Chapter 1

General introduction
For a long time, it has been known that high-dose glucocorticoid therapy may cause suppression of the hypothalamic-pituitary-adrenal (HPA) axis. HPA axis suppression, in turn, may lead to life-threatening complications. Although glucocorticoids are routinely used in treatment schedules for childhood acute lymphoblastic leukemia (ALL) and lymphoma, little is known on the occurrence and duration of HPA axis suppression during treatment for childhood ALL or lymphoma. In addition, there is increasing evidence that high-dose glucocorticoid therapy may lead to long-lasting neuroendocrine and psychosocial sequelae. Nevertheless, no information on these sequelae in survivors of childhood ALL and lymphoma is available. Therefore, the aim of this thesis is to gain more insight in the neuroendocrine and psychosocial consequences of high-dose glucocorticoid therapy during and after treatment for childhood ALL and lymphoma.

**Acute lymphoblastic leukemia**

**Epidemiology and pathogenesis**
In the Netherlands, about 600 children are diagnosed with cancer per year. ALL is the most frequent type, with approximately 120 newly diagnoses each year. ALL is slightly more common in boys than in girls and its incidence is highest at the age of three to five years. ALL is a disease caused by uncontrolled proliferation of immature lymphoblasts. Although the exact cause of childhood ALL is unclear, it is known that it occurs as a result of accumulative genetic damage of the progenitor blood-cells that leads to uncontrolled proliferation, reduced apoptosis and a block in cellular differentiation of immature lymphocytes. Retrospective identification of leukemia-specific genetic abnormalities in neonatal blood spots indicates a prenatal origin. However, additional acquired genetic mutations are needed to develop ALL.  

**Clinical features**
Most symptoms of ALL are the result of accumulation of lymphoblasts in the bone marrow leading to replacement of normal blood cells. This may result in symptoms of anemia (e.g. palor, lethargy), thrombocytopenia (e.g. easy bruising, purpura, bleeding) and neutropenia (e.g. fever, malaise, infections and/or sepsis). In addition, as a result of organ infiltration tender bones, hepatosplenomegaly, lymphadenopathy, testicular enlargement and mediastinal mass may occur. When the central nervous system is involved, children may present with neuropathies and symptoms of increased intracranial pressure (e.g. headache, nausea, vomiting, blurring of vision and diplopia).
Investigations

Besides a hemogram, cytomorphology, immunophenotyping, and cytogenetic and molecular genetic examination of the blood and bone marrow are essential to establish the diagnosis. Furthermore, the presence or absence of lymphoblasts in the cerebrospinal fluid will be examined and in boys, testicular involvement will be investigated. Based on these results, the exact type of ALL and the appropriate treatment will be determined.

Treatment and outcome

Treatment of childhood ALL consists of chemotherapy for a period of at least two years. Children with a specific genetic abnormality which is associated with a poor prognosis, a deletion on 7p12 of the IKZF1 (IKAROS) gene, receive an additional year of treatment. The most important chemotherapeutic agents for the treatment of ALL are glucocorticoids (dexamethasone, prednisone), vincristine, asparaginase, methotrexate and 6-mercaptopurine. Treatment typically consists of an “induction phase” in order to rapidly kill most of the tumor cells and achieve complete remission, a “consolidation phase” in order to eradicate drug-resistant residual leukemic cells, an “M-phase” to prevent central nervous system relapse, a “reinduction phase” which further enhances treatment outcome, and “maintenance therapy” which aims to eliminate minimal residual disease and thus to prevent relapses. Moreover, the success of intensive treatment relies on good supportive care measures.2,3

In the Netherlands, the three most recent ALL treatment protocols are the Dutch Childhood Oncology Group (DCOG) ALL-9, ALL-10 and ALL-11 protocol. The ALL-9 protocol was open from 1997 to 2004. Based on clinical and biological factors, this protocol included two different stratification arms which differ in treatment intensity: non-high risk (70% of the patients) and high risk (30% of the patients). From 2004 to 2012, the ALL-10 protocol was open for inclusion. Based on clinical factors, cytogenetics, response to initial therapy, induction failure and the minimal residual disease measured by polymerase chain reaction (PCR), it comprised three stratification arms; standard risk (29% of the patients), medium risk (67% of the patients), and high risk (4% of the patients). The ALL-11 protocol is open since April 2012 and includes a standard risk, a medium risk and a high risk group as well.4-6

The cure rate of children with ALL has increased from nearly zero in the 1960s to almost 90% today.7,8 Prognostically important factors are age at diagnosis, white blood cell count at diagnosis, genetic abnormalities in the leukemic cells and initial response to therapy.2,3

Lymphoma

Epidemiology and pathogenesis

Lymphoma is the third most frequent type of childhood cancer, with approximately 60 newly diagnosed cases per year in the Netherlands.1 It is the most common malignancy among
adolescents of 15-19 years.\textsuperscript{9} Lymphomas are a group of diseases caused by malignant lymphocytes that accumulate in lymph nodes and cause the characteristic clinical features of lymphadenopathy. The major subdivision of lymphomas is into Hodgkin lymphoma (HL) and non-Hodgkin lymphoma (NHL) based on the histological presence of Reed-Sternberg (RS) cells in HL.

\textit{Hodgkin lymphoma}

In the disease tissue of HL the characteristic multinucleated RS cells, which are usually of B cell origin, are found. The Epstein-Barr virus (EBV) genome has been detected in 50\% or more of the cases in Hodgkin tissue and it has been speculated that the virus' presence may distinguish between aetiologically separate HL entities. However, the precise contribution of EBV remains unclear.\textsuperscript{10} The Revised European American Lymphoma (REAL)/World Health Organization (WHO) lymphoma classification now recognizes two major types of HL: nodular lymphocyte predominant HL (rare in children) and classical HL. The latter consists of four histological subtypes, based on RS cell morphology and the cell composition around the RS cells: nodular sclerosing, mixed cellularity, lymphocyte rich and lymphocyte depleted.\textsuperscript{9} Nodular sclerosis is the most common subtype in adolescents, with the mixed cellularity subtype being relatively more common in children <10 years.\textsuperscript{11} HL has a typical bimodal age distribution with peaks at 15-35 years of age and again after 50 years. Among adolescents, males are affected slightly more often than females.\textsuperscript{9}

\textit{Non-Hodgkin lymphoma}

NHLs are a diverse group of lymphoid tumors whose clinical presentation and natural history are more variable than in HL. Any lymphoma that does not involve RS cells is classified as NHL. The disease can be divided according to the REAL/WHO classification into four subtypes; diffuse large B-cell lymphoma (10-12\% of cases), precursor B-lymphoblastic and precursor T-lymphoblastic lymphoma (1\% and 35-40\%, respectively), Burkitt lymphoma (40-50\% in Europe), and anaplastic large cell lymphoma (ALCL) (5-10\%).\textsuperscript{9,12} The overall incidence of NHL increases progressively from birth until age 80 years and there is a male predominance. The exact cause of NHL is unknown although infectious agents are an important cause in particular subtypes.\textsuperscript{9}

\textit{Clinical features}

Most patients with lymphoma present with asymmetric painless enlargement of lymph nodes in one or more peripheral lymph node regions. Bone marrow involvement is unusual in early disease, but if it occurs bone marrow failure may develop, leading to anemia, thrombocytopenia and neutropenia.
**Hodgkin lymphoma**

The majority of patients present with lymphadenopathy in the cervical and supraclavicular nodes. In addition, 60% of the patients have mediastinal disease. B symptoms, i.e. fever, profuse night sweats and weight loss, are present in 39% of the children. Other constitutional symptoms are pruritis due to eosinophilia, weakness and fatigue.

**Non-Hodgkin lymphoma**

Clinical presentation is heterogeneous. Diffuse large B-cell lymphoma, precursor T-lymphoblastic lymphoma and Burkitt lymphoma are highly aggressive and progress rapidly. B symptoms occur less frequently than in Hodgkin lymphoma. Common sites of NHL are the abdomen, mediastinum, cervical and peripheral lymph nodes, Waldeyer’s ring and bone marrow. Extranodal localisations include the skin, brain, testis, bone or lungs.

**Investigations**

The definitive diagnosis of lymphoma is made by histological examination, assisted by immunophenotypic and cytogenetic analysis, of an excised lymph node. The distinctive multinucleated RS cell is central to the diagnosis of the four classic types of HL. The cluster of differentiation (CD) antigens expressed on the tumor cell surface and certain characteristic chromosomal translocations are essential in the classification of NHL. Furthermore, examination of the bone marrow and the cerebrospinal fluid and body imaging are essential to determine the extent of disease.

**Treatment and outcome**

The selection of the appropriate treatment for HL depends on the type and on staging according to the Ann Arbor classification. Because HL is more radiation sensitive than other forms of lymphoma, treatment comprises radiotherapy in combination with multiagent chemotherapy in high risk groups and chemotherapy alone in lower risk groups. Treatment of NHL mainly depends on the histiotype and on staging according to the St. Jude’s Murphy system and comprises divergent multiagent chemotherapy regimens.

In the Netherlands, HL is treated according to the international EuroNet-Paediatric Hodgkin’s Lymphoma (PHL) Group lymphocyte-predominant (LP) HL protocol and classical (C) HL protocol. Children with a primary B-cell NHL, including Burkitt lymphoma, are treated according to the DCOG B-NHL_B-ALL 2008 protocol. As lymphoblastic lymphoma shares many of the same cell surface markers as well as cytogenetic abnormalities with ALL, it is treated with an ALL-based therapy: the Euro-LB-02 protocol. Finally, ALCL is treated according to the national ALCL protocol. Similar as in childhood ALL, adequate supportive care measures are essential for the success of treatment.

Survival of childhood lymphoma has dramatically improved over the last 25 years. Survival rates of HL depend on risk group, with a five year overall survival of 95% for the low risk group
and up to 85% for the high risk group. The five year event free survival of adolescents with NHL is well-above 90%. However, long-term complications (e.g. cardiac diseases, secondary primary malignancies and infertility) among survivors of childhood lymphoma, that arise from radiation therapy and exposure to anthracyclines and alkylating agents, remain a major concern.

**Glucocorticoids and the HPA axis**

**Glucocorticoids**
The name “glucocorticoid” is a composition of “glucose” (referring to its regulation of glucose levels), “cortex” (referring to its synthesis in the adrenal cortex), and “steroids” (referring to its steroidal structure). Besides its immunosuppressive and anti-inflammatory effects, glucocorticoids have a significant influence on lymphoid cells. They induce G1 cell cycle arrest and they regulate apoptosis of immature thymocytes, pre-B lymphoma cells, mature peripheral T lymphocytes, and several leukemic cell lines. This has led to their wide application in the treatment of leukemia and lymphoma. Despite ongoing research for years, the exact molecular mechanisms leading to glucocorticoid-induced apoptosis remain unraveled.

**HPA axis suppression**
Because of its apoptotic effect, children with ALL and lymphoma receive courses of high-dose glucocorticoids. However, supraphysiological doses may suppress the HPA axis. As synthetic glucocorticoids resemble the hormone cortisol, they induce a negative feedback signal to the hypothalamic paraventricular nucleus which results in reduced secretion of corticotrophin releasing hormone (CRH), and to the anterior pituitary which results in reduced secretion of adrenocorticotropic hormone (ACTH) (Figure 1.1). This in turn, results in adrenal cortex atrophy. Therefore, the adrenal glands may temporarily not be able to generate sufficient cortisol. After normalization of CRH and ACTH levels, it may take a considerable period of time until adrenal size and cortisol production return to normal. Any glucocorticoid can suppress CRH and ACTH secretion, but the degree of suppression depends on the dose, potency, and duration of action of the glucocorticoid and the duration and time of its administration.

The HPA axis regulates glucose levels, body fat and the maintenance of normal blood pressure. In addition, the HPA axis plays an important role in the response to stress and the defense against infections. Stressors such as trauma, surgery or inflammation, but also psychological stress, stimulate the HPA axis, leading to an increase of cortisol production. In addition, increased cortisol levels cause anti-inflammatory effects and inhibition of pro-inflammatory cytokines. Suppression of the HPA axis resulting in inadequate cortisol levels during stress...
(relative hypocortisolism), may lead to life-threatening hypoglycemia and/or hypotension, impaired inflammatory response, and inadequate host defense against infections.23

**HPA axis function during and after treatment of ALL and lymphoma**

Despite improvement of treatment of childhood ALL and lymphoma over time, up to 5% of children die because of the toxic side effects of treatment. The main cause of this treatment-related mortality is infection, associated with cytotoxic and immunosuppressive drugs. HPA axis suppression caused by glucocorticoid therapy could be an important contributing factor to the occurrence and severity of infections.23,24,25 Moreover, several studies report an increase of septic and lethal infections in childhood ALL since prednisone has been substituted for the more potent glucocorticoid dexamethasone.25,26

HPA axis suppression after short-term glucocorticoids in children and adults with ALL was first described in 1979 by Spiegel et al, who found that suppression lasted more than seven days in 36% of the patients and was independent of dosage and duration.27 More recent studies used the low-dose ACTH stimulation test, which is able to detect subtle degrees of adrenal atrophy caused by central adrenal insufficiency.28 Mahachoklertwattana et al. thus found adrenal insufficiency two weeks after glucocorticoid therapy in 46% of the children, persisting in 13% for 20 weeks.29 Einaudi et al. found adrenal axis suppression in 23%, last-

![Figure 1.1 Schematic overview of the HPA axis.](image-url)
ing up to ten weeks.\textsuperscript{30} Most recently, Vestergaard et al. found that the mean elapsed time between end of induction therapy and adrenal sufficiency was 8.5 months.\textsuperscript{31}

Only little information is available regarding HPA axis function after high-dose glucocorticoids in patients with lymphoma. Wilson et al. used the standard insulin hypoglycaemia test in four patients during treatment for lymphoma and described blunted plasma-ACTH responses in all patients.\textsuperscript{32} Zamkoff et al. found depression of basal plasma cortisol in all five HL patients.\textsuperscript{33} Warde et al. reported five of 12 patients with lymphoma with HPA axis suppression between courses of glucocorticoids.\textsuperscript{34} Unfortunately, none of the previous studies performed follow-up tests in patients with adrenal insufficiency. Moreover, HPA axis function has only been studied in adults with lymphoma but not in pediatric lymphoma patients.

In conclusion, suppression of the HPA axis during treatment for ALL and lymphoma has been reported, but its degree and duration are not well defined. In addition, the risk factors for prolonged HPA axis suppression remain unclear. Therefore, a consistent picture of HPA axis impairment after glucocorticoid therapy during treatment for childhood ALL and lymphoma has not yet been derived.

**Management of HPA axis suppression**

The risk of life-threatening situations as hypoglycemia, hypotension and infection in patients with HPA axis suppression can easily and successfully be reduced by glucocorticoid coverage (e.g. hydrocortisone). Patients with partial adrenal insufficiency may have normal basal cortisol levels but an inadequate response to stress. These patients need supplemental glucocorticoid coverage during periods of stress. Patients with severe adrenal insufficiency may have low basal serum cortisol levels and need maintenance glucocorticoid therapy simulating the normal circadian cortisol production. The current national and international guidelines recommend glucocorticoid coverage during periods of stress within the first 12 months after cessation of substantial glucocorticoid therapy (>15 mg/m\textsuperscript{2}/day hydrocortisone equivalent ≥ 14 days) in all patients.\textsuperscript{23, 35, 36} However, due to the lack of knowledge on the extent of HPA axis suppression during treatment of childhood ALL and lymphoma, stress dose therapy is no standard practice and there is no agreement on the optimal management of adrenal insufficiency.

**HPA axis function and psychosocial health**

Survivors of childhood cancer are at increased risk for psychosocial morbidity. There is a growing number of studies that report on sleep disturbances, fatigue and impaired quality of life in childhood ALL survivors.\textsuperscript{37-40} In addition, there is increasing evidence that sleep, fatigue, depression and quality of life are associated with HPA axis function.\textsuperscript{41-45} Moreover, glucocorticoid therapy during treatment for childhood ALL has been associated with sleep disturbances, fatigue and decreased quality of life.\textsuperscript{46-50} However, the associations between
sleep, fatigue, depression and quality of life on the one hand and HPA axis function on the other hand, have not been assessed in childhood ALL survivors before.

**Neuroendocrine sequelae later in life**

Prolonged suppression of the HPA axis has been reported after courses of high-dose glucocorticoids during treatment for ALL.\textsuperscript{29, 51, 52} However, the duration has not been established accurately. Animal studies on long-term effects of exposure to dexamethasone during early life have reported a reduction of HPA axis activity later in life.\textsuperscript{53} In addition, neonatal treatment with dexamethasone for chronic lung disease of prematurity has been associated with a blunted HPA axis activity in children at school age.\textsuperscript{54} Since treatment with high-dose glucocorticoids early in life may induce long-lasting side effects, emerging concern has risen about the long-term neuroendocrine sequelae of high-dose glucocorticoid treatment in childhood ALL and lymphoma. Nevertheless, HPA axis activity in survivors of childhood ALL and lymphoma has not been studied.

**Aim and outline of this thesis**

As the survival rates of children with ALL and lymphoma have dramatically increased over time, it is essential to gain more insight in the possible health risks associated with treatment. Glucocorticoids play an important role in the treatment of both types of childhood cancer. Since supraphysiological doses of glucocorticoids may suppress the HPA axis leading to potential life-threatening side-effects\textsuperscript{19, 20, 23} and as treatment with high-dose glucocorticoids early in life may induce long-lasting side effects,\textsuperscript{53, 54} concern has risen about the neuroendocrine sequelae of high-dose glucocorticoid treatment in childhood ALL and lymphoma. Nevertheless, information on this topic is sparse. The aim of this thesis is to assess neuroendocrine and psychosocial consequences of childhood ALL and lymphoma, in order to identify which supportive care measures need to be optimized.

The first part of the thesis, Chapter 2-4, focuses on the HPA axis function during treatment for childhood cancer and the second part of the thesis, Chapter 5-9, focuses on the HPA axis function after finishing treatment for childhood cancer.

**Chapter 2** emphasizes the clinical relevance of the subject of this thesis by reporting four clinical presentations of symptomatic glucocorticoid-induced HPA axis suppression in childhood ALL and lymphoma. The systematic review in **Chapter 3** gives an overview of the available literature regarding HPA axis suppression during treatment for childhood ALL. As concluded in this Chapter, the exact incidence and duration of HPA axis suppression during treatment of childhood ALL remains unclear and further research is needed. Moreover, no information on HPA axis suppression in childhood lymphoma is available. Therefore, we studied the occurrence and duration of adrenal insufficiency at different time points during
and after the treatment of childhood ALL and lymphoma, as described in Chapter 4 and 5, respectively. In Chapter 6 we assessed whether polymorphisms of the glucocorticoid receptor that have been associated with altered sensitivity for glucocorticoids were related to the duration of adrenal insufficiency in childhood ALL.

Considering the frequently reported problems with sleep, fatigue, depression and quality of life in childhood cancer survivors and the increasing evidence for associations between these impairments and cortisol levels, we studied these associations in childhood ALL survivors. Due to the absence of reliable and validated Dutch questionnaires to measure fatigue in children and to measure sleep in adolescents, we first had to develop and validate appropriate questionnaires, as described in Chapter 7 and 8. Consequently, in Chapter 9, we studied sleep, fatigue, depression and quality of life in relation to cortisol levels in survivors of childhood ALL. Although prolonged HPA axis suppression after treatment with high-dose glucocorticoids in childhood ALL has been reported, HPA axis activity after completing treatment for childhood ALL has never been studied before. Therefore, in Chapter 10 we examined HPA axis function in survivors of childhood ALL and in healthy children. This Chapter describes significant differences in cortisol levels between childhood ALL survivors and healthy controls. As increased cortisol levels are associated with increased sympathetic nervous system activity, we studied the relation between HPA axis function and the sympathetic nervous system in childhood ALL survivors and in healthy controls, as described in Chapter 11. Finally, in Chapter 12, the results of this thesis are discussed and in Chapter 13 an overall summary is provided.
References

Chapter 2

Bijnierschorsinsufficiëntie door glucocorticoïdgebruik in de oncologie op de kinderleeftijd

M. Suzanne Gordijn, Reinoud J.B.J. Gemke, Gertjan J.L. Kaspers en Marie J.E. Walenkamp

Tijdschrift voor kindergeneeskunde 2011;79:87-91
Samenvatting

Behandeling met suprafysiologische doses glucocorticoiden is de meest voorkomende oorzaak van secundaire bijnierschorsinsufficiëntie. Herkenning van de aspecifieke symptomen van een bijnierschorsinsufficiëntie kan moeilijk zijn. Wij presenteren de ziektegeschiedenissen van vier kinderen met een maligniteit; twee jongens van 2 en 12 jaar oud en twee meisjes van 5 en 15 jaar oud met verschillende klinische presentaties van een bijnierschorsinsufficiëntie geïnduceerd door glucocorticoiden. Bijnierschorsinsufficiëntie resulterend in inadequaat cortisolwaarden tijdens stressvolle situaties (bijvoorbeeld ziekte, trauma, operatie) kan leiden tot een levensbedreigende hypoglykemie en/of hypotensieve crise. De diagnose kan worden gesteld door de lage-dosis ACTH-test. Stressdoseringen hydrocortison kunnen een levensbedreigende situatie voorkomen. Vanwege het gebrek aan evidence-based richtlijnen is er geen consensus omtrent het optimale beleid bij bijnierschorsinsufficiëntie. Om die reden stellen wij een stroomschema voor de behandeling met hydrocortison voor.

