evidence into recommendations, according to the GRADE Evidence to Decision framework
- Recommendations
- Discussion including research agenda
- Reference list

2. External review

After the recommendations have been formulated, there will be a commentary phase where external experts review the guideline for content and implementability. Feedback is sought preferably among the scientific, professional and patient organisations involved. Feedback can also be invited from methodological experts who review the guideline for methodological validity.

3. Updating the guideline

Guidelines should be kept up to date. All IGHG/PCSF guidelines will carry a statement indicating that they will be considered for revision 5 years after publication. Searches for new evidence should be performed and updating of the recommendations might be considered.

3 Roles, publication policy and author contributions

3.1 The IGHG core leadership group

- Melissa Hudson; co-chair, advisor, COG representative
- Leontien Kremer; co-chair, advisor, PanCare and DCOG representative
- Renée Mulder; coordinator, advisor, DCOG representative
- Rod Skinner; advisor, PanCare and UKCCLG representative
- Sandy Constine; advisor and radiation expert, COG representative
- Hamish Wallace, advisor; SIGN representative
- Saro Armenian; COG representative
- Smita Bhatia; COG representative
- Wendy Landier; COG representative
- Gill Levitt; UKCCLG representative
- Kevin Oeffinger; COG representative
- Lars Hjorth; PanCare representative

Roles

- The core leadership group members will set up the methodology.
- The core leadership group members will develop future plans.
- The core leadership group members will organize meetings for the guideline panels.
- The core leadership group members will guide the work of guideline panels.

3.2 The guideline panel

Roles

- The guideline panel consists of: chairs (representing different continents), a coordinator, advisors, working group leaders and working group members.
- The chairs and advisors will appoint the guideline panel.
- The IGHG core leadership group will approve the composition of the guideline panel.
- The guideline group coordinator will facilitate the group's work and telephone meetings.
- The chairs, coordinator and working group leaders will formulate clinical questions with help of the advisors.
- The final clinical questions will be reviewed by the advisors before discussion with the guideline group members.
- The Cochrane Childhood Cancer Group will develop the search strategy together with the chairs, coordinator, working group leaders and advisors.
- The chairs, coordinator and working group leaders will formulate inclusion criteria for evidence selection with help of the advisors.
- The final inclusion criteria will be reviewed/approved by the advisors before discussion with the guideline panel members.
- The chairs, coordinator, working group leaders and members will select the publications meeting the criteria established for evidence selection.

- The chairs, coordinator, working group leaders and members will produce evidence summaries with help of the advisors.
- The final evidence summaries will be reviewed by the advisors.
- The chairs, coordinator and working group leaders will formulate conclusions of evidence with help of the advisors.
- The final conclusions of evidence will be reviewed by the advisors before discussion with the guideline panel members.
- The chairs, coordinator, working group leaders and advisors will prepare a first draft of the recommendations that will be subsequently discussed with the guideline panel members.

3.3 Manuscript writing process

- A primary manuscript of the whole guideline will be drafted that will include a description of the evidence and recommendations.
- The chairs and coordinator of the guideline group will write the first draft of the manuscript.
- The advisors and working group leaders will review/revise the first draft.
- The revised manuscript will be distributed to the guideline panel members.
- The final manuscript summarizing recommendations will be approved by the IGHG core leadership group.
- The development of additional manuscripts describing special aspects of the guideline topic should be discussed with and approved by the advisors.

3.4 Manuscript authorship

- The chairs, coordinator, working group leaders, members and advisors will be authors of the manuscript if they are substantially involved in the guideline development process (e.g., participate in the study selection, develop evidence summaries and conclusions, formulate recommendations, and write or provide critical input on the manuscript) (see the authorship guidelines of the International Committee of Medical Editors: http://www.icmje.org/)
- Authorship criteria should be communicated to members at the beginning of the guideline development process.
- The decision regarding authorship will be made by the chairs of the guideline group in consultation with the advisors.
- The person who drafts the manuscript will be first author of the guideline manuscript; this will be the coordinator or, if the coordinator is not able to write a first draft, one of the chairs. In the event that the coordinator is the first author, the chairs will be 2nd and last author or shared last authors.
- Other working group members and IGHG core leadership group members who have not been substantially involved in the guideline or manuscript development will be acknowledged in the manuscript if their contribution is limited to review and approval of the final manuscript draft.
- If financially supported by PanCare, the collaboration between IGHG and PanCare or other funding sources should be acknowledged in the titles of the published manuscripts.

3.5 Reviewers

- At least two expert reviewers and two patient representatives will be asked to review the final manuscript of the guideline.
- These reviewers will be acknowledged in the manuscript.

3.6 Presentations

- The advisors should be informed if the methods and results of IGHG endeavors are to be submitted for presentation at national and international conferences.
- All presentations in which the results of the guideline harmonization endeavor are highlighted should acknowledge the names of the IGHG core group and the specific guideline group.
- All publications and presentations should acknowledge funding sources.

4 Communication and monitoring of progress

4.1 Expected timeline for guideline development

Please note that the timelines of guideline development work are dependent on many factors, therefore, timelines will differ between topic groups. An example timeline is shown in **Appendix 10.** This provides an illustration of the stages of the work scope that are expected to occur in development work and approximate timelines. However, these timelines will be modified according to the work of each guideline panel and most likely revised as the groups progress through their work.

4.2 Teleconferences

To arrange telephone-conference times, Doodle (www.doodle.com) is useful. Invitations are sent and group members select which days/times they can/cannot attend. Changing on the time-zone support will automatically adapt the time to each participants own time-zone so there is no confusion.

The coordinator can open an account on www.freeconferencecall.com to organize free conference calls with an unlimited number of participants.

4.3 Shared-calendar

It may be useful to set up a shared-calendar system for your working group, or to send calendar invitations for meetings for teleconferences. These can be set-up in Microsoft Outlook. They allow tasks and events to be entered and for reminder alerts to be set. This can be useful to provide a visual representation of the programme of work and for working groups to keep track of tasks and progress.

