Included studies and clinical practice guidelines hepatic toxicity surveillance

Studies liver injury and iron overload

Year	Bibliography
Liver injury	
2019	Mulder et al. Hepatic late adverse effects after antineoplastic treatment for childhood cancer. Cochrane Database Syst Rev 2019;4:CD008205.
2019	Green et al. Serum ALT elevations in survivors of childhood cancer. A report from the St. Jude Lifetime Cohort Study. Hepatology 2019;69:94-106.
2013	Mulder et al. Surveillance of hepatic late adverse effects in a large cohort of long-term survivors of childhood cancer: prevalence and risk factors. Eur J Cancer 2013;49:185-93.
Iron overload	
2017	Sirvent et al. Prevalence and risk factors of iron overload after hematopoietic stem cell transplantation for childhood acute leukemia: a LEA study. Bone Marrow Transplant 2017;52:80-87.
2014	Ruccione et al. Characterization of transfusion-derived iron deposition in childhood cancer survivors. Cancer Epidemiol Biomarkers Prev 2014;23:1913-1919.
2009	Chotsampancharoen et al. Iron overload in survivors of childhood leukemia after allogeneic hematopoietic stem cell transplantation. Pediatr Transplant 2009;13:348-352.

Studies focal nodular hyperplasia and nodular regenerative hyperplasia

Year	Bibliography
Focal nodular hyperplasia	
2020	Cattoni et al. Hepatic focal nodular hyperplasia after pediatric hematopoietic stem cell transplantation: The impact of hormonal replacement therapy and iron overload. Pediatr Blood Cancer 2020;67:e28137.
2015	Pillon et al. Focal nodular hyperplasia of the liver: an emerging complication of hematopoietic SCT in children. Bone Marrow Transplantation 2015;50:414-419.
2013	Masetti et al. Benign hepatic nodular lesions after treatment for childhood cancer. JPGN 2013;56:151-155.
2013	Masetti et al. Focal nodular hyperplasia of the liver in children after hematopoietic stem cell transplantation. Pediatr Transplantation 2013;17:479-486.
2012	Smith et al. Incidence and etiology of new liver lesions in pediatric patients previously treated for malignancy. AJR 2012;199:186-191.
2009	Sudour et al. Focal nodular hyperplasia of the liver following hematopoietic SCT. Bone Marrow Transplantation 2009;43:127-132.
2003	De Bouyn et al. Hepatic focal nodular hyperplasia in children previously treated for a solid tumor. Cancer 2003;97:3107-13.
Nodular regenerative hyperplasia	
2012	Yoo et al. Dynamic MRI findings and clinical features of benign hypervascular hepatic nodules in childhood cancer survivors. AJR 2013;201:178-184.
2000	Brisse et al. Hepatic regenerating nodules: a mimic of recurrent cancer in children. Pediatr Radiol 2000;30:386-393.

Clinical practice guidelines

Year	Bibliography
Liver injury	
2019	EASL Clinical Practice Guidelines: Drug-induced liver injury. European Association for the Study of the Liver. J Hepatol 2019;70:1222-1261.
2018	Newsome et al. Guidelines on the management of abnormal liver blood tests. Gut 2018;67:6-19.
2017	Shiha et al. Asian-Pacific Association for the Study of the Liver (APASL) consensus guidelines on invasive and non-invasive assessment of hepatic fibrosis: a 2016 update. Heptol Int 2017;11:1-30.
Iron overload	
2018	Valent et al. Diagnosis, management and response criteria of iron overload in myelodysplastic syndromes (MDS): updated recommendations of the Austrian MDS platform. Expert Review of Hematology 2018; 11:109-116.
2014	Porter et al. Chapter 3 Iron overload and chelation. In: Cappellini MD, Cohen A, Porter J, et al., editors. Guidelines for the Management of Transfusion Dependent Thalassaemia (TDT) [Internet]. 3rd edition. Nicosia (CY): Thalassaemia International Federation; 2014.
2011	Bacon et al. Diagnosis and Management of Hemochromatosis: 2011 Practice Guideline by the American Association for the Study of Liver Diseases. Hepatol 2011;54:328-343.
2010	EASL Clinical practice guidelines: HFE Hemochromatosis. European Association for the Study of the Liver. J Hepatol 2010;53:3-22.

Evidence tables hepatic toxicity surveillance

Cellular liver injury (ALT) and biliary tract injury (gGT)