Summary

Treatment with supraphysiologic doses glucocorticoids is the most common cause of secondary adrenal insufficiency. Recognition of the atypical clinical signs of adrenal insufficiency may be difficult. We present four children with a malignancy; two boys aged 2 and 12 years and two girls aged 5 and 15 years, with diverse clinical presentations of glucocorticoid-induced adrenal insufficiency. Adrenal insufficiency resulting in inadequate cortisol levels during periods of stress (e.g. illness, trauma, surgery), may lead to life-threatening hypoglycemia and/or hypotension. The diagnostic work-up includes a low dose ACTH test. Glucocorticoid coverage (hydrocortisone) during periods of stress is recommended in order to prevent life-threatening situations. Due to the lack of evidence-based guidelines, there is no agreement on the optimal treatment of adrenal insufficiency. Therefore, we propose a flowchart for the indication of glucocorticoid coverage.
Inleiding

Vanwege de oncolytische werking worden glucocorticoiden bij verschillende vormen van kanker op de kinderleeftijd ingezet. Na gebruik van suprafysiologische doses kan secundaire bijnierschorsinsufficiëntie optreden en kan bij stress een levensbedreigende situatie ontstaan. Tijdige onderkenning van een bijnierschorsinsufficiëntie is daarom van belang.

Aangezien de symptomen van bijnierschorsinsufficiëntie aspecifiek zijn of in het geheel kunnen ontbreken, kan het lastig zijn dit beeld te herkennen. Aan de hand van de volgende vier ziektegeschiedenissen illustreren wij verschillende manifestaties van bijnierschorsinsufficiëntie in de kinderoncologische populatie.

Casusbeschrijvingen

Patiënt A, een 12-jarige jongen, met een acute lymfatische leukemie (ALL), zagen wij voor een chemotherapiegift op onze polikliniek kinderoncologie. Vijf maanden voor dit bezoek had hij gedurende vier weken prednisolon gehad in een dosering van 60 mg/m²/dag. Tot twee dagen voor het bezoek werd hij behandeld met vijfdaagse kuren dexamethason 6 mg/m²/dag per os. Nu klaagde hij over spierzwakte en duizeligheid. Hij was die ochtend flauwgevallen bij traplopen. Er was geen sprake van koorts. Bij lichamelijk onderzoek zagen wij een timide, niet ziek ogende jongen met een cushingoïd uiterlijk. De bloeddruk was 108/68 mmHg, de pols 90/min. Het overige lichamelijk onderzoek was niet afwijkend. Aanvullend laboratoriumonderzoek toonde, naast bij de behandeling passende bloedwaarden (Hb 8,2 mmol/l, trombocyten 144 x 10⁹/l, leukocyten 2,7 x 10⁹/l), geen bijzonderheden. Ten aanzien van de klachten werd een expectatief beleid gevolgd en de chemotherapiegift werd toegepast. Na een week volgde poliklinische controle bij de kinderarts-oncoloog, waar de tijdens stress afgenomen verlaagde ochtendserumcortisol van 102 nmol/l van de week ervoor werd opgemerkt. Onder verdenking van een bijnierschorsinsufficiëntie werd een substitutiedoosering hydrocortison 12 mg/m²/dag gestart. Bij stress werd een stressschema hydrocortison geadviseerd. Een week later voelde hij zich aanzienlijk beter, fitter en minder vermoeid.

Patiënt B, een 15-jarig meisje met de ziekte van Hodgkin, ontving in het kader van haar behandeling iedere drie weken prednisolon 40 mg/m²/dag per os in één dosis, gedurende 14 dagen. Zij kwam elders voor een chemotherapiegift. Ze was bezig met de achtste dag prednisolon. Patiënte voelde zich ziek, kouwelijk en ze hoestte. Na de chemotherapie en de daaropvolgende erytrocytentransfusie steeg haar lichaamstemperatuur tot 39,2 ºC. De hartfrequentie was 126/min, de ademhalingsfrequentie 23/min en de bloeddruk 113/36 mmHg. Na vulling met NaCl 0,9%, intraveneuze toediening van 100 mg hydrocortison en
breedspectrum antibiotica werd patiënte naar ons ziekenhuis overgeplaatst onder verdenking van een septische shock.

Bij aankomst zagen wij een niet ziek ogend meisje met een normale bloeddruk en lichaamstemperatuur. Aanvullend laboratoriumonderzoek toonde, behoudens bij de behandeling passende afwijkingen, een licht verhoogd CRP (CRP 17 mg/l, Hb 5,4 mmol/l, trombocyten 74 x 10⁹/l, leukocyten 0,2 x 10⁹/l). De bloed- en urinekweken waren negatief. Vier dagen later trad eenzelfde beeld van koorts, hypotensie en tachycardie op, zonder dat er bloedproducten waren toegediend. Hierop werd een vulling NaCl 0,9% en 100 mg hydrocortison i.v. gegeven, waarna goed herstel. De volgende ochtend werd een verlaagd ochtendserumcortisol vastgesteld (43 nmol/l). Hierop werd de dagelijkse dosering prednisolon verdeeld over drie giften en werd na het staken van de prednisolon een substitutiedosering hydrocortison afgesproken. Een lage-dosis ACTH-test een week later toonde een basale cortisolwaarde van 146 nmol/l en een inadequate oploop van het cortisol tot 261 nmol/l. De onverklaarde incidenten in combinatie met het verlaagde ochtendserumcortisol en de inadequate cortisollijging bij de lage-dosis ACTH-test, leken te passen bij een bijnierschorsinsufficiëntie. Een maand later was het ochtendserumcortisol 250 nmol/l, waarna de substitutiedosering werd gestopt. Wel werd een stressdosering hydrocortison gedurende een jaar geadviseerd.

Patiënt C, een 5-jarig meisje met ALL, was ter observatie op onze afdeling kindergeneeskunde opgenomen in verband met algehele malaise. Zij was klinisch niet ziek en het laboratoriumonderzoek toonde, naast de bij de behandeling passende bloedwaarden (CRP <2,5 mg/l, Hb 7,1 mmol/l, trombocyten 97 x 10⁹/l, leukocyten 1,6 x 10⁹/l) geen bijzonderheden. Op de tweede dag van opname kreeg ze last van hoofdpijn, misselijkheid en braken. De lichaamstemperatuur was 36,4 ºC en de bloeddruk 88/51 mmHg. Overig lichamelijk onderzoek was niet afwijkend. Patiënte was sinds 15 dagen gestaakt met een vierweekse kuur prednisolon 60 mg/m²/dag, waarna een negendaags afbouwschema volgde. Vanwege een mogelijke bijnierschorsinsufficiëntie werd 100 mg hydrocortison intraveneus toegediend, waarna een vlot herstel volgde. Twee dagen later werd een verlaagd ochtendserumcortisol vastgesteld (48 nmol/l), waarop een substitutiedosering hydrocortison 12 mg/m²/dag werd voorgeschreven. Een maand later onderging patiëntë in het kader van haar ziekte een beenmergpunctie onder narcose. Na afloop van de narcose kreeg patiëntë last van buikpijn en braken. De lichaamstemperatuur was 36,4 ºC, de pols 77/min en de bloeddruk 85/45 mmHg. Bij navraag bleek patiëntë die ochtend geen hydrocortison ingenomen te hebben. Na intraveneuze toediening van 100 mg hydrocortison knapte patiëntë snel op en de substitutiedosering hydrocortison werd herstart. Na een maand was het ochtendserumcortisol 106 nmol/l. In verband met een goede kliniek werd de substitutiedosering gestaakt en werd een stressschema geadviseerd.

De ouders van patiënt D, een 2-jarig jongetje met ALL, vertelden dat hun zoon sinds een aantal dagen wakker werd met trillende handen en dat dit pas verdween na het ontbijt. Patiënt was...
Bijnierschorsinsufficiëntie in de kinderoncologie

recent behandeld met een vierweekse kuur prednisolon 60 mg/m²/dag, hetgeen sinds vijf dagen werd afgebouwd. Aanvullend laboratoriumonderzoek toonde een normale nuchtere glucose (5,6 mmol/l) en een verlaagd ochtendserumcortisol (36 nmol/l). In verband met verdenking op hypoglykemieën bij een bijnierschorsinsufficiëntie werd een substitutiedosering hydrocortison 15 mg/m²/dag afgesproken. Na een maand werd de hydrocortison gestaakt, waarna ter controle een ochtendserumcortisol werd bepaald. Vanwege de verlaagde waarde (< 30 nmol/l) werd besloten de substitutiedosering te herstarten. Na een maand werd de substitutiedosering opnieuw gestaakt, waarna het ochtendserumcortisol 132 nmol/l was. Gezien de goede kliniek werd de substitutiedosering gestaakt en werd een stressschema geadviseerd.

**Discussie**

Bovenstaande casus illustreren verschillende presentaties van bijnierschorsinsufficiëntie in de kinderoncologische praktijk. De ziektegeschiedenis van patiënt A laat zien dat er, ondanks substantieel glucocorticoidgebruik in de anamnese, niet altijd aan een bijnierschorsinsufficiëntie wordt gedacht. De bijnierschorsinsufficiëntie van deze patiënt kan een effect zijn geweest van de eerdere langdurige glucocorticoidkuur, maar ook van de recente kortdurende glucocorticoidkuur.\(^1,^2\) Glucocorticoidspiegels dienen adequaat te zijn voor het moment van de dag en de mate van stress, zoals wordt geïllustreerd door patiënt B. Patiënt C liet vlot herstel na intraveneuze toediening van hydrocortison zien bij onverklaarde klachten die zeer goed kunnen passen bij een bijnierschorsinsufficiëntie. Hypoglykemieën kunnen wijzen op een bijnierschorsinsufficiëntie, zoals bij patiënt D. Door middel van deze casus willen wij het belang van tijdige onderkenning van iatrogene bijnierschorsinsufficiëntie benadrukken.

**Pathogenese**

De bijnierschors is onder andere verantwoordelijk voor de productie van glucocorticoiden. Dit proces wordt aangestuurd vanuit de hypothalamus, die door middel van het corticotrofine-releasing hormoon (CRH) de hypofyse aanzet tot het produceren van ACTH, hetgeen resulteert in cortisolproductie door de bijnierschors.

Primaire bijnierinsufficiëntie wordt gekenmerkt door verhoogde ACTH-spiegels. Naast deficiëntie van glucocorticoiden is uitval van mineralocorticoiden mogelijk. Secundaire bijnierinsufficiëntie wordt gekenmerkt door verlaagde ACTH-spiegels en geïsoleerde uitval van glucocorticoiden. De meest voorkomende oorzaak van secundaire bijnierinsufficiëntie is behandeling met glucocorticoiden. Door supрафysiologische concentraties wordt de secretie van CRH en ACTH geremd, waardoor uiteindelijk bijnierschorsatrofie ontstaat. Het vermogen om cortisol te produceren neemt hierbij af.\(^3,^4\) De duur en ernst hiervan kennen een inter-individuele variabiliteit.
Tijdens stressvolle situaties zoals ziekte, infectie, trauma, operatie en ernstig psychisch onwelbevinden is de natuurlijke behoefte aan cortisol verhoogd. Indien niet aan deze verhoogde cortisolbehoefte kan worden voldaan, zoals bij bijnierschorsinsufficiëntie, kan een levensbedreigende hypoglykemie en/of hypotensieve crise ontstaan.  

**Klinische verschijnselen en risicofactoren**

Soms zijn de symptomen duidelijk, zoals hypoglykemie en circulatoire insufficiëntie, maar vaak zijn de verschijnselen aspecifiek (vermoeidheid, anorexia, misselijkheid, gewichtsverlies, spier- of gewrichtspijn, duizeligheid) of in het geheel afwezig.

Indien de anamnese substantieel glucocorticoidgebruik gedurende het afgelopen jaar vermeldt, moet er aan de mogelijkheid van een bijnierschorsinsufficiëntie worden gedacht. Bovendien dient te worden overwogen of de glucocorticoidspiegel op dat moment adequaat is met betrekking tot het tijdstip op de dag en de mate van stress.

**Diagnostiek**

Ochtendwaarden serumcortisol reflecteren de basale bijnierschorsfunctie, maar geven geen informatie over de capaciteit van de bijnierschors om op stress te reageren. Stimulatietests worden gebruikt om de bijnierschorsrespons te meten. De insulinetolerantietest wordt als de gouden standaard beschouwd, maar is geassocieerd met gevaarlijke bijwerkingen, zoals ernstige hypoglykemie en coma.

De meest gebruikte test om bijnierschorsinsufficiëntie te testen is de lage-dosis (1 μg) ACTH-stimulatietest. Hierbij wordt er intraveneus of intramuskulair tetracosactide (Synacthen 1 μg per 1.73 m²), de eerste 24 aminozuren van humaan ACTH, toegediend. Dit stimuleert de biosynthese van glucocorticoiden. Een normale respons is gedefinieerd als een cortisolstijging ten opzichte van de baseline van >200 nmol/l of één gestimuleerd serumcortisol van ≥ 550 nmol/l na 30 minuten. Met deze test kan subtiele bijnierschorsatrofie worden gedetecteerd.

**Therapie**

Er bestaat geen consensus ten aanzien van het optimale beleid bij secundaire bijnierschorsinsufficiëntie aangezien evidence-based richtlijnen ontbreken. Om enige handvatten te bieden, stellen wij de volgende behandeling (Figuur 2.1) voor: gezien het risico op onvoldoende cortisolaanmaak tijdens stressvolle situaties, ontvangen alle patiënten die behandel worden met glucocorticoiden of met substantieel glucocorticoidgebruik (> 15 mg/m² dag hydrocortisonequivalent) in het afgelopen jaar, ten tijde van stress een stressschema hydrocortison 30-50 mg/m² dag (dosering afhankelijk van de mate van stress) gedurende 12 maanden na het staken van de glucocorticoidbehandeling. Deze indicatie komt te vervallen wanneer een serumcortisol > 550 nmol/l wordt gemeten (basaal, tijdens stress of tijdens de lage-dosis ACTH-test). Om te voorkomen dat patiënten onnodig (lang) een
Bijnierschorsinsufficiëntie in de kinderoncologie

stressschema hydrocortison ontvangen, adviseren wij om zes maanden na het staken van glucocorticoïdtherapie een lage-dosis ACTH-test uit te voeren.¹,⁴

Patiënten die onder basale omstandigheden onvoldoende cortisol aanmaken (ochtendserumcortisol < 80 nmol/l³,⁹) en bij wie een klinische verdenking bestaat op bijnierschorsinsufficiëntie, ontvangen een substitutiedosering hydrocortison 8-12 mg/m²/dag per os verspreid over drie doses (‘s ochtends de helft en de volgende twee giften een kwart van de dagdosis, in navolging van het fysiologisch cortisoldagritme), hetgeen na twee weken kan worden afgebouwd.⁵,¹²,¹³ In geval van een suboptimaal serumochtendcortisol van 80-270 nmol/l adviseer we de bepaling elke drie weken te herhalen tot een basale cortisolwaarde van > 270 nmol/l.³,⁵ Behandeling van patiënten met bijnierschorsinsufficiëntie tijdens glucocorticoidtherapie bestaat uit verdeling van de glucocorticoidmedicatie over meerdere tijdstippen per dag, waarbij tijdens stress een adequate hoeveelheid glucocorticoiden gegeven dient te worden.

Van groot belang is dat kinderen en hun ouders goede mondelinge en schriftelijke instructies ontvangen over de indicatie en de noodzaak van stressdoseringen.⁵,⁹

---

**Figuur 2.1** Stroomschema behandeling secundaire bijnierschorsinsufficiëntie.
Conclusie

Door middel van deze casus laten wij zien dat de manifestaties van secundaire bijnierschorsinsufficiëntie zeer verschillend kunnen zijn en dat de herkenning ervan in de klinische praktijk moeilijk is. Tijdige onderkenning is van groot belang, gezien de potentieel levensbedreigende gevolgen. Voor het diagnosticeren van een bijnierschorsinsufficiëntie kan de lagedosis ACTH-test worden gebruikt. Aangezien er geen consensus bestaat omtrent de optimale behandeling van secundaire bijnierschorsinsufficiëntie, presenteren wij een stroomschema voor de behandeling met hydrocortison.
Referenties

Chapter 3

Hypothalamic-pituitary-adrenal (HPA) axis suppression after treatment with glucocorticoid therapy for childhood acute lymphoblastic leukaemia (Review)

M. Suzanne Gordijn, Reinoud J.B.J. Gemke, Elvira C. van Dalen, Joost Rotteveel, Gertjan J.L. Kaspers

Cochrane Database Syst Rev 2012;5:CD008727
Abstract

Background
Glucocorticoids play a major role in the treatment of acute lymphoblastic leukaemia (ALL). However, supraphysiological doses may cause suppression of the hypothalamic-pituitary-adrenal (HPA) axis. HPA axis suppression resulting in reduced cortisol response may cause an impaired stress response and an inadequate host defence against infections, which remains a cause of morbidity and death. The exact occurrence and duration of HPA axis suppression after glucocorticoid therapy for childhood ALL are unclear.

Objectives
To examine the occurrence and duration of HPA axis suppression after (each cycle of) glucocorticoid therapy for childhood ALL.

Search methods
We searched the Cochrane Central Register of Controlled Trials (in The Cochrane Library, issue 3, 2010), MEDLINE/PubMed (from 1945 to July 2010) and EMBASE/Ovid (from 1980 to July 2010). In addition, we searched reference lists of relevant articles, conference proceedings and ongoing trial databases.

Selection criteria
All study designs, except case reports and patient series with fewer than 10 patients, examining the effect of glucocorticoid therapy for childhood ALL on the HPA axis function.

Data collection and analysis
Two review authors independently performed the study selection. One review author performed the data extraction and ‘Risk of bias’ assessment, which was checked by another review author.

Main results
We identified seven studies (total number of participants = 189), including one randomised controlled trial (RCT), which assessed the adrenal function. None of the studies assessed the HPA axis at the level of the hypothalamus, pituitary, or both. Due to substantial differences between studies, results could not be pooled. All studies had some methodological limitations. The included studies demonstrated that adrenal insufficiency occurs in nearly all patients in the first days after cessation of glucocorticoid treatment for childhood ALL. The majority of patients recovered within a few weeks, but a small amount of patients had ongoing adrenal insufficiency lasting up to 34 weeks. In the RCT, the occurrence and duration of adrenal insufficiency did not differ between the prednisolone and dexamethasone arms. In
one study included in the review it appeared that treatment with fluconazole prolonged the duration of adrenal insufficiency.

**Authors’ conclusions**

Based on the available evidence, we conclude that adrenal insufficiency commonly occurs in the first days after cessation of glucocorticoid therapy for childhood ALL, but the exact duration is unclear. Since no data on the level of the hypothalamus and the pituitary were available we cannot make any conclusions regarding those outcomes. Clinicians should consider prescribing glucocorticoid replacement therapy during periods of serious stress in the first weeks after cessation of glucocorticoid therapy for childhood ALL, to reduce the risk of life-threatening complications. However, more high-quality research is needed for evidence-based guidelines for glucocorticoid replacement therapy.

Special attention should be paid to patients receiving fluconazole therapy, and perhaps similar antifungal drugs, as this may prolong the duration of adrenal insufficiency.
Chapter 3

Background

Of all malignancies in children, acute lymphoblastic leukaemia (ALL) is the most frequent type. Annually, approximately 120 children with ALL are newly diagnosed in the Netherlands. Since treatment and survival rates of childhood ALL have substantially improved over time, morbidity and mortality due to treatment-related side effects become increasingly important. Unfortunately, up to 5% of children die because of the toxic side effects of treatment, and this percentage becomes even greater in higher-risk subgroups. The main cause of this treatment-related mortality is infection associated with cytotoxic and immunosuppressive drugs. Glucocorticoid therapy is an important contributing factor to the occurrence and severity of infection. In addition, several studies report an increase in sepsis and lethal infections in children with ALL since prednisone has been substituted for the more potent glucocorticoid dexamethasone.

Glucocorticoids play a major role in the treatment of ALL. They induce apoptosis of the lymphoblastic cells. Children with ALL receive cyclic courses of high-dose glucocorticoids, like prednisone (or prednisolone) and dexamethasone. However, supraphysiological doses of glucocorticoids may suppress hypothalamic secretion of corticotrophin-releasing hormone (CRH) and of adrenocorticotropic hormone (ACTH) by the pituitary gland, resulting in secondary adrenal cortex atrophy with delayed recovery of hypothalamic-pituitary-adrenal (HPA) axis function. In states of profound or prolonged ACTH deficiency, the adrenal glands may temporarily be unable to generate sufficient cortisol. Supraphysiological glucocorticoid therapy is the most common cause of secondary adrenal insufficiency. The HPA axis plays a major role in the stress response and host defence against infection. Stressors such as trauma, surgery or inflammation stimulate the HPA axis, leading to an increase in cortisol production. In turn, increased cortisol levels cause anti-inflammatory effects and inhibition of pro-inflammatory cytokines. Suppression of the HPA axis resulting in reduced adrenal cortisol production represents an impaired stress response and an inadequate host defence against infections and remains a cause of morbidity and death in childhood.

There are different well-established tests of HPA axis function. Morning serum cortisol value reflects basal adrenal function but gives no indication of the capacity to respond to stress. Stimulation tests are used to assess the HPA axis response to stress. The insulin-tolerance-test is considered the most reliable test to evaluate the HPA axis function at the level of the pituitary and the adrenal gland, but is associated with potentially dangerous side effects. The CRH stimulation test is only indicated in patients with a central disorder of the HPA axis. The glucagon test has also been proven to be a safe and reliable method for testing HPA axis function at the level of the pituitary and the adrenal gland, but may also induce mild inadvertent side effects. A well-established alternative to test the HPA axis at the level of the adrenal gland without undesirable side effects is the low-dose (1 µg) ACTH stimulation test. This test is able to detect more subtle degrees of adrenal atrophy caused by central
adrenal insufficiency compared to the “normal” ACTH stimulation test (250 µg). The results of the low-dose ACTH test closely correlate with those of the insulin-tolerance-test.\textsuperscript{24-26}

Several studies have prospectively assessed HPA axis function following high-dose glucocorticoids during treatment for ALL. Adrenal stimulation tests have been performed and repeated until cortisol levels normalised. The majority of patients seem to recover in a few weeks but prolonged suppression may occur, lasting over several months in some cases.\textsuperscript{27-29} Since children with ALL plus HPA axis suppression suffering from fever or other stressors may benefit from glucocorticoid replacement therapy (e.g. hydrocortisone), it is important to derive a consistent picture of HPA axis impairment after corticosteroid therapy.