5 Overview of other key sources of information and support

Briefly, others main sources of information available are:

- The methodology paper describing the rationale behind the harmonisation effort and the planned methodology published by Kremer et al (2013)[1]
- Recommendations for breast cancer surveillance: a report from the IGHG published by Mulder et al (2013)[14]
- Recommendations for cardiomyopathy: a report from the IGHG published by Armenian et al (2015)[15]
- Recommendations for premature ovarian insufficiency: a report from the IGHG and PCSF published by van Dorp et al (2016)[16]
- Recommendations for male gonadotoxicity: a report from the IGHG and PCSF published by Skinner et al (2017)[17]
- Recommendations for thyroid cancer: a report from the IGHG and PCSF published by Clement et al (2018)[18]
- Recommendations for ototoxicity: a report from the IGHG and PCSF published by Clemens et al (2019)[19]

- Recordings of two one-hour training web-conferences on evidence based guidelines given by Leontien Kremer and Renée Mulder to PCSF WP6 members in June 2012:
  - https://connect.sunet.se/p5gqc2b67eg/
  - https://connect.sunet.se/p2a8jwypnwg/

- Protocol from previous guidelines on request

Useful websites are:

- International Guideline Harmonization Group: http://www.ighg.org/
- Cochrane Childhood Cancer Group: http://ccg.cochrane.org/ebch-cochrane-journal/
- Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group: http://www.gradeworkinggroup.org/ and https://gradepro.org/
- Appraisal of Guidelines Research & Evaluation (AGREE): http://www.agreetrust.org/

References

1. Kremer LCM, Mulder RL, Oeffinger, KC, et al. A worldwide collaboration to harmonize guidelines for the long-term follow-up of childhood and young adult cancer survivors: a report from the International Late Effects of Childhood Cancer Guidelines Harmonization Group. Pediatr Blood Cancer 2013;60:543-549.
2. Institute of Medicine. Clinical practice guidelines we can trust. Washington: The National Academies Press, 2011.
3. Burgers J, Grol R, Eccles M. Clinical guidelines as a tool for implementing change in patient care. In Grol R, Wensing M & Eccles M (eds). Improving patient care: the implementation of change in clinical practice. 2005. Edinburgh: Elsevier.
4. Sacket Dl, Rosenberg WMC, Muir Gray JA, et al. Evidence based medicine: what it is and what it isn't. BMJ 1996;312:71.
5. Haynes B. Barriers and bridges to evidence based clinical practice. BMJ 1998;317:273.
6. Lundh A, Kremer LCM, Leclercq E. Tips and tricks for understanding and using SR results. No. 6: development of a search strategy. Evidence-Based Child Health: A Cochrane Review Journal 2007; 2:937-939.
7. van Dalen EC, Kremer LCM, Moyer VA. Tips and tricks for understanding and using SR results. No. 8: quality of studies included in a systematic review and associated risk of bias – 'garbage in, garbage out'. Evidence-Based Child Health: A Cochrane Review Journal 2007; 2:1321-1324.
8. GRADE Working Group. Grading quality of evidence and strength of recommendations. BMJ 2004; 328: 1490-1494.
9. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008; 336: 924-926.
10. Guyatt GH, Oxman AD, Schunemann HJ, et al. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. Journal of Clinical Epidemiology 2011; 64: 380-382.
11. Schünemann HJ, Jaeschke R, Cook DJ, et al. An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. American Journal of Respiratory and Critical Care Medicine 2006; 174: 605-614.
12. Alonso-Coello P, Schünemann HJ, Moberg J, et al. GRADE Evidence to Decision frameworks: a systematic and transparent approach to making well-informed healthcare choices. 1: Introduction. BMJ 2016;353:i2016.
13. Gibbons RJ, Smith S, Antman E. American College of Cardiology/American Heart Association clinical practice guidelines: Part I: where do they come from? Circulation. 2003; 107(23): 2979-86.
14. Mulder RL, Kremer LCM, Hudson MM, et al. Recommendations for breast cancer surveillance for female childhood and young adult cancer survivors given chest radiation: a report from the International Late Effects of Childhood Cancer Guideline Harmonization Group. The Lancet Oncology 2013;14:e621-629.

15. Armenian SH, Hudson MM, Mulder RL, et al. Recommendations for cardiomyopathy surveillance for survivors of childhood cancer: a report from the International Late Effects of Childhood Cancer Guideline Harmonization Group. The Lancet Oncology 2015;16:e123-136.
16. van Dorp W, Mulder RL, Kremer LC, et al Recommendations for Premature Ovarian Insufficiency Surveillance for Female Survivors of Childhood, Adolescent, and Young Adult Cancer: A Report From the International Late Effects of Childhood Cancer Guideline Harmonization Group in Collaboration With the PanCareSurFup Consortium. Journal of Clinical Oncology 2016;34:3440-3450.
17. Skinner R, Mulder RL, Kremer LCM, et al. Recommendations for gonadotoxicity surveillance for male childhood, adolescent and young adult cancer survivors: A report from the International Late Effects of Childhood Cancer Guideline Harmonization Group in collaboration with the PanCareSurFup Consortium. Lancet Oncology 2017;18:e75-e90.
18. Clement SC, Kremer LCM, Verburg FA, et al. Balancing the benefits and harms of thyroid cancer surveillance in survivors of Childhood, adolescent and young adult cancer: Recommendations from the international Late Effects of Childhood Cancer Guideline Harmonization Group in collaboration with the PanCareSurFup Consortium. Cancer Treatment Reviews 2018;63:28-39.
19. Clemens E, van den Heuvel-Eibrink MM, Mulder RL, et al. Recommendations for ototoxicity surveillance for childhood, adolescent, and young adult cancer survivors: A report from the International Late Effects of Childhood Cancer Guideline Harmonization Group in collaboration with the PanCare Consortium. Lancet Oncology 2019;20:e29-e41.