Who needs surveillance?				
Mulder et al. Hepatic late adverse effects after antineoplastic treatment for childhood cancer. Cochrane Database Syst Rev 2019;4:CD008205.				
Study design **Treatment era** **Years of follow-up**	**Participants**	**Treatment**	**Main outcomes**	**Additional remarks**
Systematic review including 33 cohort studies examining the risk of hepatic late adverse effects in childhood cancer survivors Treatment era: 1962-2006 Follow-up: Ranging from median 2.0 yr since end of treatment - 25.1 yr since primary cancer diagnosis	7,876 childhood cancer survivors (ranging from 19-2,753 per study) aged <21 years at primary cancer diagnosis Primary cancer diagnosis: Various Age at primary cancer diagnosis: Median age ranging from 0.2-10.2 yr Age at follow-up: Median age ranging from 9.7-32.0 yr	Chemotherapy: 31/33 studies Radiotherapy to fields involving the liver: 14/33 studies HSCT: 15/33 studies Hepatectomy: 4/33 studies	Outcome definitions: - Liver histology: liver fibrosis, cirrhosis - Cellular liver injury: elevated ALT or AST - Hepatobiliary dysfunction or biliary tract injury: elevated gGT, ALP or bilirubin - Liver synthetic dysfunction: abnormal prothrombin time or albumin *Cut-off limit for normal and abnormal liver enzyme values as specified by the authors of the original Studies* Prevalence elevated ALT > upper limit normal: 5.8%-52.8% in 8 studies Prevalence elevated ALT > twice upper limit normal: 0.9%-44.8% in 4 studies Prevalence elevated AST > upper limit normal: 1.1-13.0% in 2 studies Prevalence elevated AST > twice upper limit normal: 2.3% in 1 study Prevalence elevated gGT > upper limit normal: 5.3% in 1 study Prevalence elevated gGT > twice upper limit normal: 0.9% in 1 study	Strengths Comprehensive search strategy Limitations Heterogeneity of included studies **Risk of bias:** - Selection bias: Low risk: 6/33 studies High risk: 8/33 studies Unclear: 19/33 studies - Attrition bias: Low risk: 28/33 studies High risk: 2/33 studies Unclear: 3/33 studies - Detection bias: Low risk: 29/33 studies High risk: 0/33 studies Unclear: 4/33 studies - Confounding: Low risk: 2/33 studies High risk: 16/33 studies Not applicable: 15/33 studies

			<u>Prevalence elevated ALP > upper limit normal:</u> 4.3%-11.1% in 2 studies <u>Prevalence elevated bilirubin > upper limit normal:</u> 0.0%-8.7% in 3 studies <u>Risk factors for cellular liver injury (elevated ALT):</u> - Radiotherapy to fields involving the liver (especially after a high percentage of the liver irradiated), - Higher BMI - Longer follow-up time - Older age at evaluation increased the risk of cellular liver injury in multivariable analyses in 2 studies - Busulfan - Thioguanine - Hepatic surgery - Chronic viral hepatitis C - Metabolic syndrome - Use of statins - Non-Hispanic white ethnicity - Higher alcohol intake (> 14 units per week) increased the risk of cellular liver injury in multivariable analyses in 1 study Chronic viral hepatitis was shown to increase the risk of cellular liver injury in 6 univariable studies <u>Risk factors for biliary tract injury (elevated gGT):</u> - Radiotherapy involving the liver - Higher BMI - Higher alcohol intake (> 14 units per week) - Longer follow-up time - Older age at cancer diagnosis increased the risk of biliary tract injury in a multivariable analysis in 1 study	

Footnote 1: For the risk of bias, results of the Cochrane systematic review are shown. Criteria for risk of bias assessment by Cochrane may slightly differ from the IGHG criteria.
Footnote 2: More detailed results regarding risk factors are shown in the evidence tables of Green 2019 and Mulder 2013.

Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase, gGT, gamma-glutamyltransferase; HSCT, hematopoietic stem cell transplantation.

Who needs surveillance?				
Green et al. Serum ALT elevations in survivors of childhood cancer. A report from the St. Jude Lifetime Cohort Study. Hepatology 2019;69:94-106.				
Study design Treatment era Years of follow-up	**Participants**	**Treatment**	**Main outcomes**	**Additional remarks**
Retrospective cohort study Treatment era: 1962-2000 Follow-up: Median 23.2 (interquartile range 17.6-29.7) yr since diagnosis	4,421 childhood cancer survivors of whom 2,753 were included in the study group; 2,751 underwent liver function testing Primary cancer diagnosis: Various Age at primary cancer diagnosis: Median 7.4 (interquartile range 3.3-13.2) yr Age at follow-up: Median 31.4 (interquartile range 25.8-37.8) yr Hepatitis virus infection: - 7/73 (9.6%) hepatitis B seropositive - 98/1,578 (6.2%) hepatitis C seropositive Acute liver disease: 12/2,751 (0.4%) SOS BMI: - 763 (27.7%) overweight - 959 (34.9%) obese	Chemotherapy: - Methotrexate: 1,328 (48.3%) - High-dose methotrexate: 747 (27.2%); median 15,212.9 (interquartile range 4,064.5 to 21,697.3) mg/m^2 - Mercaptopurine: 1,072 (39.0%) - Thioguanine: 26 (0.9%) - Dactinomycin: 400 (14.5%) - Busulfan: 23 (0.8%) - Carmustine: 12 (0.4%) - Melphalan: 5 (0.2%) - Asparaginase: 935 (34.0%) Radiotherapy to fields involving the liver: - 437/2,751 (15.9%) - Hepatic irradiation: 368 (13.4%) - TBI: 69 (2.5%) Radiotherapy dose: - Median percentage of liver that received 10 Gy: 51.4% - Median percentage of liver that received 15 Gy: 34.6%	Outcome definitions: Hepatocellular injury: ALT > upper limit of normal (Either ≥ 19 U/L for females and ≥ 30 U/L for males; or ≥ 40 U/L according to institutional standards) Prevalence elevated ALT > upper limit normal according to sex-specific standards: 1,137/2,751 (41.3%) Prevalence elevated ALT > upper limit normal according to institutional standards: 419/2,751 (15.2%) Risk factors for hepatocellular liver injury (ALT > upper limit of normal) according to sex-specific values using multivariable Poisson regression analysis: - Radiotherapy to fields involving the liver treated to ≥ 15 Gy per 10% volume increase: RR 1.06 (95% CI 1.03-1.08)* - Busulfan vs. none: RR 1.54 (95% CI 1.02-2.33)* - Thioguanine vs. none: RR 1.38 (95% CI 1.02-1.85)* - Hepatic surgery vs. none: RR 1.90 (95% CI 1.45-2.49)* - Age at evaluation per yr: RR 1.01 (95% CI 1.00-1.01)* - BMI ≥ 25 vs. <25: RR 1.60 (95% CI 1.42-1.81)* - Hepatitis C grade ≥1 vs. <1: RR 1.76 (95% CI 1.52-2.02)*	**Risk of bias:** - Selection bias: high risk, study group consisted of 2,753/4,421 (62.3%) of the original cohort of survivors. - Attrition bias: low risk, outcome was assessed in 2,751/2,753 (99.9%) of the study group. - Detection bias: unclear if blinding of outcome assessment, but the outcome measurement was not likely to be influenced by lack of blinding. - Confounding: low risk, analyses were adjusted for cancer treatment and follow-up.