Our aim was to undertake a systematic review of the HPA axis function after glucocorticoid therapy for childhood ALL. On this basis, adequate guidelines for glucocorticoid substitution can be formulated and implemented to reduce the risk of infection in childhood ALL.

**Objectives**

*Primary objective*
- To examine the occurrence of HPA axis suppression after (each cycle of) glucocorticoid therapy for childhood ALL.
- To examine the duration of HPA axis suppression after treatment with glucocorticoids for childhood ALL.

*Secondary objectives*

To examine whether the HPA axis function after administration of glucocorticoids is dependent on:
- the moment of testing after cessation of glucocorticoid therapy
- the (cumulative) dose of glucocorticoids
- the type of glucocorticoids; prednisone, prednisolone or dexamethasone
- the duration of glucocorticoid therapy
- the method of cessation of glucocorticoid therapy; abrupt or gradual

**Methods**

**Criteria for considering studies for this review**

*Types of studies*

All study designs, except case reports and patient series with fewer than 10 patients, examining the effect of glucocorticoid therapy for childhood ALL on HPA axis function. This effect can be evaluated both during treatment for ALL (after cessation of a glucocorticoid course)
and after the end of all ALL treatment. Discrepancies concerning the definition of a cohort study between review authors were resolved by consensus. No third party arbitration was needed.

**Types of participants**

Patients who were treated with glucocorticoids for ALL between the age of 0 and 18 years, irrespective of the duration of follow-up after the end of glucocorticoid therapy. Exclusion criteria are cranial radiotherapy, since this may damage the HPA axis, and testing HPA axis function by a CRH stimulation test only, since this test is only indicated in patients with a central disorder of the HPA axis.¹⁹,²⁰

**Types of interventions**

Glucocorticoid therapy (prednisone, prednisolone, dexamethasone) during treatment for ALL. The intervention was not compared to a control group, since this was not available (with the exception of the included randomised controlled trials (RCT)).

**Types of outcome measures**

HPA axis function, measured by early morning plasma cortisol levels (between 8 and 10 a.m.) or by stimulation tests (e.g. (low-dose) ACTH stimulation test or glucagon stimulation test). We used the cut-off limit as defined by the authors of the original studies.

**Search methods for identification of studies**

See: Cochrane Childhood Cancer Group methods used in reviews.³⁰ The objective of the literature search was to identify all studies, except case reports and case series, reporting on HPA axis function after glucocorticoid therapy for childhood ALL.

**Electronic searches**

We searched the following electronic databases: the Cochrane Central Register of Controlled Trials (CENTRAL) (in The Cochrane Library, issue 3, 2010), MEDLINE/PubMed (from 1945 to July 2010) and EMBASE/Ovid (from 1980 to July 2010). The search strategies for the different electronic databases (using a combination of controlled vocabulary and text words) are shown in the appendices (see Appendix 1, Appendix 2, Appendix 3).

**Searching other resources**

We located information about trials not registered in CENTRAL, MEDLINE or EMBASE, either published or unpublished, by searching the reference lists of relevant articles and review articles. We handsearched the conference proceedings of the International Society for Paediatric Oncology (SIOP) (from 2005 to 2009) and the American Society of Clinical Oncology (ASCO) (from 2005 to 2009). For the search strategy see Appendix 4.
We also scanned the International Standard Randomized Controlled Trial Number (ISRCTN) register and the National Institutes of Health (NIH) register for ongoing trials: http://www.controlled-trials.com (both screened July 2010; for the search strategy see Appendix 5). We imposed no language restrictions. We will update the searches every two years.

Data collection and analysis

Selection of studies
After performing the search strategy, two review authors independently selected studies meeting the inclusion criteria. Discrepancies between review authors were resolved by consensus. No third party arbitration was needed. We obtained in full any study that seemed to meet the inclusion criteria on grounds of the title or abstract or both, for closer inspection. We clearly stated reasons for exclusion of any study considered for this review.

Data extraction and management
Data extraction was performed by one review author using standardised forms and checked by another review author. The review authors were not blinded to the journal, the authors or the institution. We extracted data on the following categories: study characteristics, participants, interventions, outcome measures, length of follow-up, risk factors, and ‘Risk of bias’ assessment. Discrepancies between authors were resolved by consensus. No third party arbitration was needed.

Assessment of risk of bias in included studies
The assessment of risk of bias was based on previously described checklists for observational studies according to Evidence-Based Medicine Criteria.31,32 The assessment of risk of bias of the included studies was performed by one review author and checked by another review author. The ‘Risk of bias’ assessment criteria for observational studies are described in Table 3.1. For RCTs we used the ‘Risk of bias’ items as described in the module of the Childhood Cancer Group, which are based on the Cochrane Handbook for Systematic Reviews of Interventions (see Table 3.2).30,33 Discrepancies between review authors were resolved by consensus. No third party arbitration was needed.

Measures of treatment effect
Prevalence of HPA axis suppression at several follow-up time points and time duration of HPA axis suppression.

Dealing with missing data
We contacted authors of individual studies for clarification of unclear data or to obtain missing data regarding selection of studies, the ‘Risk of bias’ assessment and data extraction.
### Table 3.1 Risk of bias assessment criteria for observational studies

<table>
<thead>
<tr>
<th></th>
<th>Internal validity</th>
<th>External validity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study group</strong></td>
<td>Selection bias (representative: yes/no):</td>
<td>Reporting bias (well defined: yes/no):</td>
</tr>
<tr>
<td></td>
<td>• if it consisted of more than 90% of the original cohort of ALL patients treated with glucocorticoids</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• or if it was a random sample with respect to the treatment</td>
<td>• if the treatment protocol was mentioned</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• and if the (cumulative) dose of glucocorticoid treatment was mentioned</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• and if the type of glucocorticoid treatment was mentioned</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• and if the duration of glucocorticoid treatment was mentioned</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• and if the method of cessation of glucocorticoid treatment was mentioned</td>
</tr>
<tr>
<td><strong>Follow-up</strong></td>
<td>Attrition bias (adequate: yes/no):</td>
<td>Reporting bias (well-defined: yes/no):</td>
</tr>
<tr>
<td></td>
<td>• if the outcome was assessed at the end date of the study for 60% to 90% of the study group</td>
<td>• if the length of follow-up was mentioned</td>
</tr>
<tr>
<td></td>
<td>• or if the outcome was assessed for more than 90% of the study group but with an unknown end date</td>
<td>• and if the frequency of measuring the outcomes was mentioned</td>
</tr>
<tr>
<td><strong>Outcome</strong></td>
<td>Detection bias (blinding: yes/no):</td>
<td>Reporting bias (well-defined: yes/no):</td>
</tr>
<tr>
<td></td>
<td>• if the outcome assessor was blinded to glucocorticoid treatment</td>
<td>• if the methods of detection were described</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• if the outcome definition was objective and precise</td>
</tr>
<tr>
<td><strong>Risk estimation</strong></td>
<td>Confounding (adjustment for other factors: yes/no):</td>
<td>Analysis (well-defined: yes/no):</td>
</tr>
<tr>
<td></td>
<td>• if important prognostic factors (i.e. age, sex, co-treatment) or follow-up were taken adequately into account</td>
<td>• if a relative risk, odds ratio, attributable risk, linear or logistic regression model, mean difference, or Chi² statistic was calculated</td>
</tr>
</tbody>
</table>

### Table 3.2 Risk of bias assessment criteria for randomised controlled trials

**Selection bias**
- Sequence generation (adequate: yes/no):
  - if the rule for allocating interventions to participants was based on some chance (random) process
- Allocation concealment (adequate: yes/no):
  - if the randomisation method did not allow investigator and participant to know or influence the allocation of treatment before eligible participants entered the study

**Performance bias**
- Blinding of care providers (yes/no):
  - if knowledge of the allocated intervention was adequately prevented during the study
- Blinding of participants (yes/no):
  - if knowledge of the allocated intervention was adequately prevented during the study

**Detection bias**
- Blinding of outcome assessors (yes/no; assessed for each outcome separately):
  - if knowledge of the allocated intervention was adequately prevented during the study

**Attrition bias**
- Incomplete outcome data (adequate: yes/no; assessed for each outcome separately):
  - if incomplete outcome (attrition and exclusions) data have been adequately addressed

**Reporting bias**
- Selective outcome reporting (yes/no):
  - if the reports of the study were free of suggestion of selective outcome reporting

**Other bias**
- Other bias (yes/no):
  - if the study was free of other problems (i.e. potential source of bias related to the specific study design, premature termination of the study due to some data-dependent process, extreme baseline imbalance) that could put it at a high risk of bias
Assessment of heterogeneity

We planned to assess heterogeneity both by visual inspection of the forest plots and by a formal statistical test for heterogeneity, that is the $I^2$ statistic. However, since we were not able to pool the results of the included studies, this was not applicable.

Assessment of reporting biases

We planned to use a funnel plot to quantify the potential presence of publication bias. However, since we were not able to pool the results of the included studies, this was not applicable.

Data synthesis

We planned to perform analyses using the statistical software Comprehensive Meta Analysis.34 Across the various studies, we planned to conduct a multivariate linear meta-regression analysis model with a backwards selection strategy ($P < 0.10$) to examine the relation between potential predictive factors and HPA axis suppression.

However, since we were not able to pool the results of the included studies, this was not applicable and we described the results of the individual studies separately. Kaplan-Meier curves were calculated using SPSS version 15.0.

Sensitivity analysis

We planned to perform a sensitivity analysis for the used ‘Risk of bias’ assessment criteria. However, since we were not able to pool the results of the included studies, this was not applicable.

Results

Description of studies

Results of the search

Running the searches in the electronic databases of CENTRAL, MEDLINE/PubMed and EMBASE/Ovid revealed a total of 1388 references. Initial screening of the titles and abstracts excluded 1375 references that clearly did not fulfil all criteria for considering studies for this review. Of the remaining 13 references we examined the full-text articles. Seven of the 13 full-text articles were eligible for inclusion in this systematic review, whereas the other six were not. The reasons for exclusion are illustrated in Characteristics of excluded studies (Supplemental Table 3.2).

Scanning the reference list of relevant articles and reviews did not identify any additional eligible studies.
Scanning the conference proceedings of SIOP and ASCO did not identify any eligible studies and scanning the ongoing trials databases did not identify any ongoing studies.

In summary, seven articles could be included in this review. We made attempts to contact authors to clarify aspects of the study design and data analysis. Characteristics of the included studies are summarised in the Characteristics of included studies table (Supplemental Table 3.1).

**Included studies**

All included studies evaluated adrenal function after glucocorticoid therapy for childhood ALL. None of the studies assessed the HPA axis at the level of the hypothalamus, pituitary or both. The total number of patients included in the studies was 189. The seven included studies examined the adrenal function after different types, doses and durations of glucocorticoid therapy and after different methods of cessation of glucocorticoid therapy. Three studies examined the effect of dexamethasone on adrenal function and four studies evaluated the effect of both dexamethasone and prednisolone on adrenal function. Adrenal function was measured by early morning plasma cortisol levels (between 8 and 10 a.m.) or by the (low-dose) ACTH stimulation test. Four studies performed follow-up tests until normalisation of adrenal function, for the other three studies the follow-up lengths were 1 month, 2 weeks and 2 to 7 days, respectively. Six of the seven identified studies were observational studies and one was an RCT evaluating prednisolone versus dexamethasone (both treatment groups received prednisolone prior to randomisation).

**Excluded studies**

Information on the six studies excluded during examination of the full-text articles is included in the Characteristics of excluded studies table (Supplemental Table 3.2). The most common reasons for exclusion were cranial irradiation therapy or the lack of (adequate) HPA axis function tests.

**Risk of bias in included studies**

**Cohort studies**

For the evaluation of the ‘internal validity’ of the six included cohort studies we assessed the risk of selection bias, attrition bias, detection bias and confounding that was present in the included studies. Based on additional information provided by the authors, there was a low risk of selection bias (based on the representativeness of the study group) in four of the six studies as the study group consisted of more than 90% of the original cohort. According to additional information provided by the authors, one study selected an unrepresentative study group of about 30% of the original cohort. For another study neither the
article nor the correspondence with the author provided information on the selection of the participants. The risk of attrition bias (based on the completeness of follow-up) was low in all six studies as the outcome was assessed for 60% to 90% of the study group at the end date of the study. Additional information provided by the authors of one study reported that the outcome assessor was not blinded. None of the other studies provided information on the blinding of the outcome assessor to glucocorticoid treatment, so detection bias could not be ruled out. In five studies, confounding (based on important risk factors and follow-up that were taken into account) might have been present. In one study, fluconazole therapy as a risk factor for the development of adrenal insufficiency was taken into account.

For the evaluation of the ‘external validity’ of the included cohort studies we assessed the risk of reporting bias. In four studies, the study group was well defined in terms of treatment protocol and (cumulative) dose, type, duration and form of cessation of glucocorticoid treatment. Two studies did not mention the treatment protocol. In all six studies the follow-up was well defined as the length of follow-up and the frequency of measuring was mentioned. In these studies the outcome was well defined and the methods of detection were described and the outcome definition was objective and precise. The risk estimation (based on the calculation of relative risk, odds ratio, attributable risk, linear or logistic regression model, mean difference, or Chi² statistic) was well defined in five of the six studies. See also Table 3.3.

**Randomised controlled trial**

For the included RCT we evaluated the risk of selection bias, performance bias, detection bias, attrition bias, reporting bias and other biases. Based on additional information provided by the authors, no selection bias (based on sequence generation and concealment of allocation) was present. Performance bias (based on the blinding of care providers and participants) and detection bias (based on the blinding of the outcome assessors) could not be ruled out. The RCT was susceptible for reporting bias as not all of the study’s pre-specified primary outcomes were reported. The risk on attrition bias (based on the completeness of outcome data) or other bias (i.e. based on potential source of bias related to the specific study design, premature termination of the study due to some data-dependent process, extreme baseline imbalance) in the RCT was low. See also Table 3.4.
<table>
<thead>
<tr>
<th>Study</th>
<th>Representative study group</th>
<th>Complete follow-up assessment</th>
<th>Blinded outcome assessor</th>
<th>Adjustment important confounders</th>
<th>Well-defined study group</th>
<th>Well-defined follow-up</th>
<th>Well-defined outcome</th>
<th>Well-defined risk estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cunha 2004</td>
<td>No, based on additional information provided by the authors, the described study group did not consist of more than 90% of the original cohort and was no random sample</td>
<td>Yes, outcome was assessed for 60% to 90% of the study group at the end date of the study</td>
<td>Unclear if outcome assessor was blinded to the glucocorticoid treatment</td>
<td>No, important prognostic factors or follow-up were not taken into account</td>
<td>Yes, treatment protocol and (cumulative) dose, type, duration and form of cessation of glucocorticoid treatment were mentioned</td>
<td>Yes, length of follow-up and frequency of measuring were mentioned</td>
<td>Yes, methods of detection were described and outcome definition was objective and precise</td>
<td>Yes, mean difference was calculated</td>
</tr>
<tr>
<td>Felner 2000</td>
<td>Yes, based on additional information provided by the authors, the described study group consisted of more than 90% of the original cohort</td>
<td>Yes, outcome was assessed for 60% to 90% of the study group at the end date of the study</td>
<td>Unclear if outcome assessor was blinded to the glucocorticoid treatment</td>
<td>No, important prognostic factors or follow-up were not taken into account</td>
<td>No, treatment protocol was not mentioned</td>
<td>Yes, length of follow-up and frequency of measuring were mentioned</td>
<td>Yes, methods of detection were described and outcome definition was objective and precise</td>
<td>Yes, mean difference was calculated</td>
</tr>
<tr>
<td>Kuperman 2001</td>
<td>Yes, based on additional information provided by the authors, the described study group consisted of more than 90% of the original cohort</td>
<td>Yes, outcome was assessed for 60% to 90% of the study group at the end date of the study</td>
<td>Unclear if outcome assessor was blinded to the glucocorticoid treatment</td>
<td>No, important prognostic factors or follow-up were not taken into account</td>
<td>No, treatment protocol was not mentioned</td>
<td>Yes, length of follow-up and frequency of measuring were mentioned</td>
<td>Yes, methods of detection were described and outcome definition was objective and precise</td>
<td>Yes, mean difference was calculated</td>
</tr>
<tr>
<td>Study</td>
<td>Representative study group</td>
<td>Complete follow-up assessment</td>
<td>Blinded outcome assessor</td>
<td>Adjustment important confounders</td>
<td>Well-defined study group</td>
<td>Well-defined follow-up</td>
<td>Well-defined outcome</td>
<td>Well-defined risk estimation</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>-------------------------------</td>
<td>--------------------------</td>
<td>------------------------------------------------</td>
<td>--------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Mahachoklertwattana 2004</td>
<td>Unclear if the study group consisted of more than 90% of the original cohort or if it was a random sample</td>
<td>Yes, outcome was assessed for 60% to 90% of the study group at the end date of the study</td>
<td>Unclear if outcome assessor was blinded to the glucocorticoid treatment</td>
<td>No, important prognostic factors or follow-up were not taken into account</td>
<td>Yes, treatment protocol and (cumulative) dose, type, duration and form of cessation of glucocorticoid treatment were mentioned</td>
<td>Yes, length of follow-up and frequency of measuring were mentioned</td>
<td>Yes, methods of detection were described and outcome definition was objective and precise</td>
<td>Yes, mean difference was calculated</td>
</tr>
<tr>
<td>Petersen 2003</td>
<td>Yes, the described study group consisted of more than 90% of the original cohort</td>
<td>Yes, outcome was assessed for 60% to 90% of the study group at the end date of the study</td>
<td>Unclear if outcome assessor was blinded to the glucocorticoid treatment</td>
<td>Yes, important prognostic factors or follow-up were taken into account</td>
<td>Yes, treatment protocol and (cumulative) dose, type and duration of glucocorticoid treatment were mentioned. Information on the form of cessation of glucocorticoid treatment was based on additional information provided by the authors</td>
<td>Yes, length of follow-up and frequency of measuring were mentioned</td>
<td>Yes, methods of detection were described and outcome definition was objective and precise</td>
<td>No, relative risk, odds ratio, attributable risk, linear or logistic regression model, mean difference, or $\text{Ch}^2$ statistic were not calculated</td>
</tr>
</tbody>
</table>
### Table 3.3 Risk of bias in included observational studies (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Representative study group</th>
<th>Complete follow-up assessment</th>
<th>Blinded outcome assessor</th>
<th>Adjustment important confounders</th>
<th>Well-defined study group</th>
<th>Well-defined follow-up</th>
<th>Well-defined outcome</th>
<th>Well-defined risk estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rix 2005</td>
<td>Yes, the described study group consisted of more than 90% of the original cohort</td>
<td>Yes, outcome was assessed for 60% to 90% of the study group at the end date of the study</td>
<td>No, the outcome assessor was not blinded to the glucocorticoid treatment</td>
<td>Yes, important prognostic factors or follow-up were taken into account</td>
<td>Yes, treatment protocol and (cumulative) dose, type and duration of glucocorticoid treatment were mentioned. Information on the form of cessation of glucocorticoid treatment was based on additional information provided by the authors</td>
<td>Yes, length of follow-up and frequency of measuring were mentioned</td>
<td>Yes, methods of detection were described and outcome definition was objective and precise</td>
<td>Yes, mean difference was calculated</td>
</tr>
</tbody>
</table>
**Effects of interventions**

**Adrenal insufficiency (occurrence and duration)**

Data on the prevalence and the duration of adrenal insufficiency after treatment with glucocorticoid therapy for childhood ALL could be extracted from all included studies. However, it should be noted that the individual studies used different types and (cumulative) doses of glucocorticoids. In addition, differences in duration and method of cessation of glucocorticoid therapy occurred. Methods of testing adrenal function varied as well. Due to this heterogeneity pooling of results was not possible. For more information see Characteristics of included studies (Supplemental Table 3.1).

Two studies used the ACTH stimulation test with comparable cut-off limits (stimulated cortisol: 18 µg/dL (500 nmol/L)) for measuring adrenal function.\(^{36,38}\) Before glucocorticoid therapy, adrenal function was normal in all patients in the study of Felner et al.\(^{36}\) However, all 10 patients (100%) had insufficient cortisol levels 1 day after abrupt cessation of 28 days of dexamethasone at 6 mg/m\(^2\)/day. Three out of 10 patients (30%) had ongoing adrenal insufficiency after 4 weeks but all recovered after 8 weeks. In the study of Petersen et al., two types of glucocorticoid therapy were assessed; induction therapy comprising 35 days of prednisolone 60 mg/m\(^2\)/day with tapering over 9 days and re-induction therapy comprising 21 days of dexamethasone 10 mg/m\(^2\)/day with tapering over 9 days.\(^{38}\) Re-induction therapy...
succeeded induction therapy, thus patients in the dexamethasone group also received induction therapy with prednisolone. HPA axis function was not assessed before glucocorticoid therapy. After induction therapy (n = 10), seven out of 10 patients (70%) had adrenal insufficiency within the first week. Six patients (60%) had ongoing adrenal insufficiency after 3 weeks and four patients (40%) after 7 weeks. These four patients remained insufficient at the end of their follow-up, that is after 10, 11, 11 and 19 weeks, respectively. The latter patient, with no recovery by 19 weeks, received, in addition to the induction therapy, two 1-week long re-induction courses including prednisolone 60 mg/m²/day during adrenal insufficiency. After re-induction therapy (n = 7), five out of seven patients (71%) had adrenal insufficiency within the first week. Four patients (57%) had ongoing adrenal insufficiency after 3 weeks and three patients (43%) after 7 weeks. These three patients remained insufficient at the end of their follow-up, that is after 16, 33 and 34 weeks, respectively. One of these patients, with no recovery after 16 weeks, received an additional 1-week long re-induction course of prednisolone 60 mg/m²/day during the period of adrenal insufficiency. The other two patients, with no recovery after 33 and 34 weeks, received three additional 1-week long re-induction courses including prednisolone 60 mg/m²/day during adrenal insufficiency. See Table 3.5 for an overview and Figure 3.1 for the Kaplan-Meier curve of the prevalence and duration of adrenal insufficiency in the studies that used an ACTH suppression test.