Appendix 1

Standard search strategies

Cancer

Cancer OR cancers OR cancer* OR oncology OR oncolog* OR neoplasm OR neoplasms OR neoplasm* OR carcinoma OR carcinom* OR tumor OR tumour OR tumor* OR tumour* OR tumors OR tumours OR malignan* OR malignant OR hematooncological OR hemato oncological OR hemato-oncological OR hematologic neoplasms OR hematolo*

Childhood cancer

((leukemia OR leukemi* OR leukaemi* OR (childhood ALL) OR AML OR lymphoma OR lymphom* OR hodgkin OR hodgkin* OR T-cell OR B-cell OR non-hodgkin OR sarcoma OR sarcom* OR sarcoma, Ewing's OR Ewing* OR osteosarcoma OR osteosarcom* OR wilms tumor OR wilms* OR nephroblastom* OR neuroblastoma OR neuroblastom* OR rhabdomyosarcoma OR rhabdomyosarcom* OR teratoma OR teratom* OR hepatoma OR hepatom* OR hepatoblastoma OR hepatoblastom* OR PNET OR medulloblastoma OR medulloblastom* OR PNET* OR neuroectodermal tumors, primitive OR retinoblastoma OR retinoblastom* OR meningioma OR meningiom* OR glioma OR gliom*) OR (pediatric oncology OR paediatric oncology) OR (childhood cancer OR childhood tumor OR childhood tumors)) OR (brain tumor* OR brain tumour* OR brain neoplasms OR central nervous system neoplasm OR central nervous system neoplasms OR central nervous system tumor* OR central nervous system tumour* OR brain cancer* OR brain neoplasm* OR intracranial neoplasm*) OR (leukemia, lymphocytic, acute[mh]) OR (leukemia, lymphocytic, acute*)

Children

Infan* OR toddler* OR minors OR minors* OR boy OR boys OR boyfriend OR boyhood OR girl* OR kid OR kids OR child OR child* OR children* OR schoolchild* OR schoolchild OR school child[tiab] OR school child*[tiab] OR adolescen* OR juvenil* OR youth* OR teen* OR under*age* OR pubescen* OR pediatrics[mh] OR pediatric* OR paediatric* OR peadiatric* OR school[tiab] OR school*[tiab]

Children, adolescents and young adults

Infan* OR toddler* OR minors OR minors* OR boy OR boys OR boyfriend OR boyhood OR girl* OR kid OR kids OR child OR child* OR children* OR schoolchild* OR schoolchild OR school child[tiab] OR school child*[tiab] OR adolescen* OR juvenil* OR youth* OR teen* OR under*age* OR pubescen* OR pediatrics[mh] OR pediatric* OR paediatric* OR peadiatric* OR school[tiab] OR school*[tiab] OR young adult[mh] OR adult[mh] OR young adult

Survivors

Survivor OR survivors OR survivor* OR long term survivor OR long term survivors OR long term survivor* OR survivo* OR surviving OR long term survival[tiab] OR survival[mh]

Late effects

"late effect" OR "late effects" OR "late effect*" OR "late side effect" OR "late side effects" OR "late side effect*" OR "late adverse effect" OR "late adverse effects" OR "late adverse effect*" OR long term effect[tiab] OR long term effect* OR long term adverse effects[mh] OR aftercare OR follow up studie* OR follow up study

Radiotherapy – general

Radiotherapy OR radiation OR radiation therapy OR irradiation OR irradiat* OR radiation injuries OR injuries, radiation OR injury, radiation OR radiation injury OR radiation syndrome OR radiation syndromes OR syndrome radiation OR radiation sickness OR radiation sicknesses OR sickness radiation OR radiation* OR irradiation OR radiations

Radiotherapy – extensive

Radiotherapy OR radiation OR radiation therapy OR irradiation OR irradiat* OR radiation injuries OR injuries, radiation OR injury, radiation OR radiation injury OR radiation syndrome OR radiation syndromes OR syndrome radiation OR radiation sickness OR radiation sicknesses OR sickness radiation OR radiation* OR irradiation OR radiations

OR stereotactic RT OR stereotactic radiotherapy[tiab] OR gamma knife OR intensity modulated radiotherapy OR IMRT OR radiotherapy, intensity-modulated[mh] OR (three dimenstional OR 3D OR 3d CRT) OR image guided radiotherapy OR IGRT OR radiotherapy, image-guided[mh] OR photon radiotherapy OR XRT OR "photons/therapeutic use"[Mesh] OR proton radiotherapy OR PRT OR proton therapy OR proton radiation OR proton beam OR carbon ion radiotherapy

Dose-response relationship – combine with radiotherapy search

radiometry OR radiation dosage OR radiation dose OR radiation doses OR radiation dosis OR radiation dosage* OR radiation dosimetry OR radiation dosimetr* OR dose-response relationship, radiation OR radiometr* OR radiotherapy dosage OR radiotherapy[sh] OR radiotherapy/adverse effects OR irradiation dose OR radiotherapy dose OR dose calculation OR near beam dose OR in beam dose OR outside beam dose OR out of beam dose OR radiation/epidemiology OR Radiation monitoring OR Organs at risk OR radiation effects[sh] OR radiation injury OR radiation injuries OR radiation OR Radiotherapy/complications[Mesh]

Radiotherapy fields – combine with radiotherapy search

<u>Cranial, head and neck</u>

Cranial OR craniospinal OR (cranial OR craniospinal OR head[tiab] OR neck[tiab] OR skull

<u>Hypothalamic-pituitary</u>

(Hypothalamus OR Hypothalamus, Middle OR Hypothalamus, Anterior, OR Hypothalamus Posterior OR Pituitary Gland, Posterior OR Skull OR Orbit OR Orbits OR Eye OR Ear OR Nasopharynx)

<u>TBI</u>

TBI OR Total body OR whole body OR total body* OR body whole*

Chemotherapy

Antineoplastic Protocols OR Antineoplastic Combined Chemotherapy Protocols OR Chemoradiotherapy OR Chemoradiotherapy, Adjuvant OR Chemotherapy, Adjuvant OR Consolidation Chemotherapy OR Induction chemotherapy OR Maintenance chemotherapy OR Chemotherapy, Cancer, Regional Perfusion OR Antineoplastic agents OR chemotherap*

