	Directness (Painless EBM)				Validity (ROBINS-I)							Precision (Painless EBM)		Applicability (Painless EBM)						Local Applicability Considerations			No. of Serious Risk			Judgement		
Study, Year RT Technique	Population	Intervention	Comparator	Outcomes	Confounding Bias	Selection Bias	Misclassification Bias (Intervention)	Analysis Bias (Missing Data)	Measurement Bias (Outcomes)	Reporting Bias (Results)	Overall Bias	Effect Size	Precision of Estimate	Sex	Co-Morbidities	Race	Age	Pathology	Socioeconomic Status (Cost)	Availability of Technology	Availability of Expertise	Organizational Requirements	Internal Validity and Reliability	External Validity	Local Applicability	Decision	Certainty	Applicability
Demoor-Goldschmidt, 2017	Serious risk Unpredictable Mixed population, of which any about 30% was WT (20% in the intervention group, 43% in the comparator group).	Low risk Gated WLI (IGRT)	Low risk Free-breathing WLI (2D)	Low risk Acute and late lung or liver toxicity	Serious risk Unpredictable Other cancers were included, for which the treatment paradigm, chemotherapy regimen and toxicity, and stage-specific prognoses were different. Baseline liver function was evaluated.	Moderate risk Favours comparator The experimental group is represented by older patients, for which the prognosis is worse.	Low risk Institutional protocols for the intervention and the comparator are defined.	Low risk All patients were accounted for.	Low risk Standard grading criteria for toxicity were applied (RTOG/EORTC).	Low risk Analysis of toxicity outcomes was straightforward. Minimal risk for multiplicity and selective reporting.	Moderate risk	Low risk The effect sizes are moderate (acute G≥3 pneumonitis, 1.40; acute G≥3 hepatotoxicity, 0.73; late hepatotoxicity, 0.73) but statistically insignificant.	Serious risk The estimates are very imprecise (acute G≥3 pneumonitis, 0.16-12.60; acute G≥3 hepatotoxicity, 0.16-32.93; late hepatotoxicity, 0.16-32.93) due to small sample size and absence or scarcity of events.	No information Probably an insignificant factor to survival outcomes and lung or liver toxicity.	No information Towards null	No information Likely Caucasians. Racial disparity in terms of prognosis or lung or liver toxicity not clearly defined for WT.	Serious risk Favours experimental The experimental group is represented by older patients, which is important in terms of the feasibility of the breath-hold.	Serious risk Unpredictable Other cancers were included, for which the treatment paradigm, chemotherapy regimen and toxicity, and stage-specific prognoses were different.	Serious risk Favours experimental The study was conducted in a first-world country (France). IGRT remains costly in the PH.	Serious risk IGRT has limited availability.	Serious risk Proper physics training is required. RT planning objectives remain to be defined.	Serious risk Gated WLI may require prolonged treatment duration. Technique may require a dedicated linear accelerator and anesthesia team for children.	3 1. Population (mixed) 2. Confounding bias 3. Precision of estimate (small sample size)	4 1. Age (older, able to cooperate) 2. Pathology (mixed) 3. Socioeconomic Status (first world, cost)	3 1. Technology availability 2. Availability of expertise 3. Organizational requirements	Downgrade	Very Low	Very Low
Morgan, 2018	Low risk Survivors treated with WLI for WT	Low risk 3D WLI (no intentional thyroid-sparing)	Moderate risk No comparator. Dosimetric data reported. Results could be interpreted based on QUANTEC. PENTEC is under development.	Low risk Primary hypothyroidism	Moderate risk Favours comparator Baseline thyroid function data not reported. However, the prevalence of hypothyroidism in children is probably low.	Low risk Consecutive sampling was used.	Low risk All 20 received the intervention as defined (3D WLI) except for one (IMRT).	Serious risk Attrition rate, 29%.	Low risk Institutional reference ranges for thyroid function tests were defined. Outcome was also defined by need for medical treatment.	Low risk Analysis of toxicity outcomes was straightforward. Minimal risk for multiplicity and selective reporting.	Moderate risk	Low risk The incidence (10%) is low and might be what could be expected given the reported range of mean thyroid doses (1.3-11.70y).	Low risk The estimate (1-32%) might be what could be expected given the reported range of mean thyroid doses (1.3-11.70y).	Low risk M:F ratio for the cohort was 0.56. Hypothyroidism is more common in women in the postmenopausal age. All patients in the cohort were <21 of age during treatment and surveillance.	No information Towards null	No information Likely Caucasians. Racial disparity in terms of developing RT-related hypothyroidism is not known.	Moderate risk Median age (5) and age range (1-9) is older than typical WT cohort.	Low risk	Low risk Favours experimental While this study was conducted in a first-world country (US), 3DCRT is already widely available and affordable in the PH.	Low risk 3DCRT is largely available.	Low risk	Low risk Technique may or may not require a dedicated linear accelerator and anesthesia team for children.	1 1. Analysis bias (high attrition)	0	0	Maintain	Low	High