**Table 3.5 Prevalence and duration of adrenal insufficiency evaluated by an ACTH stimulation test**

<table>
<thead>
<tr>
<th>Felner et al.</th>
<th>Therapy: dexamethasone (cumulative dose 168 mg/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time after cessation</td>
<td>Before</td>
</tr>
<tr>
<td>n insufficient/n total</td>
<td>0/10</td>
</tr>
</tbody>
</table>

| Petersen et al. (1) | Therapy: prednisolone (cumulative dose 2257.5 mg/m²)
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Time after cessation</td>
<td>1 week</td>
</tr>
<tr>
<td>End of follow-up:</td>
<td>10, 11, 11 and 19 weeks, respectively</td>
</tr>
<tr>
<td>n insufficient/n total</td>
<td>7/10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Petersen et al. (2)</th>
<th>Therapy: dexamethasone (cumulative dose 236.25 mg/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time after cessation</td>
<td>1 week</td>
</tr>
<tr>
<td>End of follow-up:</td>
<td>16, 33 and 34 weeks, respectively</td>
</tr>
<tr>
<td>n insufficient/n total</td>
<td>5/7</td>
</tr>
</tbody>
</table>

* One patient received additional 840 mg/m² prednisolone during the period of adrenal insufficiency.

* These patients received prednisolone 2257.5 mg/m² as induction therapy before. Three high-risk patients received additional 560 mg/m² prednisolone in advance. One patient received additional 420 mg/m² prednisolone during the period of adrenal insufficiency and two patients received additional 1260 mg/m² prednisolone during the period of adrenal insufficiency.
Three studies used the low-dose ACTH stimulation test with comparable cut-off limits (stimulated cortisol: 18 µg/dL (500 nmol/L)) for measuring adrenal function.\textsuperscript{27-29} In the study of Mahachoklertwattana et al. all patients received induction therapy with prednisolone 40 mg/m\(^2\)/day.\textsuperscript{28} Four weeks after its completion this was followed by maintenance therapy every 4 weeks consisting of 7 days of dexamethasone 8 mg/m\(^2\)/day. Baseline cortisol levels before induction therapy and 2 weeks afterwards were not significantly different between the adrenal suppressed and the unsuppressed group. Eleven out of 24 patients (46\%) had adrenal insufficiency 2 weeks after abrupt cessation of 28 days of induction therapy. Nine patients (38\%) had ongoing adrenal insufficiency after 4 weeks, seven patients (29\%) after 8 weeks and three patients (13\%) after 12 weeks. The latter three patients remained insufficient at the end of follow-up at 20 weeks.

In the study of Rix et al., three types of glucocorticoid treatment were evaluated.\textsuperscript{29} All patients (standard risk, intermediate risk and high risk) received induction therapy (22 patients in total were tested afterwards; for two patients no information was available) comprising 35 days of prednisolone 60 mg/m\(^2\)/day with tapering over 9 days. All patients also received 7-day courses of prednisolone 60 mg/m\(^2\)/day without tapering (13 patients in total were

\begin{figure}
\centering
\includegraphics[width=\textwidth]{ACTH_tests.png}
\caption{Kaplan-Meier estimate of prevalence and duration of adrenal insufficiency after treatment with glucocorticoid therapy for childhood acute lymphoblastic leukaemia based on ACTH stimulation tests.}
\end{figure}
tested afterwards) and intermediate-risk and high-risk patients received additionally 21 days of dexamethasone 10 mg/m²/day with tapering over 9 days (7 patients in total were tested afterwards). The 7-day course of prednisolone and the dexamethasone course followed induction therapy. The intermediate-risk group received the dexamethasone course before the 7-day course of prednisolone and the high-risk group received the dexamethasone course after the 7-day course of prednisolone. Based on additional information provided by the authors, 13 patients were tested before induction therapy, and all had a normal adrenal function. Sixteen patients out of 22 (73%) had adrenal insufficiency 1 day after cessation of induction therapy; one of them was lost to follow-up thereafter. Five out of 22 patients (23%) were not tested at this time point. Eight patients (36%) had ongoing adrenal insufficiency after 3 days; two of them were lost to follow-up afterwards. Seven patients (32%) (including the patient that was lost to follow-up) with no confirmed adrenal recovery were not tested at this time point. Eight patients (36%) had ongoing adrenal insufficiency at the end of follow-up of 5 days. Three patients (14%) (i.e., the three patients that were lost to follow-up) with no confirmed adrenal recovery were not tested at this time point. After 7-day courses of prednisolone, all 13 patients (100%) remained insufficient at the end of the follow-up period of 2 days. However, two patients out of 13 (15%) already had an insufficient adrenal function before prednisolone therapy. Five patients out of seven (71%) underwent a low-dose ACTH test before dexamethasone therapy; all patients had sufficient cortisol levels. One day after the dexamethasone course, two out of seven patients (29%) had adrenal insufficiency; one of them was lost to follow-up thereafter. Five out of seven patients (71%) were not tested at this time point. Three patients (43%) had ongoing adrenal insufficiency after 3 days; one of them was lost to follow-up afterwards. Two patients (29%) (including the patient that was lost to follow-up) with no confirmed adrenal recovery were not tested at this time point. One patient (14%) had ongoing adrenal insufficiency at the end of the follow-up period of 7 days. Two patients (29%) (i.e., the two patients that were lost to follow-up) with no confirmed adrenal recovery were not tested at this time point.

In the study of Einaudi et al., two randomised arms of glucocorticoid therapy were examined; 22 days of prednisolone 60 mg/m²/day with tapering over 9 days (n = 40) and 22 days of dexamethasone 10 mg/m²/day with tapering over 9 days (n = 24). Both groups of patients received 7 days of prednisolone 60 mg/m²/day in advance. At diagnosis, basal cortisol values were within the normal range in all patients. In the prednisolone arm, 32 out of 40 patients (80%) had adrenal insufficiency 1 day after cessation of glucocorticoid therapy, eight patients (20%) had ongoing adrenal insufficiency after 7 to 14 days (in the Kaplan Meier curve marked as 14 days) and five patients (13%) after 28 days. All patients (100%) recovered in 10 weeks (in the Kaplan Meier curve marked as 10 weeks). In the dexamethasone arm, 20 out of 24 patients (83%) had adrenal insufficiency 1 day after cessation of glucocorticoid therapy. Four patients (17%) had ongoing adrenal insufficiency after 7 to 14 days (in the Kaplan Meier curve marked as 14 days) and three patients (13%) after 28 days. All patients (100%) recovered in
Table 3.6 Prevalence and duration of adrenal insufficiency evaluated by a low-dose ACTH stimulation test

<table>
<thead>
<tr>
<th>Study</th>
<th>Therapy: prednisolone (cumulative dose 1120 mg/m²)¹</th>
<th>Therapy: prednisolone (cumulative dose 2257.5 mg/m²)</th>
<th>Therapy: dexamethasone (cumulative dose 236.25 mg/m²)²</th>
<th>Therapy: dexamethasone (cumulative dose 1477.5 mg/m²)³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mahachoklertwattana et al.</td>
<td>Time after cessation: 2 weeks, 4 weeks, 8 weeks, 12 weeks, 20 weeks</td>
<td>n insufficient/n total: 11/24, 9/24, 7/24, 3/24, 3/24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rix et al. (1)</td>
<td>Time after cessation: Before, 1 day, 3 days, 5 days</td>
<td>n insufficient/n total: 0/13, 16/17, 8/15, 8/17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rix et al. (2)</td>
<td>Time after cessation: Before, 2 days</td>
<td>n insufficient/n total: 2/13, 13/13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rix et al. (3)</td>
<td>Time after cessation: Before, 1 day, 3 days, 7 days</td>
<td>n insufficient/n total: 0/5, 2/2, 3/5, 1/5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Einaudi et al. (1)</td>
<td>Time after cessation: 1 day, 7 to 14 days, 28 days, 42 days, 10 weeks</td>
<td>n insufficient/n total: 32/40, 8/32, 5/8, 5/5, 0/5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Einaudi et al. (2)</td>
<td>Time after cessation: 1 day, 7 to 14 days, 28 days, 42 days, 10 weeks</td>
<td>n insufficient/n total: 20/24, 4/20, 3/4, 3/3, 0/3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹ Four weeks after completion of the induction therapy, the patients received maintenance therapy consisting of a 7-day course of high-dose dexamethasone 8 mg/m²/day, every 4 weeks. The cumulative dose depended on how long the patient had been followed up.

² All patients first received prednisolone (cumulative dose 2257.5 mg/m²).

³ After 7 days of prednisone (60 mg/m²/day, cumulative dose: 420 mg/m²).

10 weeks (in the Kaplan Meier curve marked as 10 weeks). See Table 3.6 for an overview and Figure 3.2 for the Kaplan-Meier curve of the prevalence and duration of adrenal insufficiency in the studies that used a low-dose ACTH test.

Two studies used basal morning cortisol values for measuring adrenal function. In the study of Cunha et al., including 35 children, median basal cortisol levels were inhibited on the 8th day (1.2 µg/dL, range 0.9 to 132.7 µg/dL) and the 28th day (0.9 µg/dL, range 0.9 to 6.6 µg/dL) of 28 days of dexamethasone 6 mg/m²/day and 48 hours after cessation (over 10 days) of dexamethasone treatment (2.4 µg/dL, range 0.9 to 11.2 µg/dL) compared to pre-glucocorticoid therapy levels (17.5 µg/dL, range 7.6 to 40.9 µg/dL) (P = 0.01 for the three tests versus pre-glucocorticoid levels).³⁵ Median basal cortisol levels 1 month after cessation of dexamethasone treatment (12.4 µg/dL, range 1.8 to 29.0 µg/dL), although slightly lower, did not show a significant difference compared with pre-glucocorticoid therapy levels. No data on patient level were available.

In the study of Kuperman et al., including 15 children, mean basal cortisol levels (± standard error of the mean) were significantly (P < 0.05) lower on day 7 (10.8 ± 1.0 µg/dL) and day 14 (11.5 µg/dL)
After abrupt cessation of 42 days of dexamethasone 6 mg/m²/day than pretreatment (17.8 ± 1.3 µg/dL). Levels at day 7 and day 14 did not differ significantly. Based on additional information provided by the authors, all patients (100%) had sufficient basal cortisol levels at diagnosis (> 7 µg/dL), whereas four out of 15 (27%) patients had insufficient basal cortisol levels 7 days after cessation of dexamethasone therapy. One patient was lost to follow-up thereafter. Fourteen days after cessation of dexamethasone therapy, four (29%) out of 14 patients had insufficient basal cortisol levels. It should be noted that one of them had a sufficient basal cortisol level 7 days earlier. Therefore, these data could not be presented in a Kaplan-Meier curve.

See Table 3.7 for an overview of the prevalence and duration of adrenal insufficiency in the studies that used basal morning cortisol values.

Other outcome measures
Identifying whether the adrenal function after administration of glucocorticoids was dependent on the moment of testing, the (cumulative) dose, type or duration of glucocorticoid therapy or the method of cessation of glucocorticoid therapy was not possible due to the heterogeneity.
Discussion

With the improvement of survival of childhood ALL, treatment-related side effects become increasingly relevant. Glucocorticoids play an important role in the treatment of childhood ALL, but supraphysiological doses may suppress the HPA axis, resulting in an impaired stress response and an inadequate defence against infections.\textsuperscript{10, 11} Patients with HPA axis suppression may benefit from glucocorticoid replacement therapy (e.g. hydrocortisone) to reduce the risk of life-threatening complications. Unfortunately, little information is known on the occurrence and duration of HPA axis suppression and adequate guidelines for glucocorticoid substitution are lacking. This is the first systematic review evaluating HPA axis function after treatment with glucocorticoid therapy for childhood ALL.

We identified seven studies evaluating adrenal function after treatment with glucocorticoid therapy for childhood ALL, including one RCT. None of the studies evaluated the HPA axis at the level of the hypothalamus, pituitary, or both. Due to the substantial differences in types and (cumulative) doses of glucocorticoids used, in duration and method of cessation of glucocorticoid therapy and in method of testing adrenal function, pooling of results was not possible. The included studies demonstrated that adrenal insufficiency occurs in nearly all patients in the first days after cessation of glucocorticoid treatment for childhood ALL. The majority of patients recovered from adrenal insufficiency within 7 weeks. However, a small amount of patients had ongoing adrenal insufficiency lasting up to 34 weeks. At first impression, no significant differences in the occurrence and duration of adrenal insufficiency between different types, (cumulative) doses, durations and methods of cessation of glucocorticoid therapy were present, but due to the heterogeneity between studies we were not able to assess this further. Due to this, and the small patient numbers in the included studies (i.e. low power), definitive conclusions cannot thus not be provided. Only one study was designed as an RCT, enabling comparison between two different types of glucocorticoid therapy; prednisolone versus dexamethasone.\textsuperscript{27} Both treatment groups received prednisolone prior to the randomised treatment. The occurrence and duration of adrenal insufficiency did not differ between the treatment arms.

Table 3.7 Prevalence and duration of adrenal insufficiency evaluated by basal morning cortisol values

<table>
<thead>
<tr>
<th>Study</th>
<th>Time after cessation</th>
<th>n insufficient/n total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cunha et al.</td>
<td>0/15</td>
<td>4/15</td>
</tr>
<tr>
<td>Kuperman et al.</td>
<td>Before 1 week 2 weeks</td>
<td>0/15 4/15 4/14</td>
</tr>
</tbody>
</table>

\* Results based on basal cortisol levels; cut-off level 7 μg/dL = 194 nmol/L.
Previous studies demonstrated adrenal suppression after high-dose fluconazole therapy. Only one of the studies that was included in this review, reported on fluconazole therapy. In this study, two of the three patients receiving fluconazole had ongoing adrenal insufficiency 8 months after cessation of dexamethasone therapy, whereas the third patient recovered after 3 weeks. It should be taken into account that fluconazole therapy may have influenced the duration of adrenal insufficiency in these patients.

Due to the paucity of RCTs on HPA axis suppression after glucocorticoid therapy in childhood ALL, the majority of studies included in this systematic review were uncontrolled studies. Only one RCT was identified. Due to the lack of control groups it was impossible to evaluate possible causes of HPA axis suppression other than glucocorticoid therapy. Furthermore, all included studies used biochemical markers to evaluate adrenal insufficiency; they did not discuss the clinical consequences of adrenal insufficiency. All included studies had methodological limitations, but currently, this is the best available evidence on this topic.

Authors’ conclusions

Implications for practice
Based on the currently available evidence, we can conclude that adrenal insufficiency in the first days after cessation of glucocorticoid therapy for childhood ALL routinely occurs, but the exact duration of adrenal insufficiency remains unclear. Since no data on the level of the hypothalamus and the pituitary were available we cannot make any conclusions regarding those outcomes. The majority of patients seem to recover within 3 days to 7 weeks. However, a small number of patients have prolonged adrenal insufficiency, persisting up to several months. Clinicians should consider prescribing glucocorticoid replacement therapy (e.g. hydrocortisone) during periods of serious stress in the first weeks after cessation of glucocorticoid therapy for childhood ALL, to reduce the risk of life-threatening complications. If replacement therapy is indicated, the beneficial effects and side effects should be evaluated. Until results of adequate future studies on the incidence and duration of HPA axis suppression are available, an HPA axis stimulation test can be recommended, for example, 2 months after cessation of glucocorticoids, to determine if the HPA axis has recovered and if replacement therapy during periods of stress can be discontinued. Exclusively morning cortisol levels are inappropriate to evaluate HPA axis recovery because they only reflect basal cortisol prediction and not the ability of the HPA axis to respond to stress.

Special attention should be paid to patients receiving fluconazole therapy, and perhaps similar antifungal drugs, as this may prolong the duration of adrenal insufficiency.

No definitive conclusions regarding differences in the occurrence and duration of adrenal insufficiency between the type (prednisolone versus dexamethasone), (cumulative) dose, duration and method of cessation (abrupt or gradual) of glucocorticoid therapy can be made.
Implications for research

Studies examining HPA axis suppression after high-dose glucocorticoid therapy for childhood ALL are scarce, especially RCTs. High-quality research regarding the occurrence and duration of HPA axis suppression after glucocorticoid therapy for childhood ALL is needed in order to formulate adequate evidence-based guidelines for glucocorticoid replacement therapy. It will also be of interest to compare the serious adverse effects in patients with, and without, persistent HPA axis suppression. In addition, there is a need for studies evaluating the long-term effects of glucocorticoid therapy on the HPA axis.

Future studies should focus on identifying differences in the effect of type, (cumulative) dose, repeated exposure, duration and method of cessation of glucocorticoid therapy on the occurrence and duration of HPA axis suppression. The number of included patients should be sufficient to obtain the power needed for reliable results.

Furthermore, an interesting and relevant topic for future research would be the (genetic) susceptibility of individuals for HPA axis suppression after glucocorticoid treatment.
References


Dickstein G, Spigel D, Arad E, Shechner C. One microgram is the lowest ACTH dose to cause a maximal cortisol response. There is no diurnal variation of cortisol response to submaximal ACTH stimulation. Eur J Endocrinol 1997;137(2):172-5.


Appendices

Appendix 1

Search strategy for PubMed

1. For children the following MeSH headings and text words were used:
   - infant OR infan* OR newborn OR newborn* OR new-born* OR baby OR baby* OR babies
   - OR neonat* OR perinat* OR postnat* OR child OR child* OR schoolchild* OR schoolchild OR
   - school child OR school child* OR kid OR kids OR toddler* OR adolescent OR adole* OR teen*
   - OR boy* OR girl* OR minors OR minors* OR underag* OR under ag* OR juvenil* OR youth*
   - OR kindergar* OR puberty OR puber* OR pubescen* OR prepubescent* OR prepuberty* OR
   - pediatrics OR pediatric* OR paediatric* OR paediatric* OR schools OR nursery school* OR
   - preschool* OR pre school* OR primary school* OR secondary school* OR elementary school* OR
   - elementary school OR high school* OR highschool* OR school age OR schoolage OR
   - school age* OR schoolage* OR infancy OR schools, nursery OR infant, newborn

2. For acute lymphocytic leukemia the following MeSH headings and text words were used:
   - acute lymphocytic leukemia OR childhood ALL OR Precursor Cell Lymphoblastic Leukemia-
   - Lymphoma OR Precursor Cell Lymphoblastic Leukemia Lymphoma OR Leukemia, Lympho-
   - blastic OR Leukemia, Lymphoblastic, Acute OR Leukemia, Lymphocytic, Acute OR Leukemia,
   - Lymphoid, Acute OR Lymphoblastic Leukemia, Acute OR Acute Lymphoblastic Leukemia OR
   - Leukemia, Acute Lymphoblastic OR Lymphoblastic Lymphoma OR Lymphomas, Lymphoblas-
   - tic OR Lymphoblastic Leukemia, Acute OR Acute Lymphocytic Leukemia OR Leukemia, Acute
   - Lymphocytic OR Lymphoma, Lymphoblastic OR Acute Lymphoid Leukemia OR Leukemia,
   - Acute Lymphoid OR Lymphoid Leukemia, Acute OR Lymphoblastic Leukemia OR Leukemia,
   - Lymphocytic, Acute, L1 OR Lymphocytic Leukemia, L1 OR L1 Lymphocytic Leukemia OR
   - Leukemia, L1 Lymphocytic OR Lymphoblastic Leukemia, Acute, Childhood OR Lymphoblastic
   - Leukemia, Acute, L1 OR ALL, Childhood OR Childhood ALL OR Leukemia, Lymphoblastic,
   - Acute, L1 OR Leukemia, Lymphocytic, Acute, L2 OR Lymphoblastic Leukemia, L2 OR L2 Lym-
   - phocytic Leukemia OR Leukemia, L2 Lymphocytic OR Lymphoblastic Leukemia, Acute, Adult
   - OR Lymphoblastic Leukemia, Acute, L2 OR Leukemia, Lymphoblastic, Acute, L2 OR Leukemia,
   - Lymphoblastic, Acute, Philadelphia-Positive OR ((akut* OR acut*) AND ((leukemi* OR leu-
   - kaemi*) OR (lymphocyt* OR lymphoblast*))))
3. For glucocorticoids the following MeSH headings and text words were used:
Steroid OR steroids OR steroid* OR glucocorticoid OR glucocorticoids OR glucocorticoid* OR corticoid OR corticoids OR corticoid* OR adrenal cortex hormones OR hormones, adrenal cortex OR prednisone OR prednisolone OR 53-03-2 OR Dehydrocortisone OR delta-Cortisone OR Winpred OR ICN Brand of Prednisone OR Cortancyl OR Panaforcort OR Aventis Brand of Prednisone OR Cutason OR mibe Brand of Prednisone OR Dacortin OR Merck Brand of Prednisone OR Decortin Brand of Prednisone OR Decortisyl OR Hoechst Brand of Prednisone OR Deltasone OR Pharmacia Brand of Prednisone OR Encortone OR Encorton OR Enkortolon OR Cortancyl OR Liquid Pred OR Meticorten OR Schering-Plough Brand of Prednisone OR Orasone OR Solvay Brand of Prednisone OR Panasol OR Seatrace Brand of Prednisone OR Predni Tablinen OR Lichtenstein Brand of Prednisone OR Prednidib OR Diba Brand of Prednisone OR Predniment OR Ferring Brand of Prednisone OR Prednison acsis OR acis Brand of Prednisone OR Prednison Galen OR GALENpharma Brand of Prednisone OR Prednison Hexal OR Hexal Brand of Prednisone OR Pronisone OR Rectodelt OR Trommsdorff Brand of Prednisone OR Ultracorten OR Sone OR Fawns & McAllan Brand of Prednisone OR Sterapred OR Merz Brand of Prednisone OR Apo-Prednisone OR Apotex Brand of Prednisone OR Cortan OR Halsey Drug Brand of Prednisone OR prednisolon OR 50-24-8 OR dexamethason OR dexametasone OR dexamethasone OR 50-02-2 OR Methylfluorprednisolone OR Hexadecadrol OR Maxidex OR Alcon Brand of Dexamethasone OR Dexamethasone Intensol OR Roxane Brand of Dexamethasone OR Dexasone OR Decaject OR Merz Brand 1 of Dexamethasone OR Oradexon OR Decameth OR Foy Brand of Dexamethasone OR Decaspray OR Merck Brand of Dexamethasone OR Dexpak OR ECR Brand of Dexamethasone OR Decaject-L.A. OR Decaject L.A. OR Merz Brand 2 of Dexamethasone