Alkylating agents

Antineoplastic agents, alkylating* OR antineoplastic alkylating agents OR alkylating agents, antineoplastic OR antineoplastic drugs, alkylating OR antineoplastics, alkylating OR alkylating antineoplastic drugs OR alkylating drugs, antineoplastic OR antineoplastic alkylating drugs OR drugs, antineoplastic alkylating OR alkylating antineoplastic agents OR alkylating antineoplastics OR Alkylating Agents OR alkylating agent*
OR busulphan OR busulfan* OR myleran* OR myelosan* OR Carmustine OR BCNU OR Chlorambucil OR ifosfamide OR iphosphamide OR iso endoxan OR isophosphamide OR isofosfamide OR ifosfa* OR iphospha* OR

isofosfa* OR cyclophosphamide OR cyclophosphane OR cytophosphan OR endox* OR cyclophospha* OR Lomustine OR CCNU OR lomustine* OR Mechlorethamine OR mechlorethamine*OR Chlormethine OR Mustine OR Chlorethazine OR Procarbazine OR procarbazin* OR Melphalan OR melphalan* OR Thiotepa OR Thio Tepa OR Thiophosphamide OR thiothepa* OR temozolomide OR dacarbazine OR decarbazine OR Fludarabine monophosphate*

Platinum agents
Cisplatin OR Platinum Diamminodichloride OR cis-Platinum OR cis Platinum OR Dichlorodiammineplatinum OR cis-Diamminedichloroplatinum OR cis Diamminedichloroplatinum OR Platinol OR Platidiam OR Platino OR Biocisplatinum OR CDDP OR CACP OR cisplatin* OR abiplatin OR neoplatin OR cis-DDP OR Carboplatin OR CBDCA OR Carbosin OR Carbotec OR Ercar OR Neocarbo OR Paraplatin OR Carboplat OR Paraplatine OR Platinwas OR Ribocarbo OR Blastocarb OR Nealorin OR carboplatin* OR Oxaliplatin OR oxaliplatin* OR oxaliplatine OR Eloxatine OR Eloxatin OR eloxatin* OR dacotin OR dacplat OR OR l-ohp OR oxalatoplatinum OR Platinum OR Platinum Compounds OR platinum* OR organoplatinum compounds [mh]

Cytarabine
cytosine* OR citosin* OR cytarabin* OR citarabin* OR arabino* OR arabitin* OR aracytine* OR aracytidin* OR cytin* OR cytidine* OR ara-c OR arac OR arafcyt OR cytosar* OR cytozar* OR ara-C OR beta-Ara C

Anthracyclines
anthracyclines OR anthracyclin* OR idarubicin OR idarubic* OR epirubicin OR epirubic* OR adriamycin OR doxorubicin OR doxorubic* OR adriamyc* OR daunorubicin OR daunorubic* OR daunoxome OR doxil OR caelyx OR myocet

Mitoxantrone
mitoxantrone OR mitoxantr*

MIBG
131I-Meta-iodobenzylguanidine OR 131I-MIBG OR 131I-metaiodobenzylguanidine OR Iodine-131 Metaiodobenzylguanidine OR Iobenguane (131I) OR (3-Iodo-(131I)benzyl)guanidine OR Iodine Radioisotopes/therapeutic use OR 3-Iodobenzylguanidine/therapeutic use) OR (iodine-131-metaiodobenzylguanidine OR 131I-MIBG therapy OR I-metaiodobenzylguanidine OR I-131-MIBG OR I-131-Metaiodobenzylguanidine OR (131) I-MIBG OR 3-Iodobenzylguanidine[mh] OR (131) I-metaiodobenzylguanidine OR (MIBG AND (treatment OR therapy))

Cost-benefit
Cost benefit analysis[mh] OR cost benefit OR cost benefit* OR costs* benefit OR cost effectiveness OR health care costs OR cost and cost analysis OR cost saving OR cost savings

Appendix 2

Example search strategy for male gonadal dysfunction

Search 1: Patient	(((leukemia OR leukemi* OR leukaemi* OR (childhood ALL) OR AML OR lymphoma OR lymphom* OR hodgkin OR hodgkin* OR T-cell OR B-cell OR non-hodgkin OR sarcoma OR sarcom* OR sarcoma, Ewing's OR Ewing* OR osteosarcoma OR osteosarcom* OR wilms tumor OR wilms* OR nephroblastom* OR neuroblastoma OR neuroblastom* OR rhabdomyosarcoma OR rhabdomyosarcom* OR teratoma OR teratom* OR hepatoma OR hepatom* OR hepatoblastoma OR hepatoblastom* OR PNET OR medulloblastoma OR medulloblastom* OR PNET* OR neuroectodermal tumors, primitive OR retinoblastoma OR retinoblastom* OR meningioma OR meningiom* OR glioma OR gliom*) OR (pediatric oncology OR paediatric oncology) OR (childhood cancer OR childhood tumor OR childhood tumors)) OR (brain tumor* OR brain tumour* OR brain neoplasms OR central nervous system neoplasm OR central nervous system neoplasms OR central nervous system tumor* OR central nervous system tumour* OR brain cancer* OR brain neoplasm* OR intracranial neoplasm*) OR testis neoplasm OR neoplasm, testicular OR testicular neoplasm OR testicular neoplasms OR testis cancer OR testicular cancer OR testis tumor OR testicular cancer OR cancer of testis OR testis tumour OR testis neoplasm* OR testis tumour* OR testis tumor* OR (leukemia, lymphocytic, acute[mh]))
Search 2: Patient	male[tiab] OR males OR boy OR boys OR boyfriend OR boyhood
Search 3: Intervention	Radiotherapy OR radiation OR radiation therapy OR irradiation OR irradiat* OR radiation injuries OR injuries, radiation OR injury, radiation OR radiation injury OR radiation syndrome OR radiation syndromes OR syndrome radiation OR radiation sickness OR radiation sicknesses OR sickness radiation OR radiation* OR irradiation OR radiations
Search 4: Intervention	Testicles OR testicle OR testes OR testis OR testis* OR testicle* OR testes* OR pelvic region OR region, pelvic OR pelvis region OR region pelvis OR pelvis* OR pelvic*
Search 5: Intervention	Brains OR brain OR encephalon OR encephalons OR brain* OR encephalon*
Search 6: Intervention	total body OR whole body OR total body* OR body whole*
Search 7: Outcome	spermatogenesis OR gonadal disorder OR spermiogenesis OR spermatocytogenesis OR spermatogenic failure OR azoospermia OR oligospermia OR asthenozoospermia OR teratozoospermia OR oligoasthenoteratozoospermia OR dysspermia OR normozoospermic OR semen OR semen analysis[text] OR semen quality[text] OR sperm OR sperm count OR sperm motility OR spermatozoa OR progeny OR offspring OR posterity OR fertility OR infertility OR subfertility OR reproduction OR fertilization OR conception OR paternity OR fatherhood OR parenthood OR pregnancy outcome OR fertile OR infertile OR subfertile OR sperm maturation OR aspermia OR spermatozoon abnormality
Search 8: Outcome	androgen hormone insufficiency OR leydig cell OR cells, leydig failure OR testicular interstitium cell failure OR testicular failure OR gonadal failure OR hypogonadism OR low testosterone OR testosterone deficiency OR androgen deficiency OR low testosterone* OR hypogonadism* OR leydig cell*
Search 1 AND 2 AND 3 AND (4 OR 5 OR 6) AND (7 OR 8) **Filters: published in the last 20 years; Humans**	**= 488 hits**