Kalapurakal, 2019	Serious risk Unpredictable Mixed population, of which 25% was WT (15% was newly diagnosed, non-relapsed WT).	Low risk Cardiac-sparing WLI (IMRT)	Moderate risk No comparator for the clinical outcomes. Dosimetric comparison with 2D WLI is reported. Results could be interpreted based on QUANTEC. PENTEC is under development.	Low risk 2y lung-metastasis progression-free survival; overall survival Acute lung toxicity Late lung or cardiac toxicity	Moderate risk Unpredictable Other cancers were included, for which the treatment paradigm, chemotherapy regimen and toxicity, and stage-specific prognoses were different. These were discussed in the interpretation of results. An adequate baseline cardiac function was required. All patients were treated per CDG protocols where cardiotoxic drug modifications were defined.	Low risk Exclusion criteria were defined to limit effects of confounding factors such as prior WLI or mediastinal disease that required additional RT doses.	Low risk Despite the study being multicenter, the RT protocol was defined and credentialing was done.	Low risk A minimum follow-up of 2 years was required and observed.	Low risk Follow-up protocol (frequency and modalities) was defined. Tumor response was assessed using RECIST. Cardiac evaluations included electrocardiograms and echocardiograms.	Low risk Analysis of toxicity outcomes was straightforward. Minimal risk for multiplicity and selective reporting.	Low risk	Low risk The incidence (0%) is low and might be what could be expected given the reported dosimetric cardiac-sparing when compared to 2D plans.	Serious risk The estimate is very imprecise (0-52%) due to the small number of WT cases included (5).	Low risk M:F ratio for the cohort was 1.22. Cardiopulmonary toxicity due to RT not known to be more common in either sex.	No information Towards null	No information Likely Caucasians. Racial disparity in terms of developing RT-related hypothyroidism is not known.	Moderate risk Favours experimental The experimental group is represented by older patients, which is important in terms of the feasibility of the IMRT.	Low risk	Serious risk Favours experimental The study was conducted in a first-world country (US). While IMRT is largely available, it remains costly in the PH.	Low risk IMRT is largely available.	Serious risk RT planning objectives remain to be defined.	Serious risk Cardiac sparing IMRT WLI may require prolonged treatment duration. Technique may require a dedicated linear accelerator and anesthesia team for children.	2 1. Population (mixed) 2. Imprecise estimate	2 1. Socioeconomic status (first world, cost)	2 1. Availability of expertise 2. Organizational requirements	Maintain	Low	Select
Chen, 2020	Low risk WT treated with WAI ± boost per SIOP 2001	Low risk Kidney-sparing WAI (VMAT)	Moderate risk Favours comparator No flank or abdominal RT indicated and given per SIOP 2001. Dosimetric comparison with 3D WAI is reported.	Low risk Acute intestinal or renal toxicity Late renal toxicity	Low risk Baseline creatinine clearance data were available. One patient with traumatic injury to the contralateral kidney during surgery was noted in the discussion of toxicity outcomes.	Moderate risk Favours comparator Consecutive sampling was used. Seven had WAI per SIOP 2001. The seven controls were selected from the remaining 11 who did not have RT. It was unclear whether matching was performed. However, M:F (5.2 and 6:1) and R:L (4:3 and 4:3) were similar. Median age and age ranges were different (4.8, 2.4-14.0 and 1.7, 0.8-2.9). Younger patients are known to have better tolerance of chemo side effects.	Low risk The RT protocol for VMAT (as well as for 3D, for dosimetric comparison) is described.	Low risk All patients were accounted for.	Low risk There was no definition for grading acute intestinal toxicity ("no significant acute intestinal toxicity"); although this does not seem to be the primary outcome of interest of the study. Before and after treatment creatine clearances were compared. Renal damage was defined as abnormal creatinine clearance, orather blood, urine, imaging or histopathologic markers.	Low risk Analysis of toxicity outcomes was straightforward. Minimal risk for multiplicity and selective reporting.	Low risk	Moderate risk Favours comparator The risk for late renal toxicity is 3x that of patients who had no RT. However this is not statistically significant.	Serious risk The estimate is very imprecise (RR 0.14-63.15) due to the small sample size and absence or scarcity of events.	Low risk Similar M:F between the two groups. RT nephrotoxicity not known to be more common in either sex.	No information Unpredictable	No information Likely Latinos. Racial disparity in terms of developing RT-related hypothyroidism is not known.	Low risk The experimental group is represented by older patients, which is probably important in terms of tolerance of the combined nephrotoxicity of chemotherapy and radiotherapy.	Low risk	Serious risk Favours experimental The study was conducted in newly industrialized country (Brazil) like the Philippines. VMAT is not largely available and remains costly.	Serious risk VMAT has limited availability.	Serious risk Proper physics training is required. RT planning objectives remain to be defined.	Moderate risk Kidney sparing VMAT WAI may not necessarily require prolonged treatment duration. Technique may or may not require a dedicated linear accelerator and anesthesia team for children.	1 1. Precision of estimate (small sample size)	2 1. Socioeconomic status (cost)	2 1. Availability of technology 2. Availability of expertise	Maintain	Low	Select