4. For HPA function the following MeSH headings and text words were used:
HPA OR HPA axis OR hypothalamic-pituitary-adrenal OR hypothalamic-pituitary-adrenal axis OR adrenal insufficiency OR adrenal axis OR Hypothalamo-Hypophyseal System OR hypothalamic insufficiency OR pituitary OR Pituitary-Adrenal Function Tests OR Function Test, Pituitary-Adrenal OR Function Tests, Pituitary-Adrenal OR Pituitary Adrenal Function Tests OR Pituitary-Adrenal Function Test OR Test, Pituitary-Adrenal Function OR Tests, Pituitary-Adrenal OR Pituitary-Adrenal System OR Pituitary Adrenal System OR Pituitary-Adrenal Systems OR System, Pituitary-Adrenal OR Systems, Pituitary-Adrenal OR hypothalamus OR hypothalam* OR hypophysis OR hypophys* OR Hypothalamo-Hypophyseal OR Pituitary Gland OR Pituitary Glands OR Hypothalamic OR Hypothalamic-Regulating Peptides OR Hypothalamic Pituitary Regulating Peptides OR Peptides, Hypothalamic Pituitary-Regulating OR Pituitary-Regulating Peptides, Hypothalamic OR Hypothalamic Pituitary-Regulating OR Hypothalamic Pituitary-Regulating OR Hypothalamic Pituitary-Regulating OR Hypothalamic Pituitary-Regulating OR Hypothalamic Pituitary-Regulating OR Hypothalamic Pituitary-Regulating OR Hypothalamic Pituitary-Regulating OR Hypothalamic Pituitary-Regulating OR Hypothalamic Pituitary-Regulating
Appendix 2

Search strategy for EMBASE (Ovid)

1. For children the following Emtree terms and text words were used:
   1. infant/ or infancy/ or newborn/ or baby/ or child/ or preschool child/ or school child/
   2. adolescent/ or juvenile/ or boy/ or girl/ or puberty/ or prepuberty/ or pediatrics/
   3. primary school/ or high school/ or kindergarten/ or nursery school/ or school/
   4. or/1-3
   5. (infant$ or newborn$ or (new adj born$) or baby or baby$ or babies or neonate$ or perinat$ or postnat$).mp.
   6. (child$ or (school adj child$) or schoolchild$ or (school adj age$) or schoolage$ or (pre adj school$) or preschool$).mp.
   7. (kid or kids or toddler$ or adole$ or teen$ or boy$ or girl$).mp.
   8. (minors$ or (under adj ag$) or underage$ or juvenil$ or youth$).mp.
   9. (puber$ or pubescen$ or prepubescent$ or prepubert$).mp.
   10. (pediatric$ or paediatric$ or peadiatric$).mp.
   11. (school or schools or (high adj school$) or highschool$ or (primary adj school$) or (nursery adj school$) or (elementary adj school) or (secondary adj school$) or kindergar$).mp.
   12. or/5-11
   13. 4 or 12
2. For acute lymphocytic leukemia the following Emtree terms and text words were used:
1. acute lymphocytic leukemia.mp. or exp Acute Lymphocytic Leukemia/
2. (childhood adj ALL).
3. precursor cell lymphoblastic leukemia-lymphoma.mp. or exp Acute Lymphoblastic Leukemia/
4. lymphoblastic lymphoma.mp. or exp Lymphoblastoma/
5. (acute lymphoblastic leukemia or acute lymphoid leukemia).mp.
6. lymphoblastic leukemia.mp. or exp Lymphatic Leukemia/
7. (L1 lymphocytic leukemia or L2 lymphocytic leukemia).mp.
8. (lymphoblastic lymphoma or lymphoblastic lymphomas).mp. or exp lymphatic leukemia/
9. childhood acute lymphoblastic leukemia.mp. or exp Childhood Leukemia/
10. philadelphia positive acute lymphoblastic leukemia.mp.
11. ((akut$ or acut$) and (leukemi$ or leukaemi$ or lymphocyt$ or lymphoblast$)).mp.
12. or/1-11

3. For glucocorticoids the following Emtree terms and text words were used:
1. steroid.mp. or exp steroid/
2. (steroids or steroid$).mp.
3. glucocorticoid.mp. or exp glucocorticoid/
4. (glucocorticoids or glucocorticoid$).mp.
5. corticoid.mp. or exp corticosteroid/
6. (corticoids or corticoid$).mp.
7. adrenal cortex hormones.mp.
8. prednison.mp. or exp prednisone/
9. prednisone.mp.
10. 53-03-2.rn.
11. (dehydrocortisone or delta-cortisone or winpred or cortancyl or panafcort or cutason or dacortin or decortisyl).mp.
12. (deltasone or encortone or encorton or emkortolon or kortancyl or meticorten).mp.
13. (orasone or panasol or predni tablinen or prednidib or predniment or prednison acsis).mp.
14. (prednison galen or prednison hexal or pronisone or rectodelt or ultracorten or sone).mp.
15. (sterapred or apo-prednisone or cortan or prednisolon).mp.
16. 50-24-8.rn.
17. exp dexamethasone/
18. (dexamethason or dexamethasone or dexametasone).mp.
19. 50-02-2.rn.
20. (methylfluorprednisolone or hexadecadrol or maxidex or dexamethasone intensol).mp.
21. (decaject or oradexon or decameth or decaspray or dexasone or hexadrol or millicorten or dexpak or decaject-l or decaject la).mp.
22. or/1-21

4. For HPA function the following Emtree terms and text words were used:
1. exp hypothalamus hypophysis adrenal system/
2. (HPA or HPA axis).mp.
4. adrenal insufficiency.mp. or exp adrenal insufficiency/
5. adrenal axis.mp.
6. hypothalamo-hypophyseal system.mp. or exp hypothalamus hypophysis system/
7. hypothalamic insufficiency.mp.
8. pituitary.mp.
9. (pituitary-adrenal function tests or pituitary-adrenal function test).mp.
10. (pituitary adrenal function tests or pituitary adrenal function test).mp.
11. (pituitary adrenal system or pituitary-adrenal system or pituitary adrenal systems or pituitary-adrenal systems).mp.
12. exp hypothalamus/ or hypothalamus.mp.
13. hypothalam$.mp.
14. exp hypophysis/ or hypophysis.mp.
15. hypophys$.mp.
16. exp hypothalamus hypophysis system/ or hypothalamo-hypophyseal.mp.
17. (pituitary gland or pituitary glands).mp.
18. hypothalamic hormones.mp. or exp hypothalamus hormone/
19. (hypothalamic pituitary-regulating peptides or hypothalamic pituitary regulating peptides).mp.
20. (hypothalamic pituitary-regulating hormones or hypothalamic pituitary regulating hormones).mp.
21. pituitary hormones.mp. or exp hypophysis hormone/
22. adrenal glands.mp. or exp adrenal gland/
23. adrenal function test.mp. or exp endocrine function test/
24. (adrenal function test or adrenal function testing or adrenal function evaluation).mp.
25. ACTH stimulation test.mp. or exp corticotropin test/
26. (ACTH stimulation tests or ACTH stimulation testing).mp.
27. (ACTH test or ACTH tests or ACTH testing or ACTH evaluation).mp.
28. (CRH stimulation test or CRH stimulation tests or CRH stimulation testing).mp.
29. (CRH test or CRH tests or CRH testing or CRH evaluation).mp.
30. (glucagon stimulation test or glucagon stimulation tests or glucagon stimulation testing).mp.
31. (glucagon test or glucagon tests or glucagon testing or glucagon evaluation).mp.
32. (fasting cortisol or morning cortisol).mp.
33. or/1-32

Final search: 1 AND 2 AND 3 AND 4

[mp = title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name; $ = 1 or more characters; / = Emtree term; rn = registry number]

Appendix 3

Search strategy for the Cochrane Central Register of Controlled Trials (CENTRAL)

1. For children the following text words were used:
   (infant OR infant* OR newborn OR newborn* OR new-born* OR baby OR baby* OR babies OR neonat* OR child OR child* OR schoolchild* OR schoolchild OR school child OR school child* OR kid OR kids OR toddler* OR adolescent OR adoles* OR teen* OR boy* OR girl* OR minors OR minors* OR underag* OR under ag* OR juvenil* OR youth* OR kindergarten* OR puberty OR puber* OR pubescen* OR prepubescen* OR prepuberty* OR pediatrics OR pediatric* OR paediatric* OR peadiatric* OR schools OR nursery school* OR preschool* OR pre school* OR primary school* OR secondary school* OR elementary school* OR elementary school OR high school* OR highschool* OR school age OR schoolage OR school age* OR schoolage* OR infancy)

2. For acute lymphocytic leukemia the following text words were used:
   acute lymphocytic leukemia OR childhood ALL OR Precursor Cell Lymphoblastic Leukemia-Lymphoma OR Precursor Cell Lymphoblastic Leukemia Lymphoma OR Acute Lymphoblastic Leukemia OR Lymphoblastic Lymphoma OR Acute Lymphocytic Leukemia OR Acute Lymphoid Leukemia OR Lymphoblastic Leukemia OR L1 Lymphocytic Leukemia OR L2 Lymphocytic Leukemia OR ((akut* OR acut*) AND (leukemi* OR leukaemi*)) OR (lymphocyt* OR lymphoblast*))

3. For glucocorticoids the following text words were used:
   Steroid OR steroids OR steroid* OR glucocorticoid OR glucocorticoids OR glucocorticoid* OR corticoid OR corticoids OR corticoid* OR adrenal cortex hormones OR prednison OR prednisone OR Dehydrocortisone OR delta-Cortisone OR Winpred OR ICN Brand of Prednisone OR Cortancyl OR Panafort OR Aventis Brand of Prednisone OR Cutason OR mibe Brand of Prednisone OR Dacortin OR Merck Brand of Prednisone OR Decortin Brand of Prednisone OR Decortisyl OR Hoechst Brand of Prednisone OR Deltasone OR Pharmacia Brand of Prednisone OR Encortone OR Encorton OR Enkortolon OR Kortancyl OR Liquid Pred OR Meticorten OR
Schering-Plough Brand of Prednisone OR Orasone OR Solvay Brand of Prednisone OR Panasol OR Seatrace Brand of Prednisone OR Predni Tablinen OR Lichtenstein Brand of Prednisone OR Prednidib OR Diba Brand of Prednisone OR Prednim OR Ferring Brand of Prednisone OR Pronisone OR Rectodelt OR Trommsdorff Brand of Prednisone OR Ultracorten OR Sone OR Fawns & McAllan Brand of Prednisone OR Sterapred OR Merz Brand 1 of Prednisone OR Apo-Prednisone OR Apotex Brand of Prednisone OR Cortan OR Halsey Drug Brand of Prednisone OR prednisolon OR dexamethason OR dexametasone OR dexamethasone OR Methylfluorprednisolone OR Hexadecadrol OR Maxidex OR Alcon Brand of Dexamethasone OR Dexamethasone Intensol OR Roxane Brand of Dexamethasone OR Decaject OR Merz Brand 1 of Dexamethasone OR Oradexon OR Decameth OR Foy Brand of Dexamethasone OR Decaspray OR Merck Brand of Dexamethasone OR Dexasone OR ICN Brand of Dexamethasone OR Hexadrol OR Millicorten OR Dexpak OR ECR Brand of Dexamethasone OR Decaject-L.A. OR Decaject L.A. OR Merz Brand 2 of Dexamethasone

4. For HPA function the following text words were used:
HPA OR HPA axis OR hypothalamic-pituitary-adrenal OR hypothalamic-pituitary-adrenal axis OR adrenal insufficiency OR adrenal axis OR Hypothalamo-Hypophyseal System OR hypothalamic insufficiency OR pituitary OR Pituitary-Adrenal Function Tests OR Pituitary Adrenal Function Tests OR Pituitary-Adrenal Function Test OR Pituitary-Adrenal System OR Pituitary Adrenal System OR Pituitary-Adrenal Systems OR hypothalamus OR hypophysis OR hypophys* OR Hypothalamo-Hypophyseal OR Pituitary Gland OR Pituitary Glands OR Hypothalamic Hormones OR Hypothalamic Pituitary-Regulating Peptides OR Hypothalamic Pituitary Regulating Peptides OR Hypothalamic Pituitary-Regulating Hormones OR Hypothalamic Pituitary Regulating Hormones OR Pituitary Hormones OR adrenal glands OR adrenal function test OR adrenal function tests OR adrenal function testing OR adrenal function evaluation OR ACTH stimulation test OR ACTH stimulation tests OR ACTH stimulation testing OR ACTH test OR ACTH tests OR ACTH testing OR ACTH evaluation OR CRH stimulation test OR CRH stimulation tests OR CRH stimulation testing OR CRH test OR CRH tests OR CRH testing OR CRH evaluation OR glucagon stimulation test OR glucagon stimulation tests OR glucagon stimulation testing OR glucagon test OR glucagon tests OR glucagon testing OR glucagon evaluation OR fasting cortisol OR morning cortisol

Final search: 1 and 2 and 3 and 4
The search was performed in title, abstract or keywords.
Appendix 4

Search strategy for conference proceedings

Search used in the International Society of Paediatric Oncology (SIOP) proceedings: (http://www.siop.nl)
Separate searches were done for the following search terms:
HPA
HPA axis
adrenal
hypothalamus
hypothalamic-pituitary-adrenal

Search used in American Society of Clinical Oncology (ASCO) proceedings: (http://www.asco.org)
Only annual meetings were searched (since molecular markers, breast, genitourinary, gastrointestinal and prostate specific meetings were not expected to yield results that involved HPA axis suppression during acute lymphoblastic leukemia).
Separate searches were done for the following search terms:
HPA
HPA axis
adrenal
hypothalamus
hypothalamic-pituitary-adrenal
The search was performed in title field.

Appendix 5

Search strategy for ongoing trials registers

Search used in the International Standard Randomized Controlled Trial Number (ISRCTN) register and the National Institutes of Health (NIH) register for ongoing trials: (http://www.controlled-trials.com)
Separate searches were done for the following search terms:
HPA
HPA axis
adrenal
hypothalamus
hypothalamic-pituitary-adrenal
The search was performed in title field.
## Supplemental Tables

### Supplemental Table 3.1 Characteristics of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Methods</th>
<th>Participants</th>
<th>Interventions</th>
<th>Outcomes</th>
<th>Notes</th>
</tr>
</thead>
</table>
| **Cunha 2004** | Study type: prospective multicentre study  
Setting: Brazil (University Hospital of Federal University of Minas Gerais, Santa Casa de Misericórdia and Hospital Felício Rocho, Belo Horizonte). This information was based on additional information provided by the authors | 35 children (median age at diagnosis/first HPA axis function test 6.9 years (range 1.2 to 14.4 years); 17 boys and 18 girls) with ALL | Treatment according to the Brazilian Group for Treatment of ALL, 1993 protocol (GBTLI-93).  
Specific medication not defined  
Type of glucocorticoid therapy: dexamethasone (6 mg/m²/day, twice daily) given for 28 days  
Cumulative dexamethasone dose 183.75 mg/m²  
Duration of glucocorticoid therapy: 28 days + 9 days’ tapering doses (in total 37 days)  
Methods of cessation of glucocorticoid therapy: dose reduction over 10 days (50% each 3 days, with complete withdrawal on the 10th day) | Specific HPA axis function test: the ovine CRH stimulus test at 8 a.m. (after an overnight fasting period), including cortisol basal morning value  
Moment of testing: before introduction of dexamethasone, on the 8th day and the 28th day of dexamethasone use, and 48 hours and 1 month after cessation of dexamethasone. The tests were performed during treatment for ALL  
Cut-off limits defined by original studies: baseline cortisol: 5 to 25 µg/dL (138.9 to 694.4 nmol/L) and stimulated cortisol levels were compared with the levels before treatment | 7 patients were lost to follow-up at the 8th and 28th test day, 1 more 48 hours after glucocorticoid therapy test and 7 more 1 month after glucocorticoid therapy test  
Length of follow-up after glucocorticoid therapy: 1 month  
Risk factors were not evaluated |
| **Einaudi 2008** | Study type: multicentre RCT  
Setting: Department of Pediatric Onco-Hematology, University of Turin, Italy and Department of Pediatrics, University of Milano-Bicocca, Hospital of Monza, Italy. This information was based on additional information provided by the authors | 64 children (24 patients that received dexamethasone: mean age at diagnosis 4 years 11 months (range 1 year 2 months to 12 years 1 month), 11 boys and 13 girls; 40 patients that received prednisone: mean age at diagnosis 6 years 9 months (range 1 year 2 months to 17 years 6 months), 18 boys and 22 females) with ALL. This information was based on additional information provided by the authors | | | |
### Supplemental Table 3.1 Characteristics of included studies (Continued)

| Interventions | Treatment according to the AIEOP ALL 2000 study. Induction phase IA: prednisone (from day 8 randomisation prednisone or dexamethasone), vincristine, daunorubicin, Escherichia coli L-asparaginase and intrathecal methotrexate. Induction phase IB: 6-mercaptopurine, cyclophosphamide, cytosine-arabinoside and intrathecal methotrexate
| Type of glucocorticoid therapy: induction phase: oral prednisone (60 mg/m²/day, divided into 3 daily doses) given on days 1 to 7. On day 8 the children were randomised to receive either dexamethasone (10 mg/m²/day) or prednisone (60 mg/m²/day) both divided into 3 oral doses until day 29. From day 30 onwards, the dose of both corticosteroids was tapered by 50% every 3 days, until complete withdrawal over 9 days
| Cumulative dose of glucocorticoid therapy: prednisone 420 mg/m² + dexamethasone 246.25 mg/m² or prednisone 1477.50 mg/m²
| Duration of glucocorticoid therapy: prednisone 7 days + dexamethasone 22 days or prednisone 22 days + 9 days’ tapering doses (in total 39 days)
| Methods of cessation of glucocorticoid therapy: dose was tapered by 50% every 3 days, until complete withdrawal over 9 days
| Outcomes | Specific HPA axis function test: basal cortisol between 8 and 9 a.m. at diagnosis and the low-dose ACTH test between 8 and 11 a.m. (1 µg/1.74 m² of tetracosactrin (Synacthen, Novartis, Basal, Switzerland), basal morning value cortisol, after 30 and 60 minutes)
| Moment of testing: at diagnosis basal cortisol was determined. The first low-dose ACTH stimulation test was performed 24 hours after the last tapering dose of glucocorticoid (on day 39), which was given as a single dose in the morning. The tests were performed during treatment for ALL
| Cut-off limits defined by original studies: basal cortisol: 6 to 30 µg/dL (167 to 833 nmol/L). Low-dose ACTH test: normal response ≥ 18 µg/dL (≥ 500 nmol/L)
| Notes | 0 patients were lost to follow-up
| Length of follow-up after glucocorticoid therapy: patients with suppressed levels underwent further low-dose ACTH tests between 7 and 14 days from the last glucocorticoid dose and every 2 weeks thereafter until cortisol levels normalised. The total follow-up duration was 10 weeks.
| Risk factors were not evaluated