Appendix 3

Example of an empty evidence table for 'Who needs surveillance?'

Clinical question				
Author et al. Title. Journal year;volume:pages				
Study design **Treatment era** **Years of follow-up**	**Participants**	**Treatment**	**Main outcomes**	**Additional remarks**
1. Study design 2. Treatment era 3. Follow-up	1. Type and number of participants 2. Diagnoses 3. Age at diagnosis 4. Age at follow-up 5. Controls (if applicable)	1. Chemotherapy 2. Radiotherapy 3. Surgery 4. Other treatments	1. Outcome definitions 2. Results	1. Strengths 2. Limitations 3. Risk of bias A. Selection bias Low risk/High risk/Unclear Reason: B. Attrition bias Low risk/High risk/Unclear Reason: C. Detection bias Low risk/High risk/Unclear Reason: D. Confounding Low risk/High risk/Unclear Reason:

Evidence table for 'Who needs surveillance?'

Who needs premature ovarian insufficiency surveillance?				
Byrne et al. Early menopause in long-term survivors of cancer during adolescence. Am J Obstet Gynecol 1992;166:788-793				
Study design **Treatment era** **Years of follow-up**	**Participants**	**Treatment**	**Main outcomes**	**Additional remarks**
Study design: Multi-center cohort study Treatment era: 1945-1974 Follow-up: >19 yr after cancer diagnosis	Type and number of participants 1048 female CCS ≥21 years of age at study entry; 954 were menstruating before study entry and 94 became menopausal before they were eligible for the cohort Diagnoses: Female genital cancer (n=90), Hodgkin's disease (n=206), non-Hodgkin's lymphoma (n=31), soft tissue sarcoma (n=115), leukaemia (n=15), brain or CNS tumour (n=133), bone cancer (n=65), other (n=393) Age at diagnosis: Mean 13.6 yr Age at follow-up: Mean 32.3 yr Controls: 1596 menstruating siblings at age 21 yr; Mean age at follow-up 33.0 yr	Chemotherapy only: 68 (6.5%) Alkylating agents and radiotherapy above diaphragm: 38 (3.6%) Alkylating agents and radiotherapy below diaphragm: 79 (7.5%) Radiotherapy only: 261 (24.9%) Surgery only: 493 (47.0%) Sterilizing surgery and chemotherapy and radiotherapy: 25 (2.4%) Other treatments: 84 (8.0%)	Outcome definitions: - Amenorrhea: woman's report of whether she was still having menstrual periods Amenorrhea: - 123/954 (12.9%) menopausal after study entry - 831/954 (87.1%) still menstruating Age-specific relative risks for amenorrhea survivors vs. controls: - All survivors aged 21-25: RR 4.32, 95% CI 2.28-8.17 - All survivors aged 26-30: RR 1.61, p>0.05 - All survivors aged 31-40: RR 0.78, p>0.05 - All survivors aged 41+: RR 0.98, p>0.05 - Alkylating agents alone aged 21-25: RR 9.17, 95% CI 2.67-31.49 - Radiotherapy below diaphragm and alkylating agents aged 21-25: RR 27.39, 95% CI 12.42-60.35 - Radiotherapy below diaphragm and alkylating agents aged 26-30: RR 4.64, p<0.01 - Radiotherapy alone aged 21-25: RR 3.66, 95% CI 1.34-9.99 - Radiotherapy alone aged 26-30: RR 2.41, p<0.05 - Radiotherapy alone aged 31-40: RR 0.90, p>0.05 - Radiotherapy alone aged 41+: RR 1.22, p>0.05 - Aged 0-12 at diagnosis aged 21-30 at follow-up: RR 0.62, p>0.05 - Aged 13-19 at diagnosis aged 21-30 at follow-up: RR 2.32, 95% CI 1.63-3.291	Strengths - study sample Limitations - Self-reported outcome - Control group not representative for general population **Risk of bias** A. Selection bias: Unclear Reason: unclear how many patients were included in the original cohort of survivors B. Attrition bias: Low risk Reason: 90% of eligible survivors completed follow-up assessment. At follow-up, 10% of the survivors and 1% of the controls had died C. Detection bias: Unclear Reason: unclear if the outcome assessors were blinded for important determinants related to the outcome D. Confounding: High risk Reason: Controls not matched to cases

Example of an empty evidence table for 'What surveillance modality should be used?'

Clinical question				
Author et al. Title. Journal year;volume:pages				
Study design **Treatment era** **Years of follow-up**	**Participants**	**Diagnostic tests**	**Main outcomes**	**Additional remarks**
1. Study design 2. Treatment era 3. Follow-up	1. Type and number of participants 2. Age at diagnosis 3. Age at follow-up 4. Cancer treatment 5. Prevalence/risk of late effect	1. Diagnostic test(s) 2. Outcome definitions	1. Diagnostic outcomes (sensitivity, specificity, PPV, NPV, ROC)	

Evidence table for 'What surveillance modality should be used?'