Mul, 2021	Moderate risk Towards null WT treated with highly conformal flank RT ± WLI per SIOP-RTSG protocols MRTK and CCSK histologies were included.	Low risk Highly conformal flank RT (VMAT)	Low risk No comparator for the clinical outcomes.	Low risk Locoregional control rate Disease-free survival Overall survival	Low risk Various stages were included. The local and systemic treatments remain appropriate as patients were treated according to defined SIOP-RTSG protocols. The details of the stages were discussed in the cases that relpased.	Moderate risk Towards null MRTK and CCSK histologies were included, thich are associated with more aggressive behavior. **Poor prognostic features other than aggressive histologies are present in up to 19%: IVC thrombus, 19%, abdominal boost for gross disease required: 11%).** The local and systemic treatments remain appropriate as patients were treated according to defined SIOP-RTSG protocols. The details of the histologies were discussed in the cases that relpased.	Low risk The RT protocol for highly conformal VMAT is described.	Low risk All patients were accounted for.	Low risk Follow-up protocol (frequency and modalities) was defined. Definitions for events of interest (local, abdominal, regional, and distant relapse; infield, marginal, outfield local or abdominal relapse) were provided and are valid.	Low risk Analysis of toxicity outcomes was straightforward. Minimal risk for multiplicity and selective reporting. All relapses were reported and classified.	Low risk	Low risk The locoregional control (94%), disease-free survival (91%), and overall survival (94%) rates are high and seem appropriate given the cohort characteristics.	Low risk The estimates are imprecise: locoregional control (86-100%), disease-free survival (81-100%), and overall survival (86-100%), but seem appropriate given the cohort characteristics.	Low risk M:F 1:1. No known disparity in disease control or survival in WT due to sex.	No information Unpredictable	No information Likely Caucasians. No known racial disparity in terms of disease control or survival.	Low risk The cohort is represented by the typical presentation age 3.3y (0.3-14.0) for WT.	Moderate risk Aggressive histologies represent 11% (MRTK, 8%; CCSK, 3%) of the cohort. These received appropriate treatment.	Serious risk Favours experimental The study was conducted in a first-world country (Netherlands). VMAT remains costly in the PH.	Serious risk VMAT has limited availability.	Serious risk Proper physics training is required. RT planning objectives remain to be defined. Consensus regarding target volume delineation has been defined.	Serious risk Highly conformal VMAT flank RT may not necessarily require prolonged treatment duration. Technique may or may not require a dedicated linear accelerator and anesthesia team for children. Placement of surgical clips may be recommended for proper delineation to target volume.	0	1 1. Socioeconomic status (first world, cost)	3 1. Technology availability 2. Availability of expertise 3. Organizational requirements	Maintain	Low	Select
Kalapurakal, 2003 RT Timing	Low risk FH WT treated with flank-RT or WAI enrolled in NWTS-3 and NWTS-4	Low risk Flank RT or WAI at >9 days post-op The cutoff was chosen based on the NWTS-1 results.	Low risk Flank RT or WAI at ≤9 days post-op The cutoff was chosen based on the NWTS-1 results.	Low risk Flank recurrence Abdominal recurrence	Moderate risk Patients who received preoperative chemotherapy, with bilateral WT, WT arising from fused kidney, extrarenal WT, or "adult" WT (aged ≥16) were excluded. No information on toxicity associated with early RT and its effect on treatment compliance. Patients perceived with high recurrence risk could have been pushed to get early RT more frequently. Stage 3 is heterogenous, which could not have been accounted for even when accounting for stage.	Low risk All stage II, III, or IV FH WT treated with flank-RT or WAI enrolled to the NWTS-3 and -4 were included.	Low risk The cutoff was chosen based on the NWTS-1 results. While the overall range of RT delay was reported to be 1-277, about 98% treated within 30 days; 59% at days 8-12. The majoritywould be considered to have truly received adjuvant RT, even in the current practice of delaying RT until 6 weeks when evaluating for need for WLI.	Low risk All patients were accounted for.	