**Felner 2000**

| Methods | Study type: prospective single-centre study
| Setting: Children’s Medical Center of Dallas (University of Texas Southwestern Medical School). This information was based on additional information provided by the authors
| Participants | 10 children (mean age at diagnosis 5.3 ± 2.9 years (range 2.0 to 9.9 years); 7 boys and 3 girls) with early B-cell lineage ALL
| Interventions | Induction therapy: dexamethasone, vincristine, L-asparaginase and daunorubicin. High-risk therapy: 1 additional lumbar puncture with intrathecal chemotherapy during induction
| Type of glucocorticoid therapy: induction phase: oral dexamethasone (6 mg/m²/day, divided into 2 daily doses) for 28 consecutive days
| Cumulative dexamethasone dose 168 mg/m²
| Duration of glucocorticoid therapy: 28 days
| Methods of cessation of glucocorticoid therapy: abrupt
| No control intervention
### Supplemental Table 3.1 Characteristics of included studies (Continued)

| Outcomes | Specific HPA axis function test: 250 µg cosyntropin stimulation test (synthetic corticotrophin 1-24 / ACTH test) iv between 8 and 10 a.m. (Cotrosyn, Organon) (basal morning value cortisol and after 45 minutes)  
Moment of testing: at diagnosis (baseline), 24 hours after completion of the dexamethasone course and every 4 weeks thereafter until normalisation of adrenal function. The tests were performed during treatment for ALL  
Cut-off limits defined by original studies: baseline cortisol: not defined. Low-dose ACTH test: normal response ≥ 18 µg (≥ 500 nmol/L)  
| Notes | 0 patients were lost to follow-up  
Length of follow-up after glucocorticoid therapy: patients with suppressed levels underwent further tests every 4 weeks thereafter until cortisol levels normalised. The total follow up duration was 8 weeks  
Risk factors were not evaluated |
| Study | Kuperman 2001 |
| Methods | Study type: prospective single-centre study  
Setting: The Oncology Department of the Children's Institute, Hospital das Clínicas-Sao Paulo University School of Medicine, Brazil |
| Participants | 15 children (age at diagnosis 5 months to 12 years; 5 boys and 10 girls) with ALL |
| Interventions | Dexamethasone, daunomycin, vincristine, L-asparaginase and cytosine-arabinoside  
Type of glucocorticoid therapy: induction phase: oral dexamethasone (6 mg/m²/day, divided into 3 daily doses) for 42 consecutive days  
Cumulative dexamethasone dose 252 mg/m²  
Duration of glucocorticoid therapy: 42 days  
Methods of cessation of glucocorticoid therapy: abrupt  
No control intervention |
| Outcomes | Specific HPA axis function test: 1 µg/kg ovine CRH stimulation test iv, after an 8-hours fast, between 8 and 9 a.m. (basal morning value cortisol and after 15, 30, 60 and 90 minutes)  
Moment of testing: before dexamethasone therapy (baseline), 7 and 14 days after the last dose of dexamethasone. It was not defined if tests were performed during treatment for ALL  
Cut-off limits defined by original studies: basal cortisol: 7.0 µg/dL. This information was based on additional information provided by the authors. Ovine CRH test: cortisol above 12.8 µg/dL (353.2 nmol/L) was considered normal (basal and peak cortisol levels of the 3 different time points were compared with each other) |
| Notes | Based on additional information provided by the authors, 1 patient was lost to follow-up 14 days after administration of dexamethasone  
Length of follow-up after glucocorticoid therapy: 14 days  
Risk factors were not evaluated |
| Study | Mahachoklertwattana 2004 |
| Methods | Study type: prospective single-centre study  
Setting: Department of Pediatrics, Faculty of Medicine, Ramathibodi Hospital, Bangkok, Thailand |
| Participants | 24 children (median age at diagnosis 3.5 years (range 1 to 14 years); 13 boys and 11 girls) with newly diagnosed ALL |
**Supplemental Table 3.1** Characteristics of included studies (Continued)

| Interventions | According to the modification of St Jude Children’s Research Hospital Total XIII Protocol for ALL. Standard induction therapy: prednisolone, vincristine, L-asparaginase, doxorubicin, etoposide and cytosine-arabinoside
Type of glucocorticoid therapy: induction phase: oral prednisolone (40 mg/m²/day, divided into 3 daily doses) for 28 consecutive days. 4 weeks after completion of the induction therapy, the patients received maintenance therapy consisting of a 7-day course of high-dose dexamethasone 8 mg/m²/day, every 4 weeks in conjunction with other chemotherapeutic agents according to the risk classification
Cumulative prednisolone dose 1120 mg/m². Cumulative dose of dexamethasone per patient depended on how long the patient was followed up. A dexamethasone course (56 mg/m²/day) was administrated at 4, 8, 12 and 16 weeks after induction therapy. The maximum cumulative dose of dexamethasone (4 courses) is 224 mg/m²
Duration of glucocorticoid therapy: 28 days of prednisolone and additional 7-day course of dexamethasone
Methods of cessation of glucocorticoid therapy: abrupt
No control intervention |
| Outcomes | Specific HPA axis function test: serum cortisol level at 8 a.m., at diagnosis (baseline) and low-dose ACTH stimulation test (1 µg cosyntropin, Cotrosyn, Organon, West Orange, NJ) at 8 a.m. after an overnight fast (basal cortisol and after 30 minutes)
Moment of testing: baseline adrenal function was assessed by determination of serum cortisol level at 8 a.m. before induction therapy. The first low-dose ACTH stimulation test was performed 2 weeks after discontinuation of prednisolone. Patients with adrenal insufficiency underwent repeated ACTH tests 4 weeks after completion of the prednisolone course and every 4 weeks thereafter in the morning of the day in which the patients were admitted for the next course of maintenance chemotherapy until normalisation. The tests were performed during treatment for ALL
Cut-off limits defined by original studies: basal cortisol: not defined. Low-dose ACTH test: normal response ≥ 18 µg/dL (≥ 500 nmol/L) |
| Notes | 0 patients were lost to follow-up
Length of follow-up after glucocorticoid therapy: up to 20 weeks
Risk factors were not evaluated |

**Petersen 2003**

| Methods | Study type: prospective single-centre study
Setting: The University Hospital, Rigshospitalet, Copenhagen, Denmark |
| Participants | 17 children (median age at diagnosis 5.4 years (range 2 to 15 years)) with ALL |
### Supplemental Table 3.1 Characteristics of included studies (Continued)

**Interventions**

| Study | Type of glucocorticoid therapy: induction phase (n = 10): prednisolone (60 mg/m²/day, in 3 daily doses) during the first 5 weeks of induction therapy followed by 9 days of tapering. 1-week courses of prednisolone (60 mg/m²/day, based on additional information provided by the authors) were administered every 4 to 10 weeks as part of the re-induction therapy, beginning approximately 8 weeks after prednisolone. Re-induction therapy (n = 7): dexamethasone (10 mg/m²/day, divided in 3 daily doses) for 3 weeks on protocol days 169 to 190 (4 intermediate-risk patients) or days 246 to 267 (3 high-risk patients) followed by 9 days of tapering. The high-risk patients received two 1-week courses of prednisolone (40 mg/m²/day) 4 and 8 weeks prior to the dexamethasone therapy. 1-week courses of prednisolone (60 mg/m²/day) were administered every 4 to 10 weeks as part of the re-induction therapy, beginning approximately 11 weeks after dexamethasone therapy. Cumulative dose: induction therapy: 2100 mg/m² + 157.5 mg/m² prednisolone. 1 patient received additional 840 mg/m² prednisolone during the period of adrenal insufficiency. Re-induction therapy: 210 mg/m² + 26.5 mg/m² dexamethasone. High-risk patients received 560 mg/m² prednisolone in advance 1 patient received an additional 420 mg/m² prednisolone during the period of adrenal insufficiency and 2 patients received an additional 1260 mg/m² prednisolone during the period of adrenal insufficiency. Duration of glucocorticoid therapy: induction therapy: 35 days of prednisolone + 9 days' tapering doses. 1 patient received an additional 14 days of prednisolone. Re-induction therapy: 21 days of dexamethasone + 9 days' tapering doses. The high-risk patients received 14 days of prednisolone 4 and 8 weeks prior to dexamethasone therapy. After the dexamethasone course, the high-risk patients also received prednisolone for 1 week (n = 1) or 3 weeks (n = 2). Methods of cessation of glucocorticoid therapy: 50% each 3 days, over 9 days in total. No control intervention. |

**Outcomes**

| Specific HPA axis function test: ACTH stimulation test (250 µg tetracosactide (Synacthen, Novartis) between 8 and 11 a.m. (basal cortisol and after 30 and 60 minutes). Moment of testing: the adrenal function was assessed by an ACTH stimulation test within 2 weeks after discontinuation of glucocorticoid therapy. The test was repeated every 3 to 5 weeks until recovery or end of follow-up. The tests were performed during treatment for ALL. Cut-off limits defined by original studies: low-dose ACTH test: normal response > 500 nmol/L. |

**Notes**

| No patients were lost to follow-up. Length of follow-up after glucocorticoid therapy: fluctuating. Fluconazole therapy as a risk factor for adrenal insufficiency was evaluated. |

**Rix 2005**

| Study type: prospective multicentre study. Setting: Department of pediatrics, Aalborg University Hospital and the department of pediatrics, Aarhus University Hospital Skejby. This information was based on additional information provided by the authors. |

| 24 children (median age at diagnosis 4.5 years (range 1.8 to 14.6 years); 17 boys and 7 girls) with newly diagnosed ALL. 12 had standard-risk ALL according to the Nordic risk criteria, 7 had intermediate risk and 5 had high risk. |
### Supplemental Table 3.1 Characteristics of included studies (Continued)

<table>
<thead>
<tr>
<th>Interventions</th>
<th>According to the NOPHO ALL-92 protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of glucocorticoid therapy: all children received prednisolone (60 mg/m²/day, in 3 daily doses) during the first 5 weeks of induction therapy followed by 9 days of tapering. All children received 1-week courses of prednisolone (60 mg/m²/day) without tapering, and children with intermediate-risk and high-risk criteria received an additional 3-week course of dexamethasone (10 mg/m²/day) with tapering over a 9-day period. Cumulative dose of glucocorticoid therapy: all patients received 2100 mg/m² + 157.5 mg/m² prednisolone. In addition, several courses (not defined) of 420 mg/m²/day prednisolone. Intermediate-risk and high-risk patients: additional 210 mg/m² + 26.25 mg/m² dexamethasone. Duration of glucocorticoid therapy: induction therapy: 35 days of prednisolone + 9 days' tapering doses. Additional 7-day courses of prednisolone. Intermediate-risk and high-risk patients: additional 21 days dexamethasone + 9 days tapering doses. Methods of cessation of glucocorticoid therapy: 50% every 3 days over 9 days in total. No control intervention.</td>
<td></td>
</tr>
</tbody>
</table>

| Outcomes | Specific HPA axis function test: low-dose ACTH stimulation test (1 µg tetracosactide (Synacthen, Novartis) between 8 and 10 a.m. (basal cortisol and after 30 minutes). Moment of testing: for each child: before the 5-week course of prednisolone (weeks 1 to 5) and days 1, 3 and 5 after tapering was completed. Before the 1-week course of prednisolone (weeks 14 (standard risk), 28 (high risk) or 37 (intermediate risk)) and on day 2 after cessation. Before the 3-week course of dexamethasone (weeks 25 to 27 (intermediate risk) or 36 or 38 (high risk)), before tapering, and on days 1, 3 and 7 after tapering was completed. The tests were performed during treatment for ALL. Cut-off limits defined by original studies: low-dose ACTH test: normal response > 500 nmol/L. |

| Notes | Based on additional information provided by the authors, 5 patients were lost to follow-up. Length of follow-up after glucocorticoid therapy: varied. Risk factors were not evaluated. |

ACTH: adrenocorticotropic hormone  
AIEOP: Associazione Italiana Ematologia e Oncologia  
ALL: acute lymphoblastic leukaemia  
CRH: corticotrophin-releasing hormone  
HPA: hypothalamic-pituitary-adrenal  
iv: intravenous  
NOPHO: Nordic Society of Pediatric Hematology and Oncology  
RCT: randomised controlled trial

### Supplemental Table 3.2 Characteristics of excluded studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Reason for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bessho 1984</td>
<td>HPA axis was not examined</td>
</tr>
</tbody>
</table>
| Birkebaek 1998    | No data on intervention (doses, duration, methods of cessation of glucocorticoid therapy)  
No data on cut-off limits of HPA axis function tests  
No accurate data on follow-up period  
Data on patients without cranial irradiation were not reported separately |
| Felder-Puig 2007  | HPA axis function was only examined during glucocorticoid treatment                    |
| Lightner 1981     | Patients received cranial irradiation                                                  |
| Pawlaczysz 1993    | Patients received cranial irradiation                                                  |
| Silva 2006        | Double publication of Cunha et al.                                                    |

HPA: hypothalamic-pituitary-adrenal
Chapter 4

Adrenal insufficiency during induction and maintenance treatment for childhood acute lymphoblastic leukemia

M. Suzanne Gordijn, Reinoud J.B.J. Gemke, A.S. Paul van Trotsenburg, Marc B. Bierings, Cor van den Bos, Joost Rotteveel, Gertjan J.L. Kaspers

Submitted
Abstract

Background
Glucocorticoids play a major role in the treatment of acute lymphoblastic leukemia (ALL), but may suppress the hypothalamic-pituitary-adrenal (HPA) axis. Adrenal insufficiency may lead to life-threatening complications. The duration of adrenal insufficiency following high doses of glucocorticoids is not well defined. This study aimed to elucidate the occurrence and duration of adrenal insufficiency during induction and maintenance treatment for childhood ALL.

Methods
Adrenal function was assessed in 27 children between 1 and 19 years of age with ALL by the low dose ACTH test at different time points after cessation of glucocorticoid therapy.

Results
The median time to recovery from adrenal insufficiency after induction therapy was 28 (95% CI 12.0-44.0) days. The majority of patients recovered within five weeks. The estimated proportion with no recovery of adrenal function at the end of the follow-up period after induction therapy of 96 days was 29 (95% CI 10.1-47.7) percent. Interestingly, time to recovery from adrenal insufficiency was longer in boys compared to girls (P =0.009). All patients without recovery of adrenal function after induction therapy showed adrenal insufficiency during maintenance therapy.

Conclusion
Adrenal insufficiency frequently occurs in childhood ALL. Ongoing adrenal insufficiency after induction therapy is predictive for adrenal insufficiency during maintenance therapy. We recommend glucocorticoid coverage during periods of stress in the first five weeks after induction therapy in all patients. Thereafter, we recommend adrenal function tests to determine which patients suffer from prolonged adrenal insufficiency and require additional glucocorticoid treatment during stress.
Introduction

Glucocorticoids, like predniso(lo)ne and dexamethasone, play a major role in the treatment of childhood acute lymphoblastic leukemia (ALL). However, supraphysiological doses may suppress the hypothalamic-pituitary-adrenal (HPA) axis. The resulting impaired cortisol response may lead to life-threatening hypotension and/or hypoglycemia, impaired inflammatory response, and inadequate host defense against infections.

Adrenal insufficiency caused by high doses of glucocorticoids could be a contributing factor to the occurrence and severity of infections during treatment of childhood ALL. Since prednisone has been substituted for the more potent dexamethasone several studies have reported an increase of septic and lethal infections.

Patients with adrenal insufficiency may benefit from glucocorticoid coverage during periods of stress to reduce the risk of life-threatening complications. However, due to the inconsistent picture of adrenal insufficiency during treatment for childhood ALL, routine glucocorticoid coverage has not become standard practice.

Relatively little information is available regarding the incidence and duration of adrenal insufficiency during treatment for childhood ALL. Previous studies have applied adrenal function tests which are regarded inferior compared to the low dose 1 μg ACTH test. Furthermore, the majority of these studies performed adrenal function tests following induction therapy but not during maintenance therapy and the follow-up period was relatively short.

This prospective multi-center study aimed to further elucidate the occurrence and duration of adrenal insufficiency during induction and maintenance therapy in children treated for ALL.

Patients and methods

Patients

All children between 1 and 19 years of age with newly diagnosed ALL treated according to the Dutch Childhood Oncology Group (DCOG) ALL-10 or ALL-11 protocol in VU University Medical Center Amsterdam, Wilhelmina Children’s Hospital/University Medical Center Utrecht or Emma Children’s Hospital/Academic Medical Center in the Netherlands between December 2010 and February 2013 could be included. Exclusion criteria were glucocorticoid therapy for other reasons than ALL since one year before diagnosis and a history of HPA axis related diseases. This study protocol was approved by the Institutional Review Board of the VU University Medical Center and written informed consent was obtained from the children (≥12 years of age) and their caregivers.
Treatment
Both ALL-10 and ALL-11 protocol consisted of an induction phase comprising prednisolone (PRED-4w) (60 mg/m²/day, day 1-28 with tapering over nine days). Induction therapy also included vincristine, daunorubicin, asparaginase, cyclophosphamide, cytosine arabinoside, 6-mercaptopurine and intrathecal methotrexate and methotrexate/cytosine arabinoside/diadreson-F aquosum. This induction phase was followed by treatment with 6-mercaptopurine, high dose methotrexate, folinic acid and intrathecal methotrexate/cytosine arabinoside/diadreson-F aquosum. Based on central nervous system and testicular involvement, cytogenetics, response to initial therapy, induction failure and the minimal residual disease, patients were classified in the Standard Risk Group (SRG), the Medium Risk Group (MRG) or the High Risk Group (HRG). SRG patients received an additional course, protocol IV, comprising dexamethasone (DEX-2w) (10 mg/m²/day, day 1-15 with tapering over nine days), vincristine and PEG asparaginase, followed by maintenance treatment comprising 6-mercaptopurine and methotrexate. MRG patients received 84 weeks of maintenance treatment comprising cyclic courses of dexamethasone (DEX-5d) (6 mg/m²/day, 5 days every 3 weeks, week 1-82) and vincristine, doxorubicine, methotrexate, PEG asparaginase, 6-mercaptopurine and intrathecal methotrexate/cytosine arabinoside/diadreson-F aquosum. We did not perform follow-up adrenal function tests in HRG patients.

Adrenal function test
Adrenal function was assessed by the intravenous low-dose (1 μg) ACTH stimulation test (LD ACTH test) using tetracosactrin (Synacthen, Novartis), which is considered as the most reliable test to detect subtle degrees of adrenal atrophy. The first LD ACTH test was performed approximately one week after the last glucocorticoid (tapering) dose. A normal response was defined as a stimulated serum cortisol ≥550 nmol/l. In case of suppression, the LD ACTH test was repeated every two to three weeks, until normal stimulated cortisol levels were observed. For practical reasons, e.g. rescheduling of hospital visits, catheter malfunction or glucocorticoid use for other reasons than standard treatment (e.g. antiemetics or stress doses), some of the scheduled tests were postponed. In case of adrenal insufficiency following DEX-2w in SRG patients, the test interval was extended to minimize patient’s burden. In MRG patients, a single LD ACTH test was performed after the first DEX-5d according to maintenance therapy and halfway maintenance treatment (week 43 of 84 weeks in total). All tests were performed between 0800 h and 1000 h in the hospital or at the outpatient clinic. A blood sample was obtained before the administration of 1 μg/1.73 m² tetracosactrin to determine basal morning value (t=0), and after 30 (t=30) and 60 (t=60) minutes.

Cortisol measurements
The blood samples were analyzed according to standard procedures at the laboratories of the participating centers. In the Department of Clinical Chemistry, VU University Medical Center,
Amsterdam, serum cortisol was analyzed using a competitive immunoassay (Advia Centaur, Siemens Diagnostics, Deerfield, IL). Intra- and interassay coefficients of variation (c.v.) were 3 and 6%, respectively, and the detection limit was 30 nmol/l (500 nmol/l = 18.1 μg/dl). In the Laboratory of Endocrinology and Radiochemistry, department of Clinical Chemistry Academic Medical Centre, University of Amsterdam, serum cortisol was assayed using a chemiluminescent immunoassay on an Immulite system (Siemens Healthcare Diagnostics B.V., Breda, The Netherlands). For cortisol the interassay c.v. were 8.3% at 120 nmol/L and 7.5% at 400 nmol/L with a detection limit of 50 nmol/L and the intra assay c.v. were 5.5% at 120 nmol/L and 6.1% at 400 nmol/L. In the Department of Clinical Chemistry and Haematology, University Medical Center Utrecht, serum cortisol was determined by the Beckman Coulter DXi. Intra-assay c.v. were 6.7% at 160 nmol/L, 5% at 660 nmol/L and 5% at 1100 nmol/L and interassay c.v. were 8-9% at 70 nmol/L, 5-6% at 450 nmol/L and 5-5.4% at 800 nmol/L. The lower limit of quantitation was 20 nmol/L.

Glucocorticoid coverage

Patients with adrenal insufficiency demonstrated by the LD ACTH test received hydrocortisone coverage (30-50 mg/m²/day divided into four doses) during periods of stress, until recovery of the adrenal function. Adrenal function tests following hydrocortisone stress doses with a duration of seven days or longer or stress doses which were started within seven days after cessation of glucocorticoid therapy according to ALL treatment, were performed at least seven days after cessation of the stress dose therapy.

Statistical analysis

Within-group differences were assessed by t-tests for continuous variables, and by chi-square test for categorical variables. Time to recovery from adrenal insufficiency and its 95% confidence interval (95% CI) was estimated with the Kaplan-Meier method; between group comparisons were performed using the log-rank test. In case of no recovery of the adrenal function, survival data were censored at last follow-up. All statistical analyses were performed with SPSS (Version 20, IBM SPSS Statistics). A two-sided p value of <0.05 was accepted as statistically significant.

Results

Demographics

Twenty-nine patients were included in this study. Another eight patients declined participation. As our group simultaneously performed a study regarding adrenal function in lymphoma patients, one patient with gamma delta T-cell lymphoma who received treatment according to the ALL-10 induction protocol was included in this study as well. In two of the 29 included
children the LD ACTH test could not be performed due to severe illness and early death and they were excluded from analysis. Sociodemographic and treatment characteristics of participants and non-participants are shown in Table 4.1. In comparison to participants, non-participants were treated more frequently according to the ALL-11 treatment protocol, however, both ALL treatment protocols did not differ in glucocorticoid dosing schedule.

### Table 4.1 Sociodemographic and treatment characteristics of participants and non-participants

<table>
<thead>
<tr>
<th></th>
<th>Participants (N = 27)</th>
<th>Non-participants (N = 8)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years: median (IQR)</td>
<td>9.0 (4.0-13.0)</td>
<td>9.0 (5.0-13.5)</td>
<td>0.549</td>
</tr>
<tr>
<td>Sex, boys (%)</td>
<td>44%</td>
<td>25%</td>
<td>0.431</td>
</tr>
<tr>
<td><strong>Treatment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALL-10 protocol: N</td>
<td>23</td>
<td>3</td>
<td>0.015</td>
</tr>
<tr>
<td><strong>Risk Group</strong></td>
<td></td>
<td></td>
<td>0.919</td>
</tr>
<tr>
<td>Standard Risk: N</td>
<td>8</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Medium Risk: N</td>
<td>13</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>High Risk: N</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuation with other treatment protocol: N</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALL-11 protocol: N</td>
<td>4</td>
<td>5</td>
<td>0.015</td>
</tr>
<tr>
<td><strong>Risk Group</strong></td>
<td></td>
<td></td>
<td>0.333</td>
</tr>
<tr>
<td>Standard Risk: N</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium Risk: N</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>High Risk: N</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Unknown yet: N</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IQR = 25% and 75% interquartile range
ALL = acute lymphoblastic leukemia

**Adrenal function in ALL patients**

The first LD ACTH test was conducted within 25 days after discontinuation of PRED-4w in all patients. At the first test, 15 (56%) of the 27 children had insufficient responses, with peak cortisol levels at either 30 or 60 minutes <550 nmol/L (Figure 4.1). Six of these 15 patients recovered from adrenal insufficiency at a later time point, one patient was lost to follow up and the remaining eight patients still had an insufficient cortisol response by the end of their follow-up.