What surveillance modality should be used to detect impaired spermatogenesis?				
Green et al. Lack of specificity of plasma concentrations of inhibin B and follicle-stimulating hormone for identification of azoospermic survivors of childhood cancer: a report from the St Jude Lifetime Cohort Study. J Clin Oncol 2013;31:1324-1328				
Study design **Treatment era** **Years of follow-up**	**Participants**	**Diagnostic tests**	**Main outcomes**	**Additional remarks**
1. Study design Single-centre cohort study 2. Treatment era Not reported 3. Follow-up ≥10 years from diagnosis	1. Type and number of participants 275 male CCS who received gonadotoxic treatment, were not receiving exogenous androgens, had received <40 Gy hypothalamic/pituitary irradiation and were ≥18 years of age at follow-up 2. Age at diagnosis Range 0-21 years 3. Age at follow-up Median 30.5 (19.7-59.1) years 4. Gonadotoxic treatment 275 (100%); Alkylating agents, direct testicular irradiation or <40 Gy hypothalamic/pituitary irradiation (≥40 Gy hypothalamic/pituitary irradiation specifically excluded from analysis) 5. Prevalence azoospermia: 105 (38.2%)	1. Diagnostic test Inhibin B, FSH and inhibin B:FSH ratio 2. Cut-off levels for azoospermia - Inhibin B ≤31 ng/L - FSH >11.5 IU/L - Inhibin B:FSH ratio ≤2.52 pg/mIU ROC analysis determined the optimal cut-off levels	**Azoospermia** Sensitivity - Inhibin B: 100% - FSH: 78.1% - Inhibin B:FSH ratio: 75.3% Specificity - Inhibin B: 45.0% - FSH: 74.1% - Inhibin B:FSH ratio: 74.5% Negative predictive value - Inhibin B: 100% - FSH: 84.6% - Inhibin B:FSH ratio: 83.5% Positive predictive value - Inhibin B: 52.1% - FSH: 65.1% - Inhibin B:FSH ratio: 63.8% Area under the ROC curve - Inhibin B: 0.72 - FSH: 0.83 - Inhibin B:FSH ratio: 0.83	298 out of 485 (61.4%) eligible survivors participated in this study; 23 treated with ≥40 Gy hypothalamic/pituitary irradiation and/or tumour in hypothalamic/pituitary region excluded. Inhibin B was measured in 238 patients and FSH was measured in 275 patients. Patient sample divided into a learning set (n=140) and a validation set (n=135) by random assignment: diagnostic values were similar. Results are shown for the combined data sets.

Appendix 4

Risk of bias assessment of observational studies

	Internal validity
Study group	Selection bias Is the study group representative? yes/no/unclear Yes if: • the study group consisted of more than 75% of the original cohort of childhood cancer survivors • *or* it was a random sample with respect to the cancer treatment
Follow-up	Attrition bias Is the follow-up adequate? yes/no/unclear Yes if: • the outcome was assessed for more than 75% of the study group
Outcome	Detection bias Are the outcome assessors blinded for important determinants related to the outcome? yes/no/unclear Yes if: • the outcome assessors were blinded for important determinants related to the outcome
Risk estimation	Confounding Are the analyses adjusted for important confounding factors? yes/no/unclear Yes if: • important prognostic factors (i.e. age, gender, co-treatment, follow-up) were taken adequately into account

Appendix 5

Example of a Summary of findings table of the body of evidence for the risk of premature ovarian insufficiency after procarbazine

Outcome	Study	No. of participants	Follow up (median/mean, range) yr	Alkylating agents Radiotherapy to ovaries	Events	Effect size	Risk of bias
1.5 Risk POI after procarbazine (n=4 studies)	Chemaitilly 2006*	3,390 CCS	>5 yr after cancer diagnosis	Alkylating agents: 49.7%; RT to ovaries: 24.5%	215/3390 (6.3%) amenorrhea within 5 yr after their cancer diagnosis	*Odds ratio (95% CI) for amenorrhea age at diagn 0-12 yr* Procarbazine yes vs. no: OR 3.2 (1.3-7.3)	SB: high risk AB: low risk DB: unclear CF: low risk
	Thomas-Teinturier 2013*	706 CCS	>5 yr after cancer diagnosis	Any alkylating agent: 47.7%; Procarbazine: 7.2%; RT to ovaries: 56.7%	62/706 (8.9%) nonsurgical menopause; 15/706 (2.1%) nonsurgical premature menopause <age 40 yr	*Relative risk (95% CI) for nonsurgical menopause* Procarbazine dose per g/m^2: RR 2.5 (1.4-5.8)	SB: high risk AB: low risk DB: unclear CF: low risk
	Thomas-Teinturier 2015*	108 CCS	>3 yr after cancer treatment	Any alkylating agent: 100%; Procarbazine: 21.9%; RT to ovaries: 17.6%	8/108 (7.6%) altered ovarian function (↑ FSH, ↓ AMH and amenorrhoea)	*Mean FSH* Procarbazine dose: β 0.012, p<0.001; (Each unit increase in procarbazine dose, mean FSH values increased by 0.012 IU/L)	SB: high risk AB: low risk DB: unclear CF: low risk
	Levine 2018*	2,930 CCS	>5 yr after cancer diagnosis	Alkylating agents: 46.5%; Procarbazine: 201 (7.2%); RT to ovaries: 55.4%	110/2930 (3.8%) nonsurgical premature menopause <age 40 yr	*Odds ratio (95% CI) for nonsurgical premature menopause* Procarbazine dose <4000 mg/m^2 vs. 0: OR 3.07 (0.76-12.43); Procarbazine dose ≥4000 mg/m^2 vs. 0: OR 8.96 (5.02-16.00)	SB: high risk AB: low risk DB: unclear CF: low risk
GRADE assessment:							
Study design:	+4	Retrospective cohort studies					
Study limitations:	-1	Limitations: Selection bias high in 4/4; Attrition bias low in 4/4; Detection bias unclear in 4/4; Confounding low in 4/4					
Consistency:	0	No important inconsistency, all show effect of procarbazine					
Directness:	0	Results are direct, population and outcomes broadly generalizable					
Precision:	0	No important imprecision, large sample size, high total number of events and narrow confidence intervals					
Publication bias:	0	Unlikely					
Effect size:	0	No large magnitude of effect					
Dose-response:	+1	Dose response relationship as higher doses are associated with an increased risk as compared to lower doses					
Plausible confounding:	0	No plausible confounding					
Quality of evidence:		⊕⊕⊕⊕ HIGH					
Conclusion:		Increased risk of POI after procarbazine vs. no procarbazine in female cancer survivors diagnosed before age 25 years. (4 studies significant effect; 7,134 participants; 395 events; 4 multivariable analyses)					