		- Median percentage of liver that received 20 Gy: 25.3% <u>Hepatectomy:</u> 24/2,751 (0.9%) <u>HSCT:</u> - Total: 76/2,751 (2.8%) - Allogeneic: 47 (1.7%) - Autologous: 29 (1.1%) (2 participants included who underwent both allogeneic and autologous HSCT) <u>Blood transfusion:</u> nm	- Metabolic syndrome vs. none: RR 1.40 (95% CI 1.26-1.55)* - Statins (atorvastatin, rosuvastatin, simvastatin) vs. none: RR 1.20 (95% CI 1.02-1.42)* - Non-Hispanic white ethnicity vs. non-Hispanic black or other: RR 1.37 (95% CI 1.18-1.58)* (Analysis with radiotherapy involving liver treated to ≥ 20 Gy provided comparable results) - No significant effect of methotrexate, mercaptopurine, dactinomycin, HSCT, alcohol intake, gender, educational level and age at diagnosis in univariable analysis and therefore not included in the multivariable model	

Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase, BMI, body mass index; gGT, gamma-glutamyltransferase; HSCT, hematopoietic stem cell transplantation; nm, not mentioned; RR, relative risk; SOS: sinusoidal obstruction syndrome; TBI, total body irradiation; *, significant.

Who needs surveillance?				
Mulder et al. Surveillance of hepatic late adverse effects in a large cohort of long-term survivors of childhood cancer: prevalence and risk factors. Eur J Cancer 2013;49:185-93.				
Study design **Treatment era** **Years of follow-up**	**Participants**	**Treatment**	**Main outcomes**	**Additional remarks**
Retrospective cohort study Treatment era: 1966-2003 Follow-up: Median 12.4 (range 5.0-36.1) yr since diagnosis	1,795 childhood cancer survivors of whom 1,404 were included in the study group; 1,362 underwent liver function testing Primary cancer diagnosis: Various Age at primary cancer diagnosis: Median 5.9 (range 0.0-17.8) yr Age at follow-up: Median 19.5 (range 5.8-47.0) yr Hepatitis virus infection: 0/1362 (0.0%) (participants with hepatitis virus infection excluded according to eligibility criteria for the study) Acute liver disease: 0/1362 (0.0%) SOS (participants with SOS excluded according to eligibility criteria for the study)	Chemotherapy: - Any: 1,204/1,362 (88.4%) - Methotrexate: 392 (28.8%) - Mercaptopurine: 352 (25.8%) - Thioguanine: 98 (7.2%) - Dactinomycin: 397 (29.1%) - Busulphan: 10 (0.7%) - Other antimetabolites: 426 (31.3%) - Other cytotoxic antibiotics: 633 (46.5%) - Other alkylating agents: 715 (52.5%) - Plant alkaloids: 1115 (81.9%) - Other chemotherapeutics: 837 (61.5%) Chemotherapy dose: nm Radiotherapy to fields involving the liver: - 123/1,362 (9.0%) - Abdomen: 102 (7.5%) - TBI: 21 (1.5%) Radiotherapy dose: Median 20.0 (5.0 to 46.0) Gy Hepatectomy: 35/1,362 (2.6%)	Outcome definitions: - Hepatocellular injury: ALT > upper limit of normal (≥ 34 U/L for females, ≥ 45 U/L for males and children < 15 years) - Biliary tract injury: gGT > upper limit of normal (≥ 40 U/L for females, ≥ 60 U/L for males, ≥ 56 U/L for children < 15 years) Prevalence elevated ALT > upper limit normal: 79/1,362 (5.8%) Prevalence elevated ALT > twice upper limit normal: 12/1,362 (0.9%) Prevalence elevated gGT > upper limit normal: 68/1,295 (5.3%) Prevalence elevated gGT > twice upper limit normal: 12/1,295 (0.9%) Risk factors for hepatocellular liver injury (ALT > upper limit of normal) in multivariable logistic regression analysis: - Radiotherapy to fields involving the liver vs. none: OR 2.34 (95% CI 1.07-5.13)* - Methotrexate vs. none: OR 1.22 (95% CI 0.53-2.84) - Mercaptopurine vs. none: OR 0.84 (95% CI 0.36-1.99) - Thioguanine vs. none: OR 1.40 (95% CI 0.38-5.18)	**Risk of bias:** - Selection bias: low risk, study group consisted of 1,404/1,795 (78.2%) of the original cohort of survivors. - Attrition bias: low risk, outcome was assessed in 1,362/1,404 (97.0%) of the study group. - Detection bias: unclear if blinding of outcome assessment, but the outcome measurement was not likely to be influenced by lack of blinding. - Confounding: low risk, analyses were adjusted for cancer treatment and follow-up.