As shown in Figure 4.2, the majority of patients recovered within five weeks after cessation of PRED-4w, whilst the probability of adrenal recovery after five weeks was low. The median time to recovery from adrenal insufficiency was 28 (95% CI 12.0 - 44.0) days. The estimated
Figure 4.1 Peak cortisol levels during LD ACTH tests at different time points after PRED-4w (induction therapy) in each patient. The horizontal line indicates the threshold of the normal cortisol response. + represents the patients with a sufficient cortisol response during the first adrenal function test after induction therapy, ◊ represents the patients that recovered from adrenal insufficiency at a later time point, and ○ represents the patients with an insufficient cortisol response during the adrenal function test by the end of their follow-up.

Figure 4.2 Kaplan-Meier survival estimate of recovery of adrenal function after PRED-4w (induction therapy).
proportion without recovery of adrenal function at the end of the follow-up period after induction therapy of 96 days was 29 (95% CI 10.1-47.7) percent.

Time to recovery from adrenal insufficiency was longer in boys in comparison to girls. Median time to recovery could not be estimated for boys because only a small fraction recovered during follow-up and median time to recovery for girls was 11 days (95% CI 7.2-14.8), \( P = 0.009 \) (Figure 4.3). The estimated proportion that recovered at the end of follow-up was 44.4 (95% CI 15.2-73.6) percent for boys compared to 91.1 (75.0-100) percent for girls. Because of the potential modifying effect of estrogens on adrenal function, we also performed the analyses by excluding the three postmenarcheal female patients. These analyses revealed the same difference in time to recovery from adrenal insufficiency between boys and girls (\( P = 0.017 \)).

Age was not different between boys and girls (median: 11.5 (IQR: 5-14.8) vs 6.0 (IQR: 3.0-12.0) years, \( P = 0.11 \)). Age, treatment protocol and risk group stratification were not associated with the recovery time of adrenal function (data not shown).

As shown in Figure 4.4, longitudinal data were obtained in six SRG patients. Two other SRG patients declined further participation. Two SRG patients with sufficient cortisol responses during the first adrenal function test after induction therapy, showed sufficient cortisol responses at the first LD ACTH test after DEX-2w. However, two other SRG patients with initially sufficient adrenal responses showed impaired responses after DEX-2w. In addition, one of the two SRG patients with prolonged adrenal insufficiency after induction therapy showed prolonged adrenal insufficiency after DEX-2w, whilst the other patient showed a sufficient cortisol response at the first adrenal function test after cessation of DEX-2w.
As shown in Figure 4.5, data on adrenal function after the first DEX-5d according to maintenance therapy was collected in 14 MRG patients. One MRG patient declined further participation and one MRG patient did not yet complete her first DEX-5d. In 11 of the 14 patients data on adrenal function halfway maintenance therapy were obtained, whilst the other three patients did not yet complete the first 40 weeks of maintenance treatment. Six of the ten MRG patients with a recovered adrenal function after induction therapy who underwent follow-up adrenal function tests, showed sufficient cortisol responses after the first DEX-5d. Three of them were also tested halfway maintenance therapy and all three patients showed sufficient cortisol responses. Four of the ten MRG patients with a recovered adrenal function after induction therapy, showed insufficient cortisol responses after the first DEX-5d. Four of them were also tested halfway maintenance therapy; one patient showed insufficient cortisol responses, whilst the other three patients were recovered from adrenal insufficiency at that time point. All four MRG patients with no recovery of adrenal function after induction therapy showed adrenal insufficiency after the first DEX-5d and they all showed adrenal insufficiency halfway maintenance therapy.

Figure 4.4 Peak cortisol levels during LD ACTH tests at different time points after PRED-4w (induction therapy) in each SRG patient. The vertical broken lines represent the administration of the additional DEX-2w course (protocol IV). The horizontal line indicates the threshold of the normal cortisol response. ○ represents the patients that declined further participation, + represents the patients with recovered adrenal function after induction therapy who showed sufficient cortisol responses after DEX-2w, △ represents the patients with recovered adrenal function after induction therapy who showed impaired responses after DEX-2w, □ represents a patient with prolonged adrenal insufficiency after induction therapy who showed prolonged adrenal insufficiency after DEX-2w, and ◊ represents a patient with prolonged adrenal insufficiency after induction therapy who showed a sufficient cortisol response after DEX-2w.
This study demonstrates that adrenal insufficiency frequently occurs after induction therapy with prednisolone for childhood ALL. The majority of patients recovered within five weeks. However, one third of the patients suffered from prolonged adrenal insufficiency. In our study, all patients with ongoing adrenal insufficiency after induction therapy were also insufficient after the second glucocorticoid course as part of their maintenance therapy. In line with previous studies, the recovery time of adrenal function shows considerable inter-individual variation. This might be due to differences in genetic susceptibility for developing adrenal insufficiency during glucocorticoid therapy. An interesting finding of the current study is that the duration of adrenal insufficiency was longer in boys in comparison to girls, irrespective of menarcheal state. A possible explanation for this finding is interference of sex hormones with the HPA axis. Animal studies have demonstrated that estrogens enhance the HPA axis responsiveness whilst androgens inhibit the glucocorticoid response. Unfortunately, literature addressing sex hormone influences upon the HPA axis in humans is scarce. Another potential explanation for the difference in duration of adrenal insuffi-
ciency between boys and girls might be gender specific effects of glucocorticoid receptor polymorphisms on adrenal function. Because this is the first study to report on prolonged duration of adrenal insufficiency in boys, this finding needs to be interpreted with caution until it can be confirmed in larger studies. Furthermore, this study shows that, in contrast to the conclusion of Spiegel et al., patients with a proven adequate adrenal function after induction therapy may develop adrenal insufficiency after a short-term (DEX-5d) glucocorticoid course.

Several limitations of our study need to be mentioned. First, due to the two to three week test intervals, we could have missed earlier recovery of adrenal function between the intervals. This may have resulted in an overestimation of the time to recovery from adrenal insufficiency. In addition, the threshold of a normal cortisol response to the LD ACTH test is arbitrary. Some studies have used a cut-off value of 500 nmol/l. As we used the rather stringent cut-off level of 550 nmol/l, the number of patients with adrenal insufficiency might have been overestimated. Furthermore, previous studies demonstrated adrenal insufficiency after high-dose fluconazole therapy. As fluconazole was administered as fungal prophylaxis during induction therapy and during the first 19 weeks of MRG maintenance therapy in one of the three participating centers, this might have prolonged the duration of adrenal insufficiency. However, there was no significant difference in the duration of adrenal insufficiency between the participating centers. Further, the number of patients per risk group was relatively small, thus generalization of the current results regarding adrenal function during maintenance therapy should be made with caution. However, because of the lack of studies on adrenal function during maintenance therapy for childhood ALL, our results provide important information of practical relevance. Additionally, the detected difference in duration of adrenal insufficiency between boys and girls was an unexpected finding. Tanner stage and sex hormone levels were not collected systematically in this study, hence further studies including these measurements should be performed to elucidate this observation. Finally, different cortisol assays have been used. However, no significant differences in the occurrence and duration of adrenal insufficiency between the three participating centers have been detected.

Based on the results of this study, we recommend glucocorticoid coverage (e.g. hydrocortisone) during periods of stress in the first five weeks after cessation of induction therapy in all children treated for ALL. Adrenal function tests should be performed thereafter to determine which patients suffer from prolonged adrenal insufficiency and to determine in which patients stress dose therapy can be discontinued. Special attention should be paid to patients with prolonged adrenal insufficiency after induction therapy as they may suffer from extended adrenal insufficiency during maintenance therapy.
References

Adrenal insufficiency during childhood ALL


Chapter 5

Adrenal insufficiency during treatment for childhood lymphoma

M. Suzanne Gordijn, Reinoud J.B.J. Gemke, A.S. Paul van Trotsenburg, Marc B. Bierings, Cor van den Bos, Joost Rotteveel, Gertjan J.L Kaspers

Submitted
Summary

Glucocorticoids play an important role in the treatment of lymphoma, but may suppress the hypothalamic-pituitary-adrenal axis. Adrenal insufficiency may lead to life-threatening complications. In this study adrenal function was assessed in 17 children with (non-)Hodgkin lymphoma at different time points during treatment. Seven out of 13 patients demonstrated adrenal insufficiency halfway treatment and 9/16 patients demonstrated adrenal insufficiency after the final glucocorticoid course.

Adrenal insufficiency frequently occurs during treatment for childhood lymphoma. We recommend glucocorticoid coverage during periods of stress in between cyclic glucocorticoid courses and in the first weeks after glucocorticoid therapy in all children treated for lymphoma.
Introduction

Glucocorticoids, like prednisolone and dexamethasone, are routinely used in the treatment of childhood lymphoma, but supraphysiological doses may suppress the hypothalamic-pituitary-adrenal (HPA) axis. Adrenal insufficiency may result in life-threatening hypoglycemia, hypotension and/or inadequate host defense against infections. Patients with adrenal insufficiency may benefit from glucocorticoid coverage during periods of stress to prevent life-threatening complications. Due to the lack of information regarding adrenal insufficiency during treatment for childhood lymphoma, routine glucocorticoid coverage during stress is not standard practice. The available literature regarding adrenal function during and after treatment for lymphoma is scarce and only comprises adult lymphoma patients. In addition, previous studies applied adrenal function tests that are now obsolete and did not perform follow-up tests in patients with adrenal insufficiency. The aim of this prospective multi-center study was to examine the occurrence and duration of adrenal insufficiency during treatment for childhood lymphoma.

Materials and methods

Patients

Children between 1 and 19 years of age with newly diagnosed Hodgkin lymphoma or non-Hodgkin lymphoma treated in VU University Medical Center Amsterdam, Wilhelmina Children’s Hospital / University Medical Center Utrecht or Emma Children’s Hospital / Academic Medical Center Amsterdam in the Netherlands between December 2010 and February 2013 were included. Exclusion criteria were glucocorticoid therapy for other reasons than treatment for lymphoma since one year before diagnosis and a history of HPA axis related disease. This study was approved by our Institutional Review Board and written informed consent was obtained from the children (≥12 years of age) and their caregivers.

Treatment

Six patients with Hodgkin lymphoma were treated according to the EuroNet-Paediatric Hodgkin’s Lymphoma (PHL)-C1 protocol, comprising two OEPA courses (OEPA1 and OEPA2) with prednisolone (60 mg/m²/day, day 1-5). For two patients this was followed by two additional courses, COPDAC courses (COPDAC1 and COPDAC2) for one patient and COPP courses (COPP1 and COPP2) for the other patient, both courses comprising prednisolone (40 mg/m²/day, day 1-15). One patient with Hodgkin lymphoma was treated according to the BEACOPP protocol, comprising six cycles (BEACOPP1-6) of prednisolone (40 mg/m²/day, day 1-14).
Eight patients with non-Hodgkin lymphoma were treated according to the DCOG B-NHL/B-ALL 2008 protocol. All patients initially received a COP course comprising prednisolone (60 mg/m²/day, day 1-7), followed by two COPADM courses (COPADM1 and COPADM2) comprising prednisolone (60 mg/m²/day, day 1-5 with tapering over three days). This was followed by two CYM courses without glucocorticoids for all patients. One patient received four additional OK courses (OK1-4). OK1 and OK3 comprised prednisolone (60 mg/m²/day, day 1-5 with tapering over three days).

Two patients with non-Hodgkin lymphoma were treated according to the ALCL 99 protocol. They received a pre-phase course comprising dexamethasone (5 mg/m²/day on day 1 and 2 and 10 mg/m²/day on day 3-5), followed by three AM courses (AM1-3) and three BM courses (BM1-3), each course comprising dexamethasone (10 mg/m²/day, day 1-5). See also Table 5.1.

### Table 5.1 Treatment schedules of the study participants

<table>
<thead>
<tr>
<th></th>
<th>Hodgkin lymphoma patients</th>
<th>Non-Hodgkin lymphoma patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=4</td>
<td>n=1</td>
<td>n=1</td>
</tr>
<tr>
<td>OEPA1</td>
<td>OEPA1</td>
<td>BEACOPP1</td>
</tr>
<tr>
<td>OEPA2</td>
<td>OEPA2</td>
<td>BEACOPP2</td>
</tr>
<tr>
<td>COPDAC1</td>
<td>COPP1</td>
<td>BEACOPP3</td>
</tr>
<tr>
<td>COPDAC2</td>
<td>COPP2</td>
<td>BEACOPP4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CYM2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OK1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OK2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OK3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OK4</td>
</tr>
</tbody>
</table>

OEPA, prednisolone, vincristine, doxorubicin and etoposide

COPDAC, prednisolone, dacarbazine, vincristine and cyclophosphamide

COP, prednisolone, procarbazine, vincristine and cyclophosphamide

BEACOPP, prednisolone, bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine and procarbazine

COP, prednisolone, vincristine, cyclophosphamide and intrathecal therapy

COPADM, prednisolone, vincristine, methotrexate, folinic acid, cyclophosphamide, doxorubicin and intrathecal therapy

CYM, methotrexate, folinic acid, cytosine arabinoside and intrathecal therapy

Pre-phase, dexamethasone, cyclophosphamide and intrathecal therapy

AM, dexamethasone, methotrexate, folinic acid, ifosfamide, cytosine arabinoside and etoposide

BM, dexamethasone, methotrexate, folinic acid, cyclophosphamide and doxorubicin

OK1, prednisolone, vincristine, methotrexate, folinic acid, cyclophosphamide, doxorubicin and intrathecal therapy

OK2 & OK4, cytosine arabinoside and etoposide

OK3, prednisolone, vincristine, cyclophosphamide, doxorubicin
Adrenal function test

Adrenal function was assessed by the intravenous low-dose (1 μg/1.73 m²) ACTH stimulation test (LD ACTH test) using tetracosactrin (Synacthen, Novartis).6-8 A LD ACTH tests was performed at least one week after discontinuation of a glucocorticoid course approximately halfway during treatment. Additionally, a LD ACTH test was performed at least one week after cessation of the final glucocorticoid course. In case of adrenal insufficiency, this test was repeated at varying intervals until normal stimulated cortisol levels were observed or until the end of follow-up. For practical reasons, e.g. rescheduling of hospital visits, catheter malfunction or glucocorticoid use for other reasons than standard treatment (e.g. anti-emetics or stress doses), some of the scheduled tests were canceled or postponed. All tests were performed between 8 and 10 am. A blood sample was obtained before the administration of tetracosactrin (t=0), and after 30 and 60 minutes. A normal response was defined as a stimulated serum cortisol ≥550 nmol/l. Blood samples were analyzed according to standard procedures at the laboratories of the participating centers.

Glucocorticoid coverage

Patients with adrenal insufficiency demonstrated by the LD ACTH test received hydrocortisone coverage (30-50 mg/m²/day divided into four doses) during periods of stress, until recovery of the adrenal function was demonstrated by LD ACTH test.

Statistical analysis

Within-group differences were assessed by the Mann-Whitney U-test for continuous variables, and by chi-square test for categorical variables. All statistical analyses were performed with SPSS (Version 20, IBM SPSS Statistics). A two-sided p value of <0.05 was accepted as statistically significant.

Results

A total of 17 patients (10 boys and 7 girls), with a median age of 13.0 (interquartile range 9.5-14.5) years, were included in the current study.

In total, 44 LD ACTH tests were performed. In 13 patients, a LD ACTH test was performed approximately halfway during treatment (i.e. after OEP1, BEACOP3, COPADM1, AM3) (Figure 5.1). Six patients showed a sufficient cortisol response, with peak cortisol levels at 30 or 60 minutes of ≥550 nmol/L, whilst seven patients suffered from adrenal insufficiency.
In 16 patients, LD ACTH tests were performed after the final glucocorticoid course (Figure 5.2). One of the four patients who finished their treatment after two OEPA courses showed a sufficient cortisol response. This patient also showed adrenal sufficiency after OEPA1. The other three patients showed adrenal insufficiency after their final OEPA courses. One of them was tested after OEPA1 and showed adrenal insufficiency during that time point as well. One of the three suppressed patients recovered from adrenal insufficiency within 74 days, whilst the other two patients did not recover at the end of their follow-up period of 14 and 18 days, respectively.

Two patients finished their treatment after two additional COPDAC or COPP courses. Both patients showed sufficient cortisol responses. Both patients were tested after cessation of COPDAC1 and OEPA1, respectively, and showed adrenal sufficiency during those time points as well.

One patient finished her treatment after six BEACOPP courses. Although she demonstrated adrenal insufficiency after BEACOPP3, she showed an adequate cortisol response 26 days after her final BEACOPP course.

Six patients finished their glucocorticoid treatment after two COPADM courses. One of these patients demonstrated a sufficient cortisol response. The other five patients showed adrenal insufficiency after their final COPADM course. Four of them were tested after COPADM1: one patient showed a sufficient cortisol response, whilst the other three patients demonstrated adrenal insufficiency. Two of the patients with adrenal insufficiency recovered

\[\text{Figure 5.1 Peak cortisol levels during a LD ACTH test approximately halfway during treatment for childhood lymphoma. The horizontal line indicates the threshold of the normal cortisol response.}\]
within 93 and 240 days, respectively. The other three patients did not demonstrate recovery of adrenal function at the end of their follow-up of 50, 69 and 230 days, respectively. One patient finished his glucocorticoid treatment after the third OK course. He was tested 56 days after cessation of OK3 and showed a sufficient cortisol response. This patient also showed adrenal sufficiency after COPADM1 and COPADM2.

Two patients finished their treatment after three AM and BM courses. One patient, who demonstrated adrenal insufficiency after AM3, showed a sufficient cortisol response 32 days after his final glucocorticoid course. The other patient, who showed adrenal insufficiency after AM3, demonstrated an inadequate cortisol response after his final glucocorticoid course and did not recover at the end of his follow-up period of 300 days.

**Discussion**

This is the first study that reports on adrenal function during treatment for childhood lymphoma. Adrenal insufficiency occurred frequently: seven out of 13 patients demonstrated adrenal insufficiency halfway during treatment and nine out of 16 patients demonstrated adrenal insufficiency after cessation of the final glucocorticoid course. The majority of patients with adrenal insufficiency in between consecutive glucocorticoid courses demonstrated prolonged adrenal insufficiency after completing treatment for childhood lymphoma. In line with literature regarding adrenal insufficiency during treatment for childhood acute lymphoblastic leukemia, the recovery time of adrenal function shows considerable inter-individual
variation. Possible underlying genetic differences would be an interesting topic for future research.

Limitations of this study include the small number of subjects. However, because of the lack of literature on adrenal function during treatment for childhood lymphoma, our results provide important information. Furthermore, due to the test intervals, we could have missed earlier recovery of adrenal function between the test intervals. This may have resulted in an overestimation of the time to recovery from adrenal insufficiency. In addition, after their final glucocorticoid course (COPADM2), six non-Hodgkin patients treated according to the DCOG B-NHL/B-ALL 2008 protocol received two additional CYM courses. Some of these patients received dexamethasone (5 mg/m²/day, for several days) as antiemetic therapy during these courses. These substantial doses of dexamethasone may have extended the time to recovery from adrenal insufficiency.

In conclusion, adrenal insufficiency frequently occurs during treatment for childhood lymphoma. Until larger studies have confirmed our findings, we recommend glucocorticoid coverage during periods of stress in between consecutive glucocorticoid courses and in the first weeks after cessation of glucocorticoid therapy for childhood lymphoma. An adrenal function test is recommended, for example one month after cessation of glucocorticoids, to identify patients who suffer from prolonged adrenal insufficiency and to determine in which patients stress dose therapy can be discontinued.
Adrenal insufficiency during childhood lymphoma

References

Chapter 6

Adrenal insufficiency during treatment for childhood acute lymphoblastic leukemia is associated with glucocorticoid receptor polymorphisms ER22/23EK and BclI

M. Suzanne Gordijn, Ruben D. de Ruiter, Reinoud J.B.J. Gemke, Cor van den Bos, Marc B. Bierings, Joost Rotteveel, Jan W. Koper, Elisabeth F.C. van Rossum, Gertjan J.L. Kaspers
Abstract

Background
Glucocorticoids are an important component of treatment for acute lymphoblastic leukemia (ALL), but may suppress the hypothalamic-pituitary-adrenal (HPA) axis. The duration of adrenal insufficiency shows considerable inter-individual variation. Several polymorphisms in the glucocorticoid receptor (GR) gene have been described to alter glucocorticoid sensitivity. We explored whether these genetic variations within the GR gene are related to the duration of adrenal insufficiency after cessation of high-dose glucocorticoid therapy in children treated for ALL.

Methods
We genotyped four common GR polymorphisms (ER22/23EK, GR-β, N363S and BclI) in 25 children between 1 and 19 years of age with ALL. The duration of adrenal insufficiency, which was assessed by the low dose ACTH test at different time points after cessation of a four-week induction course with high-dose prednisolone therapy, was compared between the different GR genotypes.

Results
The duration of adrenal insufficiency in carriers of the ER22/23EK polymorphism was significantly shorter in comparison to noncarriers (median 7.0 (range 6.0-8.0) vs 28.0 (range 6.0-96.0) days, \( P = 0.03 \)). Adrenal insufficiency in homozygous BclI carriers lasted significantly longer than in patients that were heterozygous or noncarrier of the BclI variant (median 70.5 (range 26.0-93.0) vs 22.0 (range 6.0-96.0) days, \( P = 0.04 \)).