Abbreviations: AB, attrition bias; AMH, anti-Müllerian hormone; CCS, childhood cancer survivors; CF, confounding; DB, detection bias; FSH, follicle stimulating hormone; SB, selection bias; yr, year.
* Overlap in included patients in studies of Chemaitily 200 and Levine 2018; and Thomas-Teinturier 2013 and 2015.

Appendix 6

GRADE quality assessment

Initial score based on type of evidence

- +4: RCTs/ SR of RCTs
- +2: CCTs or observational evidence (e.g., cohort, case-control) for intervention questions
- +4: Observational evidence for etiologic, prognostic and diagnostic questions

Factors that might decrease the quality of the body of evidence

1. Study limitations: risk of bias based on selection bias, attrition bias, detection bias and confounding as defined in the risk of bias table.

    - 0: No problems
    - -1: Problem with 1 element
    - -2: Problem with 2 elements
    - -3: Problem with 3 or more elements

2. Consistency: degree of consistency of effect between or within studies

    - 0: All/most studies show similar results
    - -1: Lack of agreement between studies (statistical heterogeneity / conflicting result, e.g. effect sizes in different directions)

3. Directness: the generalizability of population and outcomes from each study to the population of interest

    - 0: Population and outcomes broadly generalizable
    - -1: Problem with 1 element (population different from the defined inclusion criteria OR outcomes different from the defined inclusion)
    - -2: Problem with 2 elements (population and outcomes)

4. Precision: the precision of the results

    - 0: No important imprecision when studies include many patients and many events and thus have narrow confidence intervals; Determine with the chairs and advisors what is seen as many patients, many events and narrow confidence intervals
    - -1: Important imprecision when studies include relatively few patients and few events and thus have wide confidence intervals (especially when the confidence interval cross the 0). Another criteria to consider is the clinical decision threshold. This is the threshold of the effect size that would change the decision whether or not to adopt a clinical action. Downgrade if the effect estimate and confidence intervals cross the clinical decision threshold. Determine with the chairs and advisors the clinical decision threshold.
      OR if only one study has been identified

- -2: If there is important imprecision (see -1) AND if only one study has been identified

5. Publication bias: if investigators fail to report studies and outcomes (typically those that show no effect)

    - 0: Publication bias unlikely
    - -1: Risk of publication bias when for example published evidence is limited to industry funded trials

Factors that might increase the quality of the body of evidence

1. Magnitude of effect:

    - +1: Large magnitude of effect; all studies show significant effect sizes (point estimate) >2 or <0.5
    - +2: Very large magnitude of effect; all studies show significant effect sizes (point estimate) >5 or <0.2

2. Dose response gradient:

    - +1: Evidence of clear relation with increases in the outcome with higher exposure levels across or within studies

3. Plausible confounding:

    - +1: If adjustment for confounders would have increased the effect size; for example the estimate of effect is not controlled for the following possible confounders: smoking, degree of education, but the distribution of these factors in the studies is likely to lead to an underestimate of the true effect

Total score

⊕⊕⊕⊕ High quality evidence

⊕⊕⊕⊖ Moderate quality evidence

⊕⊕⊖⊖ Low quality evidence

⊕⊖⊖⊖ Very low quality evidence

Appendix 7

Overall conclusions of evidence table for male gonadotoxicity survveillance

Who needs surveillance?	
Risk of impaired spermatogenesis in male cancer survivors diagnosed before age 25 years	**Quality of evidence**
Increased risk after *alkylating agents* vs. no alkylating agents	⊕⊕⊕⊕ HIGH [1-5]
Increased risk after *increasing doses of alkylating agents*	⊕⊕⊕⊕ HIGH [1-5]
Increased risk after *cyclophosphamide* vs. no cyclophosphamide	⊕⊕⊕⊕ HIGH [1,3-5]
Increased risk after *increasing doses of cyclophosphamide*	⊕⊕⊕⊕ HIGH [1,3-5]
Increased risk after *procarbazine* and *mechlorethamine* (given as part of multi-agent treatment) vs. no procarbazine and mechlorethamine	⊕⊖⊖⊖ VERY LOW [2]
Increased risk after *increasing doses of procarbazine* and *mechlorethamine* (given as part of multi-agent treatment)	⊕⊖⊖⊖ VERY LOW [2]
Unknown risk after *dacarbazine*	No studies
No significant effect of *dacarbazine dose*	⊕⊖⊖⊖ VERY LOW [2]
Unknown risk after busulfan, chlorambucil, ifosfamide, melphalan, thiotepa, carmustine (BCNU), lomustine (CCNU)	No studies
Unknown risk after *antimetabolites*	No studies
Unknown risk after *platinum compounds*	No studies
Increased risk after *radiotherapy exposing the testes* vs. no radiotherapy exposing the testes	⊕⊖⊖⊖ VERY LOW [3,6]
Unknown risk after *higher vs. lower doses of radiotherapy exposing the testes*	No studies
Unknown risk after *gonadotoxic chemotherapy combined with radiotherapy exposing the testes*	No studies
Unknown risk after *unilateral orchiectomy combined with radiotherapy exposing the testes*	No studies
Unknown risk after novel agents (tyrosine kinase inhibitors, demethylating agents, oxaliplatin)	No studies