	BMI: nm	BMT: 61/1,362 (4.5%) Blood transfusion: nm	- Dactinomycin vs. none: OR 0.71 (95% CI 0.29-1.76) - Busulfan vs. none: OR 3.9 (95% CI 0.29-32.90) - Other antimetabolites vs. none: OR 0.61 (95% CI 0.24-1.56) - Other cytotoxic antibiotics vs. none: OR 1.91 (95% CI 1.00-3.68) - Other alkylating agents vs. none: OR 0.63 (95% CI 0.32-1.26) - Plant alkaloids vs. none: OR 2.14 (95% CI 0.85-5.38) - Other chemotherapeutic agents vs. none: OR 0.90 (95% CI 0.38-2.17) - Liver resection vs. none: OR 1.87 (95% CI 0.38-9.07) - Age at diagnosis in yr: OR 1.06 (95% CI 1.00-1.13) - Time since cancer diagnosis in yr: OR 1.10 (95% CI 1.05-1.15)* - Male vs. female: OR 1.18 (95% CI 0.67-2.08) - BMI z-score: OR 1.67 (95% CI 1.37-2.03)* - Alcohol intake <7 units per week vs. none: OR 1.21 (95% CI 0.63-2.30) - Alcohol intake 7-14 units per week vs. none: OR 0.87 (95% CI 0.33-2.31) - Alcohol intake >14 units per week vs. none: OR 1.67 (95% CI 1.37-2.03)* Risk factors for biliary tract injury (gGT > upper limit of normal) in multivariable logistic regression analysis: - Radiotherapy to fields involving the liver vs. none: OR 5.45 (95% CI 2.51-11.82)* - Methotrexate vs. none: OR 0.70 (95% CI 0.27-1.81) - Mercaptopurine vs. none: OR 0.64 (95% CI 0.25-1.64) - Thioguanine vs. none: OR 0.51 (95% CI 0.09-2.80) - Dactinomycin vs. none: OR 0.46 (95% CI 0.17-1.21) - Busulfan vs. none: OR 4.03 (95% CI 0.33-48.94)	

			- Other antimetabolites vs. none: OR 0.81 (95% CI 0.31-2.10) - Other cytotoxic antibiotics vs. none: OR 1.45 (95% CI 0.72-2.91) - Other alkylating agents vs. none: OR 0.89 (95% CI 0.43-1.87) - Plant alkaloids vs. none: OR 2.65 (95% CI 0.96-7.31) - Other chemotherapeutic agents vs. none: OR 1.61 (95% CI 0.60-4.30) - Liver resection vs. none: OR 1.09 (95% CI 0.12-9.69) - Age at diagnosis in yr: OR 1.08 (95% CI 1.01-1.15)* - Time since cancer diagnosis in yr: OR 1.13 (95% CI 1.07-1.18)* - Male vs. female: OR 0.71 (95% CI 0.38-1.31) - BMI z-score: OR 1.43 (95% CI 1.14-1.81)* - Alcohol intake <7 units per week vs. none: OR 0.96 (95% CI 0.48-1.93) - Alcohol intake 7-14 units per week vs. none: OR 1.14 (95% CI 0.43-3.01) - Alcohol intake >14 units per week vs. none: OR 3.04 (95% CI 1.16-7.96)*	

Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase, BMI, body mass index; BMT, bone marrow transplantation; gGT, gamma-glutamyltransferase; nm, not mentioned; OR, odds ratio; SOS: sinusoidal obstruction syndrome; TBI, total body irradiation; *, significant.