Low risk Follow-up protocol was defined per NWTS-3 and -4. Definitions for events of interest (local, and abdominal relapse) were provided and are valid. Tumor recurrences in the liver (hematogenous spread), or in the opposite kidney were excluded.	Low risk Any local or abdominal relapse was the sole outcome of interest.	Low risk	Low risk The results suggest that the delayed RT is associated with lower risk (HR 0.57) for flank recurrence, but slightly higher risk for abdominal recurrence (HR 1.09). However, these are not statistically significant. Overall, there is no difference between delayed or early RT. This is true for the NWTS-3, NWTS-4, and the entire cohort.	Low risk The confidence intervals include 1.00, suggesting no statistical difference.	Low risk M:F 1:1.16. No known disparity in disease controll in WT due to sex.	No information Unpredictable	No information Likely Caucasians. No known racial disparity in terms of disease control.	No information Unpredictable The age profile of the cohort was not described. In the eligibility criteria, adult WT (aged ≥16) was excluded.	Low risk FH WT was included. Bilateral WT, WT arising from fused kidney, extrarenal WT, or "adult" WT (aged ≥16) were excluded.	Serious risk Favours comparator The study was conducted in a first-world country (US).	Moderate risk Early RT may require improvement of surgical techniques, adoption of neoadjuvant chemotherapy approach, and enhanced peri-operative care.	Moderate risk Early RT may require improvement of surgical techniques, adoption of neoadjuvant chemotherapy approach, and enhanced peri-operative care.	Serious risk Early RT by 9 days may require very close coordination among the pediatric, surgical, and radiotherapy services. Large sample size, allowing for precise estimates. Data retrieved from (prospective) clinical trials, with predefined selection criteria and intervention protocols.	0 Confounding bias well managed through restriction and multivariate analysis.	1 1. Socioeconomic status (first world)	1 1. Organizational requirements	Upgrade	Moderate	High
Stokes, 2018	Low risk U/FH WT treated with flank RT or WAI	Low risk Flank RT or WAI at >9 days post-op Flank RT or WAI at >14 days post-op The cutoff was chosen to validate the clinical practice recommendation for moving the cutoff for early RT from ≤9 to ≤14 days.	Low risk Flank RT or WAI at ≤9 days post-op Flank RT or WAI at ≤14 days post-op	Low risk Mortality	Low risk Patients with bilateral disease or age >25 were excluded.	Low risk U/FH WT treated with flank RT or WAI identified from the NCDB, from 2004-2013, for which the time from surgery to radiotherapy, and survival status is known or derivable.	Low risk Patients with SRI of negative values (reflecting neoadjuvant RT) or >60 days (possible salvage RT) were excluded.	Moderate risk All patients were accounted for. An additional 243 might have been included had surgery-radiotherapy interval been derivable. This represents an attrition of 14%.	Low risk Survival status was determined from the NCDB.	Low risk Mortality (survival) was the sole outcome of interest.	Low risk	Low risk Among patients with non-metastatic disease, delayed RT of >14 days is associated with >2x higher mortality risk, which is statistically significant (p=0.013). This is supported by an increased risk by 1.04 when the delay (SRI) is taken as a continuous variable on multivariate analysis. The risk is not statistically different for delay of >9 days, or for patients with metastatic disease.	Low risk The estimated risk is 1.17 to 1.87 with delayed RT of >14 days among patients with non-metastatic disease. The estimate seems appropriate considering that both UFH and FH WT were included.	Low risk The M:F for the entire cohort is 1:1.22; the non-metastatic cohort, 1:1.19, and the metastatic cohort, 1:1.34. No known disparity in disease controll in WT due to sex.	No information Unpredictable	Low risk The majority are represented by whites, 61.2%, blacks, 18.7%, and hispanics, 13.9%. Other races constituted 6.2%. No known racial disparity in terms of survival.	Low risk The mean age (4.3 for the entire cohort; 4.2, non-metastatic cohort; 4.7, metastatic cohort) is typical. In the eligibility criteria, adult WT (aged ≥25) was excluded.	Low risk FH and UFH WT were included. Bilateral WT and adult WT were excluded.	Serious risk Favours comparator The study was conducted in a first-world country (US).	Moderate risk Early RT may require improvement of surgical techniques, adoption of neoadjuvant chemotherapy approach, and enhanced peri-operative care.	Moderate risk Early RT may require improvement of surgical techniques, adoption of neoadjuvant chemotherapy approach, and enhanced peri-operative care.	Moderate risk Early RT by 14 days may require close coordination among the pediatric, surgical, and radiotherapy services.	Confounding bias well managed through restriction, stratification, and multivariate analysis. Large sample size, allowing for precise estimates.	1 1. Socioeconomic status (first world)	0	Upgrade	Moderate	High