Appendix 8

Evidence to decision (EtD) framework - template

	Criteria	Judgements	Research evidence	Additional considerations
PROBLEM	Is the problem a priority?	☐ No ☐ Probably no ☐ Uncertain ☐ Probably yes ☐ Yes		
BENEFITS AND HARMS	What is the overall certainty of this evidence?	☐ No included studies ☐Very low ☐ Low ☐ Moderate ☐ High		
	Is there important uncertainty about how much people value the main outcomes?	☐ Important uncertainty or variability ☐ Possibly important uncertainty or variability ☐ Probably no important uncertainty or variability ☐ No important uncertainty or variability ☐ No known undesirable outcomes		
	Are the desirable anticipated effects large?	☐ No ☐ Probably no ☐ Uncertain ☐ Probably yes ☐ Yes ☐ Varies		
	Are the undesirable anticipated effects small?	☐ No ☐ Probably no ☐ Uncertain ☐ Probably yes ☐ Yes ☐ Varies		
	Are the desirable effects large relative to undesirable effects?	☐ No ☐ Probably no ☐ Uncertain ☐ Probably yes ☐ Yes ☐ Varies		

RESOURCE USE	Are the resources required small?	☐ No ☐ Probably no ☐ Uncertain ☐ Probably yes ☐ Yes ☐ Varies		
	Is the incremental cost small relative to the net benefits?	☐ No ☐ Probably no ☐ Uncertain ☐ Probably yes ☐ Yes ☐Varies		
EQUITY	What would be the impact on health inequities?	☐ Increased ☐ Probably increased ☐ Uncertain ☐ Probably reduced ☐ Reduced ☐ Varies		
ACCEPTABILITY	Is the option acceptable to key stakeholders?	☐No ☐ Probably no ☐ Uncertain ☐ Probably yes ☐ Yes ☐ Varies		
FEASIBILITY	Is the option feasible to implement?	☐ No ☐Probably no ☐ Uncertain ☐ Probably yes ☐ Yes ☐ Varies		

Overall conclusions

Balance of consequences				
Undesirable consequences *clearly outweigh* desirable consequences in most settings ☐	Undesirable consequences *probably outweigh* desirable consequences in most settings ☐	The balance between desirable and undesirable consequences *is closely balanced or uncertain* ☐	Desirable consequences *probably outweigh* undesirable consequences in most settings ☐	Desirable consequences *clearly outweigh* undesirable consequences in most settings ☐

Appendix 9

Criteria for grading the recommendations

Grade of Recommendation **Conclusions of evidence according to GRADE**	**Strong recommendation to do** Benefits >>> risk & harms	**Moderate recommendation to do** Benefits > or = risk & harms	**Recommendation not to do** No benefit / Potentially harm
High quality of evidence Consistent evidence from well performed and high quality studies or systematic reviews (low risk of bias, direct, consistent, precise).	Strong recommendation based on high quality evidence	Moderate recommendation based on high quality evidence	Recommendation not to do based on high quality evidence
Moderate quality of evidence Evidence from studies or systematic reviews with few important limitations.	Strong recommendation based on moderate quality evidence	Moderate recommendation based on moderate quality evidence	Recommendation not to do based on moderate quality evidence
Low to very low quality of evidence Evidence from studies with serious flaws, only expert opinion, or standards of care.	Strong recommendation based on expert opinion	Moderate recommendation based on (very) low quality evidence Diverging expert opinions	Recommendation not to do based on expert opinion
	Wording in recommendations:		
	Is recommended …	Is reasonable …	Is not recommended …

Appendix 10

Example timeline for guideline development work

Tasks to be finished	By whom	Estimated time to complete task
Preparation phase		
Compose working group Identify coordinator and WG leaders	Chairs and advisors	Allow 2 months for this before:
Conference call: introduction and composition of working groups	Total group	
Development phase		
Step 1		
Develop protocol Define scope of the guideline, i.e. outcomes and population of interest Evaluate concordances/discordances Formulate clinical questions	Chairs, coordinator, WG leaders and advisors	Allow 2 months for this before:
Step 2		
Send scope of the guideline and clinical questions to WG members	Coordinator	Allow 1 month for this before:
Conference call: discuss scope of the guideline and clinical questions	Total group	
Step 3		
Finalize clinical questions Develop search strategy Define in- and exclusion criteria	Chairs, coordinator, WG leaders and advisors	Allow 2 months for this before:
Conference call: discuss search strategy and in- and exclusion criteria	Total group	
Perform literature search	Cochrane Childhood Cancer Group	Allow 2 months for this before:
Conference call: discuss steps for evidence selection	Total group	
Send results literature search and instructions to WG leaders and members	Coordinator	
Select evidence based on search Send final inclusion of eligible studies to coordinator	Coordinator, WG leaders and members	Allow 2-3 months for this depending on number of articles
Step 4		
Conference call: discuss steps for summarizing the evidence	Total group	
Make evidence tables	Coordinator, WG leaders and members	Allow 1-2 months for this before:
Conference call: if necessary to discuss difficulties	Total group or separate WGs	Allow 1 month to make modifications before:
Circulate evidence tables to the whole group	Coordinator	
Each WG checks evidence tables (missing studies, completeness, etc) Return comments evidence tables	WG leaders to coordinate within their WG	Allow 1 month for this before:
Agree final evidence tables	Total group	
Conference call: discuss and agree final evidence tables and outline next steps for formulating overall conclusions of the evidence	Total group	
Develop conclusion of evidence tables	Chairs, coordinator, WG leaders and advisors	Allow 2-3 months for this before:
Circulate conclusions of evidence tables to the working group members	Coordinator	

Conference call: discuss and agree final conclusions of evidence tables	Total group	Allow 2 months to make modifications before:
Step 5		
Formulate draft recommendations	Chair, coordinator, advisors and WG leaders	Allow 2 months before:
Conference call: discuss draft recommendations	Total group	Allow 2 months to make modifications before:
Discuss and develop final recommendations, preferably in a face-to-face meeting	IGHG, PCSF and external experts	