Iron overload

Who needs surveillance? At what frequency should surveillance be performed?				
Sirvent et al. Prevalence and risk factors of iron overload after hematopoietic stem cell transplantation for childhood acute leukemia: a LEA study. Bone Marrow Transplant 2017;52:80-87.				
Study design **Treatment era** **Years of follow-up**	**Participants**	**Treatment**	**Main outcomes**	**Additional remarks**
Prospective cohort study Treatment era: 1982-2011 Follow-up: Mean 9.98 ± 0.35 yr from HSCT to last visit	420 childhood leukemia survivors treated with HSCT of whom 384 had at least one ferritin value Primary cancer diagnosis: ALL (68.2%), AML (31.8%) Age at HSCT: Median 8.8 ± 0.25 yr Age at follow-up: nm Hepatitis virus infection: nm Acute liver disease: 140/322 (43.5%) allogeneic HSCT recipients grade II-IV acute GVHD or extensive chronic GVHD	HSCT: - Total: 384/384 (100%) - Allogeneic: 322 (83.9%) - Autologous: 62 (16.1%) TBI-based regimen: 257/384 (66.9%) Busulfan-based regimen: 127/384 (33.1%) Blood transfusion: nm	Outcome definitions: Iron overload: serum ferritin level ≥350 ng/ml with an erythrocyte sedimentation rate at an hour <50 mm (value is slightly above the upper limit of normal of most laboratories) Prevalence iron overload: - 162 (42.2%, 95% CI 37.2-47.2%) serum ferritin level ≥350 ng/ml - 51 (13.3%, 95% CI 10.1-17.2%) serum ferritin level ≥1000 ng/mL Risk factors for iron overload (serum ferritin level ≥350 ng/ml) using multivariable logistic regression analysis: - Age at HSCT >4.7-≤8.2 yr vs. ≤4.7 yr: OR 1.46 (95% CI 0.70-3.05) - Age at HSCT >8.2-≤12.7 yr vs. ≤4.7 yr: OR 5.36 (95% CI 2.63-10.95)* - Age at HSCT >12.7 yr vs. ≤4.7 yr: OR 7.64 (95% CI 3.73-15.64)* - AML vs. ALL: OR 3.23 (95% CI 1.47-7.13)* - Allogeneic, sibling vs. autologous: OR 2.53 (95% CI 1.20-5.33)* - Allogeneic, alternative donor vs. autologous: OR 4.34 (95% CI 2.07-9.12)* - TBI-based regimen vs. busulfan-based regimen: OR 2.45 (95% CI 1.09-5.53)* - Status at transplant >complete remission 1 vs. complete remission 1:	Evaluation of liver iron concentration by MRI was recommended in patients with iron overload. 127/384 had at least 2 ferritin evaluations, of whom 68 with untreated iron overload. Among the 162 patients with a serum ferritin level ⩾ 350 ng/mL, 17 underwent liver iron concentration by MRI. Liver iron concentration was above the upper limit of the normal range (i.e. 2 mg/gdw) in 17/17 patients. There was a statistically significant correlation between serum ferritin level and liver iron concentration by MRI. **Risk of bias:** - Selection bias: unclear how many eligible survivors were included in the original cohort of survivors. - Attrition bias: low risk, outcome was assessed in

			OR 1.27 (95% CI 0.79-2.06) <u>Risk factors for iron overload (serum ferritin level ≥1000 ng/ml) using multivariable logistic regression analysis:</u> - Age at HSCT >4.7-≤8.2 yr vs. ≤4.7 yr: OR 3.87 (95% CI 0.78-19.31) - Age at HSCT >8.2-≤12.7 yr vs. ≤4.7 yr: OR 5.08 (95% CI 1.05-24.65)* - Age at HSCT >12.7 yr vs. ≤4.7 yr: OR 20.40 (95% CI 4.57-91.14)* - AML vs. ALL: OR 1.55 (95% CI 0.60-4.02) - Allogeneic, sibling vs. autologous: OR 2.14 (95% CI 0.66-6.95) - Allogeneic, alternative donor vs. autologous: OR 2.88 (95% CI 0.90-9.18) - TBI-based regimen vs. busulfan-based regimen: OR 1.22 (95% CI 0.44-3.37) - Status at transplant >complete remission 1 vs. complete remission 1: OR 2.04 (95% CI 1.03-4.03)* <u>Risk factors for iron overload (serum ferritin level ≥350 ng/ml) in 322 allogeneic HSCT recipients using multivariable logistic regression analysis:</u> - Age at HSCT >4.7-≤8.2 yr vs. ≤4.7 yr: OR 1.58 (95% CI 0.73-3.44) - Age at HSCT >8.2-≤12.7 yr vs. ≤4.7 yr: OR 5.53 (95% CI 2.59-11.79)* - Age at HSCT >12.7 yr vs. ≤4.7 yr: OR 8.50 (95% CI 3.94-18.33)* - AML vs. ALL: OR 3.22 (95% CI 1.44-7.22)* - Allogeneic, alternative donor vs. sibling: OR 1.79 (95% CI 1.07-2.98)* - TBI-based regimen vs. busulfan-based regimen: OR 2.16 (95% CI 0.93-5.01) - Significant GVHD yes vs. no: OR 1.80 (95% CI 1.09-2.99)* - Status at transplant >complete remission 1 vs. complete remission 1:	384/420 (91.4%) of the study group. - <u>Detection bias:</u> unclear if blinding of outcome assessment, but the outcome measurement was not likely to be influenced by lack of blinding. - <u>Confounding:</u> high risk, analyses were adjusted for cancer treatment and age at HSCT, but not for red blood cell transfusions.

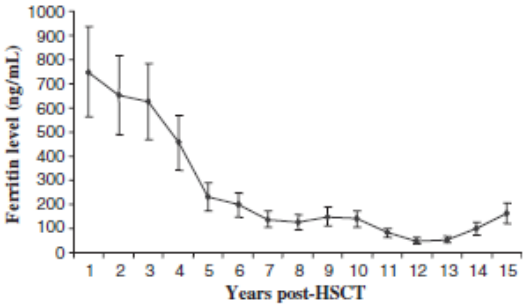
			OR 1.08 (95% CI 0.64-1.82) <u>Risk factors for iron overload (serum ferritin level ≥1000 ng/ml) in 322 allogeneic HSCT using multivariable logistic regression analysis:</u> - Age at HSCT >4.7-≤8.2 yr vs. ≤4.7 yr: OR 4.11 (95% CI 0.81-20.74) - Age at HSCT >8.2-≤12.7 yr vs. ≤4.7 yr: OR 3.86 (95% CI 0.76-19.50) - Age at HSCT >12.7 yr vs. ≤4.7 yr: OR 20.34 (95% CI 4.48-92.39)* - AML vs. ALL: OR 1.51 (95% CI 0.56-4.07) - Allogeneic, alternative donor vs. sibling: OR 1.39 (95% CI 0.70-2.79) - TBI-based regimen vs. busulfan-based regimen: OR 1.18 (95% CI 0.40-3.48) - Significant GVHD yes vs. no: OR 1.70 (95% CI 0.86-3.35) - Status at transplant >complete remission 1 vs. complete remission 1: OR 1.74 (95% CI 0.85-3.54)* <u>Iron overload over time:</u> Mean serum ferritin levels decreased from 883 ng/mL at first evaluation to 581 ng/mL at last evaluation (mean time between evaluations 3.68 yr) among 68 survivors with untreated iron overload with at least 2 measurements.	

Abbreviations: ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; GVHD, graft-versus-host disease; HSCT, hematopoietic stem cell transplantation; nm, not mentioned; OR, odds ratio; TBI, total body irradiation; *, significant.

Who needs surveillance?				
Ruccione et al. Characterization of transfusion-derived iron deposition in childhood cancer survivors. Cancer Epidemiol Biomarkers Prev 2014;23:1913-1919.				
Study design **Treatment era** **Years of follow-up**	**Participants**	**Treatment**	**Main outcomes**	**Additional remarks**
Retrospective cohort study Treatment era: 2004-2009 Follow-up: Median 4.4 (range 0.2-7.6) yr; Median 4.9 (range 1.4-7.9) since last transfusion	75 childhood cancer survivors Primary cancer diagnosis: ALL (30.7%), AML (13.3%), germ cell tumor (18.7%), osteosarcoma (12.0%), Ewing sarcoma (9.3%), Wilms tumor (5.3%), rhabdomyosarcoma (9.3%), nasopharyngeal carcinoma (1.4%) Age at diagnosis: Median 7.7 (range 1.8-20.2) yr Age at follow-up: Median 14 (range 8-25.6) yr Hepatitis virus infection: nm Acute liver disease: nm Hemochromatosis mutations: 0/75 (0%) C282Y homo; H63D-C282Y hetero	HSCT: 4 (5.3%) Treatment intensity: - Least-moderately intensive: 39 (52.0%) - Very intensive: 22 (29.3%) - Most intensive: 14 (18.7%) PRBC transfusions: - Total: 67 (89.3%) - <10: 31 (41.3%) - ≥10: 36 (48.0%)	Outcome definitions: - Iron concentration by MRI Liver: LIC >1.2mg/g abnormal Pancreas: R2*>30 Hz abnormal Heart: T2* <20 ms abnormal - Iron status by serum markers (abnormal values not reported) Prevalence abnormal iron concentration by MRI: - Liver: 36/73 (49.3%) - Pancreas: 19/72 (26.4%) - Heart: 0/74 (0%) Prevalence abnormal iron serum markers: - 21/74 (28.4%) elevated serum ferritin levels - 2/71 (2.8%) elevated serum iron levels - 3/75 (4%) elevated iron-binding capacity - 5/71 (7%) elevated percentage of transferrin saturation Risk factors for increased liver iron concentration using multivariable regression analysis: - Weight-adjusted cumulative PRBC volume (p<0.0001) associated with a 0.03 mg/g increase in LIC for each mL/kg transfused - Age at diagnosis (p<0.0001) associated with a 0.29 mg/g increase in LIC for each 1-year increase in age - Final model explained 52% of the variance in LIC	Blood samples were obtained on the same day as the MRI evaluation. Liver/pancreas R2* and heart T2*, indicators of tissue iron content, were assessed using multiecho gradient echo technique. Both hepatic and pancreatic R2* were positively correlated with serum ferritin, serum iron, the percentage of transferrin saturation, and weight-adjusted cumulative PRBC volume. Having undergone HSCT was associated with increased liver iron concentration (P <0.0001) in univariable analysis, but because there were only 4 patients in this group, this variable was not included in the multivariable analyses. **Risk of bias:** - Selection bias: high risk, 75/157 (47.8%) eligible survivors were included in

				the original cohort of survivors. - Attrition bias: low risk, outcome was assessed in 73/75 (97.3%) of the study group. - Detection bias: low risk, in addition to the study radiologist's blinded interpretation of MRIs relative to study aims, a separate, blinded radiologist reviewed MRIs for evidence of non-cardiac abnormalities. - Confounding: low risk, analyses were adjusted for cancer treatment and age at diagnosis.

Abbreviations: ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; GVHD, graft-versus-host disease; HSCT, hematopoietic stem cell transplantation; nm, not mentioned; OR, odds ratio; PRBC, packed red blood cells.

At what frequency should surveillance be performed?				
Chotsampancharoen et al. Iron overload in survivors of childhood leukemia after allogeneic hematopoietic stem cell transplantation. Pediatr Transplant 2009;13:348-352.				
Study design Treatment era Years of follow-up	**Participants**	**Treatment**	**Main outcomes**	**Additional remarks**
Longitudinal prospective cohort study Treatment era: 1990-2005 Follow-up: Mean 5.6 (range 1-15) yr from HSCT	133 childhood cancer survivors Primary cancer diagnosis: ALL (51%), AML (58%), CML (24%) Age at HSCT: Mean 9.1 (range 0.6-21.4) yr Age at follow-up: 18% had features of chronic GVHD Hepatitis virus infection: nm Acute liver disease: nm	HSCT: 133 (100%) allogeneic TBI: 127 (95.5%) TBI dose: Range 8-14.4 Gy Blood transfusions: 133 (100%)	Outcome definitions: Serum ferritin >110 ng/ml Prevalence abnormal serum ferritin at 1 yr post-HSCT: 124 (93.2%) Serum ferritin levels over time: Mean serum ferritin level at 1 yr post-HSCT was 1158 (range 22-3264) ng/ml and declined over time  	The total number of serum ferritin measurements per patients is not reported. **Risk of bias:** - Selection bias: unclear how many survivors were included in the original cohort of survivors. - Attrition bias: unclear for how many patients serum ferritin was longitudinally measured over time. - Detection bias: unclear if blinding of outcome assessment, but the outcome measurement was not likely to be influenced by lack of blinding.

Abbreviations: ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; CML, chronic myeloid leukemia; HSCT, hematopoietic stem cell transplantation; nm, not mentioned.

Clinical practice guidelines

Cellular liver injury

What surveillance modality should be used?	
EASL Clinical Practice Guidelines: Drug-induced liver injury. European Association for the Study of the Liver. J Hepatol 2019;70:1222-1261.	
Recommendation[1]	**Level of evidence[2]**
Grade C: ALT, ALP and total bilirubin are the standard analytes to define liver damage and liver dysfunction in drug-induced liver injury. AST values can be used to reliably substitute ALT in calculating the pattern of injury when the latter is unavailable at drug-induced liver injury recognition, whereas gGT is less reliable as an ALP substitute.	Extrapolation from level 2b studies (exploratory cohort studies with good reference standards)
Grade B: Persistently elevated total bilirubin and ALP in the second month from drug-induced liver injury onset should be used as a marker for chronic drug-induced liver injury.	Level 1b studies (individual inception cohort studies).

Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase, gGT, gamma-glutamyltransferase.

[1] Grades of recommendation from the Oxford Centre for Evidence-based Medicine
A: Consistent level 1 studies.
B: Consistent level 2 or 3 studies *or* extrapolations from level 1 studies.
C: Level 4 studies *or* extrapolations from level 2 or 3 studies.
D: Level 5 evidence *or* troublingly inconsistent or inconclusive studies of any level.

[2] Level of evidence based on the Oxford Centre for Evidence-based Medicine and recommended for EASL CPGs Level
1: Systematic reviews (SR) (with homogeneity) of randomized controlled trials (RCT); *Further research is unlikely to change our confidence in the estimate of* benefit and risk.
2: RCT or observational studies with dramatic effects; SR of lower quality studies (i.e. non-randomized, retrospective).
3: Non-randomized controlled cohort/follow-up study/control arm of randomized trial (SR is generally better than an individual study); Further research (if performed) is likely to have an impact on our confidence in the estimate of benefit and risk and may change the estimate.
4: Case-series, case-control, or historically controlled studies (systematic review is generally better than an individual study).
5: Expert opinion (mechanism-based reasoning); Any estimate of effect is uncertain.

What surveillance modality should be used?	
Newsome et al. Guidelines on the management of abnormal liver blood tests. Gut 2018;67:6-19.	
Recommendation[1]	**Level of evidence**[2]
Grade B: Initial investigation for potential liver disease should include bilirubin, albumin, ALT, ALP and gGT, together with a full blood count if not already performed within the previous 12 months.	Level 2b
Grade D: Abnormal liver blood test results should only be interpreted after review of the previous results, past medical history and current medical condition.	Level 5
Grade D: The extent of liver blood test abnormality is not necessarily a guide to clinical significance. This is determined by the specific analyte which is abnormal (outside the reference range) and the clinical context.	Level 5

Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; gGT, gamma-glutamyltransferase.

[1] **Grades of recommendation**

A: Consistent level 1 studies.
B: Consistent level 2 or 3 studies *or* extrapolations from level 1 studies.
C: Level 4 studies *or* extrapolations from level 2 or 3 studies.
D: Level 5 evidence *or* troublingly inconsistent or inconclusive studies of any level.

[2] **Level of evidence**

1: Systematic reviews (SR) (with homogeneity) of randomized controlled trials (RCT); Further research is unlikely to change our confidence in the estimate of benefit and risk.
2: RCT or observational studies with dramatic effects; SR of lower quality studies (i.e. non-randomized, retrospective).
3: Non-randomized controlled cohort/follow-up study/control arm of randomized trial (SR is generally better than an individual study); Further research (if performed) is likely to have an impact on our confidence in the estimate of benefit and risk and may change the estimate.
4: Case-series, case-control, or historically controlled studies (systematic review is generally better than an individual study).
5: Expert opinion (mechanism-based reasoning); Any estimate of effect is uncertain.

What surveillance modality should be used?	
Shiha et al. Asian-Pacific Association for the Study of the Liver (APASL) consensus guidelines on invasive and non-invasive assessment of hepatic fibrosis: a 2016 update. Heptol Int 2017;11:1-30.	
Recommendation[1]	**Level of evidence**[2]
Grade 1: Conventional ultrasound cannot be used for the diagnosis of hepatic fibrosis.	B
Grade 1: Conventional ultrasound for diagnosis of early cirrhosis should be confirmed by additional studies.	C
Grade 1: Conventional CT and MRI have higher specificity and sensitivity than conventional ultrasound for the diagnosis of cirrhosis.	A
Grade 1: Transient elastography is an established technique and is recommended as the initial assessment for significant liver fibrosis and cirrhosis.	A
Grade 1: Transient elastography is a highly reproducible and user-friendly technique for assessing liver fibrosis in patients with chronic liver disease. However, because transient elastography reproducibility is significantly reduced in patients with steatosis, increased BMI, lower degrees of hepatic fibrosis and narrow intercostal spaces, caution could be warranted in the clinical use of transient elastography as surrogate for liver biopsy.	A
Grade 1: Liver biopsy is considered as the gold standard for diagnosing liver fibrosis, but sampling errors and both intra- and inter-observer agreement on biopsy samples may lead to poor reproducibility for many liver biopsies, and the procedure is invasive and expensive.	A
Grade 2: Compared to liver biopsy, FibroTest and transient elastography are more cost-effective.	A

[1] Grades of recommendation adapted from GRADE
1: Strong; Factors influencing the strength of the recommendation included the quality of the evidence, presumed patient-important outcomes, and cost.
2: Weak; Variability in preferences and values, or more uncertainty. Recommendation is made with less certainty, or higher cost or resource consumption.

[2] Level of evidence adapted from GRADE
A: High; Further research is very unlikely to change confidence in the estimate of the clinical effect.
B: Moderate; Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
C: Low or vert low; Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. Any estimate of effect is uncertain.

Iron overload

What surveillance modality should be used?	
Valent et al. Diagnosis, management and response criteria of iron overload in myelodysplastic syndromes (MDS): updated recommendations of the Austrian MDS platform. Expert Review of Hematology 2018; 11:109-116.	
Recommendation[1]	**Level of evidence**
Serum ferritin is recommended for daily practice in myelodysplastic syndromes (+++). Specificity of the ferritin test per se is low. However, when liver enzymes and inflammation parameters are also tested, elevated ferritin levels are a sensitive and rather specific parameter of the total body iron burden.	Not reported
Laboratory parameters and imaging studies to determine iron overload: Transferrin saturation (++) Labile plasma iron (+/-) MRI and SQUID only recommended for patients in whom ferritin levels are very high and a concomitant liver disease or massive inflammation is present (+/-) Non-transferrin-bound free (plasma) iron (-)	Not reported

[1] Grades of recommendation not explained in the manuscript.

What surveillance modality should be used?	
Porter et al. Chapter 3 Iron overload and chelation. In: Cappellini MD, Cohen A, Porter J, et al., editors. Guidelines for the Management of Transfusion Dependent Thalassaemia (TDT) [Internet]. 3rd edition. Nicosia (CY): Thalassaemia International Federation; 2014.	
Recommendation[1]	**Level of evidence**
Grade B: Liver iron concentration (biopsy, SQUID, MRI) can be used to calculate total body iron, and serum ferritin is an approximate marker of liver iron concentration.	Not reported

[1] Grades of recommendation not explained in the manuscript.

What surveillance modality should be used?	
Bacon et al. Diagnosis and Management of Hemochromatosis: 2011 Practice Guideline by the American Association for the Study of Liver Diseases. Hepatol 2011;54:328-343.	
Recommendation[1]	**Level of evidence**[2]
Grade 1: In a patient with suggestive symptoms, physical findings, or family history of hemochromatosis, a combination of transferrin saturation and ferritin should be obtained rather than relying on a single test.	B
Grade 1: Diagnostic strategies using serum iron markers should target high-risk groups such as those with a family history of hemochromatosis or those with suspected organ involvement.	B

[1] **Grades of recommendation**
1: Strong; Factors influencing the strength of the recommendation included the quality of the evidence, presumed patient-important outcomes, and cost.
2: Weak; Variability in preferences and values, or more uncertainty. Recommendation is made with less certainty, or higher cost or resource consumption.

[2] **Level of evidence**
A: High; Further research is unlikely to change confidence in the estimate of the clinical effect.
B: Moderate; Further research may change confidence in the estimate of the clinical effect.
C: Low; Further research is very likely to impact confidence on the estimate of clinical effect.

What surveillance modality should be used?	
EASL Clinical practice guidelines: HFE Hemochromatosis. European Association for the Study of the Liver. J Hepatol 2010;53:3-22.	
Recommendation[1]	**Level of evidence[2]**
Grade 1: Patients with suspected iron overload should first receive measurement of fasting transferrin saturation and serum ferritin.	B

[1] **Grades of recommendation according to GRADE**
1: Strong; Defined as being 'confident that adherence to the recommendation will do more good than harm or that the net benefits are worth the costs'.
2: Weak; Defined as being 'uncertain that adherence to the recommendation will do more good than harm OR that the net benefits are worth the costs'.

[2] **Level of evidence according to GRADE**
A: High; Randomized trials that show consistent results, or observational studies with very large treatment effects; Further research is very unlikely to change our confidence in the estimate of effect.
B: Moderate; Randomized trials with methodological limitations, or observational studies with large effect; Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
C: Low and very low; Observational studies without exceptional strengths, or randomized trials with very serious limitations; unsystematic clinical observations (e.g. case reports and case series; expert opinions) as evidence of very-low quality evidence; Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. Any estimate of effect is very uncertain.