Conclusion
The GR polymorphisms ER22/23EK and BclI affect the duration of adrenal insufficiency after high-dose glucocorticoid therapy in children treated for ALL.
Adrenal insufficiency and GR polymorphisms

Introduction

Acute lymphoblastic leukemia (ALL) is the most frequent type of malignancy in children. Due to advances in the treatment of childhood ALL, survival rates have dramatically improved over time. Currently, the five-year event free survival has reached about 90%.\textsuperscript{1,2} Therefore, treatment related morbidity has become increasingly important.

Children with ALL receive treatment with high-dose glucocorticoids, which results in apoptosis of malignant lymphoblasts.\textsuperscript{3,4} However, supraphysiologic doses suppress the secretion of adrenocorticotropic hormone (ACTH) by the pituitary as well as the corticotrophin releasing hormone (CRH) by the hypothalamus because of a negative feedback signal induced by glucocorticoids binding to the glucocorticoid receptor (GR). This can result in adrenal cortex atrophy.\textsuperscript{5,6} The resulting impaired cortisol response may lead to life-threatening hypoglycemia and/or hypotension, impaired inflammatory response, and inadequate host defense against infections.\textsuperscript{7,8}

A recent study performed by our study group demonstrated that adrenal insufficiency frequently occurs after discontinuation of a four-week induction course with high-dose prednisolone for childhood ALL (see Chapter 4, submitted). The median time to recovery from adrenal insufficiency after induction therapy was 28 (95% CI 12.0-44.0) days. However, approximately one third of the patients suffered from prolonged (at least 96 days) adrenal insufficiency. In line with previous studies\textsuperscript{6,9-11}, the recovery time of adrenal function after cessation of glucocorticoids showed considerable inter-individual variation. A possible explanation for this variation in adrenal insufficiency after cessation of glucocorticoid therapy is differences in genetic susceptibility of individuals for developing adrenal insufficiency during glucocorticoid therapy.

The effects of glucocorticoids are mediated by the GR. A large number of polymorphisms in the GR gene (NR3C1), which is located on chromosome 5, are known. A few of these polymorphisms have been associated with an altered glucocorticoid sensitivity in vivo. For the ER22/23EK polymorphism (rs6189 and rs6190), a GAG AGG to GAA AAG sequence alteration located on exon 2 resulting in a protein alteration from glutamic acid-arginine (E-R) to glutamic acid-lysine (E-K), a relative glucocorticoid resistance has been demonstrated.\textsuperscript{12,13} In addition, the GR-β polymorphism (rs6198), an A to G substitution in the “ATTTA motif” (ATTTA-to-GTTTA) located on exon 9β, has also been associated with glucocorticoid resistance.\textsuperscript{13} In contrast, the N363S polymorphism (rs56149945, previously rs6195), an AAT to AGT nucleotide change located in codon 363 of exon 2 yielding a protein alteration from asparagine (N) to serine (S), has been reported to be associated with an enhanced sensitivity to glucocorticoids.\textsuperscript{13} The BclI polymorphism (rs41423247), a C to G substitution 646 nucleotides downstream from exon 2, has also been associated with increased glucocorticoid sensitivity.\textsuperscript{13,14} To the best of our knowledge, no literature on the relation between these GR polymorphisms and the duration of adrenal insufficiency after high-dose glucocorticoid therapy in patients is available.
The aim of this pilot study was to explore the relation between the duration of adrenal insufficiency after a four-week induction course with high-dose prednisolone therapy in children treated for ALL and the incidence of GR polymorphisms with previously reported altered sensitivity for glucocorticoids and sufficient prevalence in the population (ER22/23EK, GR-β, N363S and BclI).

Materials and methods

Patients

Children between 1 and 19 years of age with newly diagnosed ALL treated in VU University Medical Center Amsterdam, Wilhelmina Children’s Hospital / University Medical Center Utrecht or Emma Children’s Hospital / Academic Medical Center Amsterdam, in the Netherlands between December 2010 and February 2013 were included. Exclusion criteria were glucocorticoid therapy for other reasons than treatment for ALL during one year before diagnosis and a history of HPA axis related diseases. This study protocol was approved by the Institutional Review Board of the VU University Medical Center and written informed consent was obtained from the children (above 12 years of age) and their caregivers. The study was conducted in accordance with the Declaration of Helsinki.

ALL treatment

All children were treated according to the Dutch Childhood Oncology Group (DCOG) ALL 10 or the consecutive ALL 11 protocol. Both treatment protocols did not differ in glucocorticoid treatment schedule and started with an induction course comprising prednisolone (60 mg/m²/day, day 1-28 with tapering over nine days).15,16 Induction treatment also included the administration of vincristine, daunorubicin, asparaginase, cyclophosphamide, cytosine arabinoside, 6-mercaptopurine, intrathecal methotrexate and intrathecal methotrexate/cytosine arabinoside/diadresin F aquosum. The treatment protocols did not include cranial irradiation.

Adrenal function tests

Adrenal function was assessed by the intravenous low-dose (1 μg) ACTH stimulation test (LD ACTH test) using tetracosactrin (Synacthen, Novartis).17-19 LD ACTH tests were performed approximately one week after the last prednisolone tapering dose according to the induction course. A blood sample was obtained before the administration of 1 μg/1,73 m² tetracosactrin to determine basal morning value (t=0), and after 30 (t=30) and 60 (t=60) minutes. A normal response was defined as a stimulated serum cortisol ≥ 550 nmol/l. In case of suppression, the LD ACTH test was repeated every two to three weeks, until normal stimulated cortisol levels were observed or until the end of follow-up which was 96 days after the last prednisolone
tapering dose according to induction treatment because of the start of a second high-dose glucocorticoid course. All tests were performed between 8 and 10 am in the hospital or at the outpatient clinic.

**Cortisol measurements**

The blood samples were analyzed according to standard procedures at the laboratories of the three participating centers. In the Department of Clinical Chemistry, VU University Medical Center, Amsterdam, the Netherlands, serum cortisol was analyzed using a competitive immunoassay (Advia Centaur, Siemens Diagnostics, Deerfield, IL). Intra- and interassay coefficients of variation (c.v.) are 3 and 6%, respectively, and the detection limit was 30 nmol/l (500 nmol/l = 18.1 μg/dl). In the Laboratory of Endocrinology and Radiochemistry, department of Clinical Chemistry Academic Medical Centre, University of Amsterdam, the Netherlands, serum cortisol was assayed using a chemiluminescent immunoassay on a Immulite system (Siemens Healthcare Diagnostics B.V., Breda, The Netherlands). For cortisol the interassay c.v. were 8.3% at 120 nmol/L and 7.5% at 400 nmol/L with a detection limit of 50 nmol/L and the intra assay c.v. were 5.5% at 120 nmol/L and 6.1% at 400 nmol/L. In the Department of Clinical Chemistry and Haematology, University Medical Center Utrecht, the Netherlands, serum cortisol was determined by the Beckman Coulter Dxi. Intra-assay c.v. were 6.7% at 160 nmol/L, 5% at 660 nmol/L and 5% at 1100 nmol/L and interassay c.v. were 8-9% at 70 nmol/L, 5-6% at 450nmol/L and 5-5.4% at 800 nmol/L. The lower limit of quantitation was 20 nmol/L.

**Genotyping of GR**

DNA material of the included patients was extracted from leukocytes in peripheral venous blood samples taken in hematologic remission using standard techniques. Genotypes were determined by allelic discrimination using a Taqman ABI Prism 7900HT Sequence Detection System. The Assay-by-Design service (http://www.appliedbiosystems.com) was used to set up a Taqman allelic discrimination assay for ER22/23EK, GR-β, N363S and BclI. Primer and probe sequences are available on request. The polymerase chain reaction (PCR) mixture included 5 ng of genomic DNA in a 2 μl volume, essentially as previously described.20

**Statistical analyses**

Data was analyzed using SPSS for Windows (IBM SPSS Statistics, Version 21). Differences in duration of adrenal insufficiency after induction therapy between the different GR genotypes were analyzed by Mann-Whitney U-test. A two-sided p value of <0.05 was considered as statistically significant.
### Table 6.1 Genotype distribution in study sample

<table>
<thead>
<tr>
<th></th>
<th>ER22/23EK</th>
<th>GR-β</th>
<th>N363S</th>
<th>BclI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>noncarrier</td>
<td>carrier</td>
<td>P value</td>
<td>noncarrier</td>
</tr>
<tr>
<td>Participants, N</td>
<td>23</td>
<td>2</td>
<td>0.43</td>
<td>16</td>
</tr>
<tr>
<td>Age, years: median (range)</td>
<td>6.0 (1.0-6.0)</td>
<td>10.5 (5.0-16.0)</td>
<td>0.43</td>
<td>4.5 (1.0-16.0)</td>
</tr>
<tr>
<td>Gender, boys (%)</td>
<td>44</td>
<td>50</td>
<td>1.00</td>
<td>38</td>
</tr>
<tr>
<td>Treatment protocol</td>
<td>1.00</td>
<td>1.00</td>
<td>NA</td>
<td>0.32</td>
</tr>
<tr>
<td>ALL 10</td>
<td>19</td>
<td>2</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>ALL 11</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

* of which two patients are homozygous
** of which four patients are homozygous
NA = not applicable
Results

Twenty-seven patients were included in this study. Two patients in whose DNA sample the genotyping failed were excluded from further analyses. Median age of the remaining 25 patients was 6.0 (range 1.0-16.0) years and 44% were boys. Twenty-one patients were treated according to the ALL-10 protocol, and four patients were treated according to the ALL-11 protocol.

GR genotype distribution of the phenotyped subjects is shown in Table 6.1. These genotype frequencies are similar to those reported previously. There were no significant differences between the genetic groups in terms of sex, age, and treatment protocol, except for BclI carriers, who were significantly younger than noncarriers. Age did not differ between homozygous BclI carriers and the combined group of patients that were not heterozygous or noncarrier of the BclI variant.

Figure 6.1 shows the duration of adrenal insufficiency in childhood ALL patients after induction therapy with high-dose predniso(lo)ne for the different GR genotypes. The duration of adrenal insufficiency in ER22/23EK carriers was shorter than in the noncarriers (median 7.0 (range 6.0-8.0) vs 28.0 (range 6.0-96.0) days, \( P = 0.03 \)). There was no statistically significant difference in duration of adrenal insufficiency between carriers and noncarriers of GR-\( \beta \) (median

![Figure 6.1 Duration of adrenal insufficiency (data expressed as median) after a four-week predniso(lo)ne induction course in ALL patients with different GR genotypes.](image-url)
22.0 (range 6.0-67.0) vs 42.0 (range 6.0-96.0) days, \( P = 0.17 \) or between patients homozygous for GR-\( \beta \) and noncarriers (median 10.0 (range 9.0-11.0) vs 42.0 (range 6.0-96.0) days, \( P = 0.21 \)). The influence of genotype N363S could not be statistically analyzed because there was only one carrier and these data are therefore presented at a descriptive level. The duration of adrenal insufficiency in patients homozygous for \( Bcl \) was longer than in patients who were not homozygous for \( Bcl \) (median 70.5 (range 26.0-93.0) vs 22.0 (range 6.0-96.0) days, \( P = 0.04 \)).

**Discussion**

This pilot study in children with newly diagnosed ALL showed that the duration of adrenal insufficiency after discontinuation of a four-week high-dose prednisolone course is affected by GR genotype, which to the best of our knowledge has never been reported before. The duration of adrenal insufficiency in carriers of the ER22/23EK polymorphism, which is associated with a relative glucocorticoid resistance\(^ {12,13} \), was indeed significantly shorter than in noncarriers. This indicates that the ER22/23EK carriers have a lower susceptibility for developing adrenal insufficiency after high-dose glucocorticoid therapy. This is in line with Van Rossum et al. who found that the feedback regulation of the HPA axis in ER22/23EK carriers was relatively resistant in comparison to noncarriers.\(^ {12} \) In contrast to many polymorphisms in general, the molecular mechanism of this ER22/23EK alteration leading to partial GR resistance has been elucidated. When the ER22/23EK polymorphism is present, there is an increased expression of the less transcriptionally active GR-A translational isoform.\(^ {22} \) The decreased sensitivity of this GR polymorphism to glucocorticoids has been confirmed by *in vitro* studies.\(^ {23} \) No significant difference in duration of adrenal insufficiency was found between carriers and noncarriers of GR-\( \beta \), which however might be due to the relatively small number of participants. The association between the N363S polymorphism and the duration of adrenal insufficiency could not be analyzed because there was only one N363S carrier. Adrenal insufficiency in patients homozygous for \( Bcl \), which is in general associated with increased glucocorticoid sensitivity\(^ {13,14} \), was significantly longer present in comparison to patients that were not homozygous for \( Bcl \). This indicates that patients homozygous for \( Bcl \) have an increased susceptibility for developing prolonged adrenal insufficiency. Our finding is compatible with previous studies that found increased sensitivity to the negative feedback action of glucocorticoids on the HPA axis in carriers of the \( Bcl \) genotype.\(^ {14,21} \)

Based on the present findings it can be hypothesized that the functionally relevant GR polymorphisms are associated with the degree of glucocorticoid-associated toxicity as well. The sparse evidence in the literature is conflicting. Eipel et al. found that hepatotoxicity and glucose metabolism abnormalities were more frequent among N363S carriers.\(^ {24} \) In addition, as the \( Bcl \) genotype has been related to higher body mass index, abdominal obesity and higher systolic blood pressure,\(^ {21} \) Te Winkel et al. studied the relation between this genotype
Adrenal insufficiency and GR polymorphisms

and fat gain in childhood ALL patients but found no relation\textsuperscript{25}. Moreover, Felder-Puig et al. found no correlation between the ER22/23EK, N363S and \textit{BclI} genotypes and adverse psychological reactions in childhood ALL patients.\textsuperscript{26} One might argue that psychological reactions are more likely multifactorial than other side-effects. Previous research regarding the relation between GR polymorphisms and the efficacy of treatment is inconsistent as well. Some studies did not find an association between the ER22/23EK, N363S and \textit{BclI} polymorphisms and treatment outcome\textsuperscript{27,28}, whilst others found an association between reduced survival and the GR-\(\beta\) and \textit{BclI} genotype\textsuperscript{29,30}, or a better five-year event-free survival rate in N363S carriers\textsuperscript{24}. The discrepancy between previous studies concerning the relation between GR polymorphisms and glucocorticoid-related toxicity or treatment outcome might be explained by the complexity of the effects of glucocorticoids, depending on tissue-specific GR isoforms, but also on the cascade of downstream events. Further investigation is warranted to determine whether GR genotype may be one of the future candidates for determining individualized glucocorticoid therapy.

The strength of this study is the unique design which allowed for close monitoring of adrenal function with an extended follow-up period in a clinical population of children with ALL. A clear limitation of this study is the limited number of patients included in this study and the accompanying problem of limited statistical power. Therefore, this study can be considered as a pilot study and validation of the results in larger cohorts is recommended. Nevertheless, despite the small sample size and the relatively low frequencies of the GR polymorphisms in the general population we observed statistically significant and clinically relevant associations between certain GR genotypes and the duration of adrenal insufficiency, which were biologically plausible. This indicates that the duration of adrenal insufficiency during treatment for childhood ALL can, at least partially, be attributed to variants of the GR gene which encodes for the protein through which glucocorticoids exert their action. Of note, these polymorphisms might affect the duration of adrenal insufficiency in other clinical populations that receive high-dose glucocorticoid therapy, e.g. patients with other hematologic malignancies, chronic inflammatory diseases or autoimmune syndromes, as well.

In conclusion, this study showed for the first time that the GR polymorphisms ER22/23EK and \textit{BclI} affect the duration of adrenal insufficiency after high-dose glucocorticoid therapy in children treated for ALL.
References


Chapter 7

Fatigue in children: reliability and validity of the Dutch PedsQL™ Multidimensional Fatigue Scale

M. Suzanne Gordijn, Eline M.P. Cremers, Gertjan J.L. Kaspers, Reinoud J.B.J. Gemke

Qual Life Res 2011;20:1103-8
Abstract

Purpose
The aim of the study is to report on the feasibility, reliability, validity and the norm-references of the Dutch version of the PedsQL™ Multidimensional Fatigue Scale.

Methods
The study participants are 497 parents of children aged 2-18 years and 366 children aged 5-18 years from various day care facilities, elementary schools and a high school who completed the Dutch version of the PedsQL™ Multidimensional Fatigue Scale.

Results
The number of missing items was minimal. All scales showed satisfactory internal consistency reliability, with Cronbach's coefficient alpha exceeding 0.70. Test-retest reliability was good to excellent (ICCs 0.68-0.84) and inter-observer reliability varied from moderate to excellent (ICCs 0.56-0.93) for total scores. Parent/child concordance for total scores was poor to good (ICCs 0.25-0.68). The PedsQL™ Multidimensional Fatigue Scale was able to distinguish between healthy children and children with an impaired health condition.

Conclusions
The Dutch version of the PedsQL™ Multidimensional Fatigue Scale demonstrates an adequate feasibility, reliability and validity in another sociocultural context. With the obtained norm-references, it can be utilized as a tool in the evaluation of fatigue in healthy and chronically ill children aged 2-18 years.
Introduction

Fatigue is a common symptom in pediatric health conditions and is associated with poorer health-related quality of life (HRQOL). The Pediatric Quality of Life Inventory (PedsQL™) Multidimensional Fatigue Scale was designed by Varni and colleagues to measure fatigue in children. The original American version demonstrated adequate reliability and validity. Recently, this questionnaire has been translated into Dutch in accordance with internationally accepted methods.

Substantial cultural differences regarding sleep and fatigue in children have been reported, precluding generalization of instruments before assessment in other sociocultural contexts has been performed. Accordingly, our objectives were to obtain a norm-reference and to test the psychometric properties of the Dutch version of the PedsQL™ Multidimensional Fatigue Scale.

We hypothesized that the reliability and validity of the Dutch version is comparable to the original version. In addition, we expected adolescents to be more fatigued than younger children.

Methods

The PedsQL™ Multidimensional Fatigue Scale was distributed at day care facilities and schools in urban and suburban areas in the Netherlands, between October 2009 and May 2010. The questionnaire was self-administered, but children aged 5-7 years have been assisted by the researcher. One half of the participants were given the option to receive the questionnaire again after two weeks to assess test-retest reliability. The other half received two copies of the questionnaire to be completed independently by both parents to test inter-observer reliability.

The 18-item PedsQL™ Multidimensional Fatigue Scale reflects three subscales; general fatigue (GF), sleep/rest fatigue (SRF), and cognitive fatigue (CF). The questionnaire comprises parallel child self-reports for the ages 5-7 years (young child), 8-12 years (child) and 13-18 years (adolescent) and parent proxy-reports, the latter also including 2-4 years of age (toddler). The participants rated how often a particular problem occurred in the past month, using a 5-point Likert scale and for the young child self-report a 3-point scale. Each item is reverse-scored and rescaled to 0-100 scale, so that higher scores indicate fewer symptoms of fatigue.

Feasibility was evaluated from the percentage of missing answers. Range of measurement was based on the percentage of scores at extremes of the scaling range. Scale internal consistency was assessed by calculating Cronbach's coefficient alpha. Test-retest and inter-observer reliability and the parent/child concordance were assessed by intraclass cor-
relation coefficients (ICCs). ICCs were designated as ≤0.40 poor to fair agreement, 0.41-0.60 moderate agreement, 0.61-0.80 good agreement, and 0.81-1.00 excellent agreement. The ability of the questionnaire to distinguish between groups differing in health condition was computed using unpaired t-tests. Calculated effect sizes up to 0.20 were considered to be small, about 0.50 moderate and about 0.80 large.

The effect of socio-demographic variables was assessed using linear regression analysis. Within-group differences were assessed by analysis of variance with post hoc Bonferroni correction for age and education and t-tests for gender, country of birth, and family structure. Data were analyzed using SPSS 15.0.1. A p value of <0.05 was accepted as statistically significant.

**Results**

In total, 1257 parent reports and 1000 child reports were distributed, of which 497 and 366 reports were returned, respectively (response rates 40 and 37%). Sociodemographic characteristics of the initial test are presented in Table 7.1. Most reported chronic health conditions were asthma, allergies and attention-deficit hyperactivity disorder.

PedsQL™ Multidimensional Fatigue Scale scores are summarized in Table 7.2. Adolescents had more symptoms of GF than the other age ranges (mean difference = 6.20, p <0.001 for child reports and 5.27, p <0.001 for parent reports). Boys had more symptoms of CF than girls (mean difference = 3.42, p = 0.003 for parent reports). Children of immigrants had more prob-

### Table 7.1 Sociodemographic characteristics of sample

<table>
<thead>
<tr>
<th></th>
<th>Total sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Child report</strong></td>
<td></td>
</tr>
<tr>
<td>Age, years: mean (95% CI)</td>
<td>11.4 (11.0-11.8)</td>
</tr>
<tr>
<td>Gender, boys (%)</td>
<td>166 (45%)</td>
</tr>
<tr>
<td>Impaired health condition (%)</td>
<td>40 (11%)</td>
</tr>
<tr>
<td>Medication use (%)</td>
<td>29 (8%)</td>
</tr>
<tr>
<td>Ethnicity, Dutch (%)</td>
<td>359 (98%)</td>
</tr>
</tbody>
</table>

| **Parent report**         |              |
| Age child, years: mean (95% CI) | 9.5 (9.0-9.9) |
| Gender, male (%)          | 64 (13%)     |
| Ethnicity, Dutch (%)      | 466 (94%)    |
| Education, high (%)       | 203 (42%)    |
| Employment                | 419 (86%)    |
| Single-parent family (%)  | 38 (8%)      |

95% CI: 95% confidence interval.
### Table 7.2 Scale descriptives for PedsQL Multidimensional Fatigue Scale

<table>
<thead>
<tr>
<th>Scale</th>
<th>Toddler (2-4)</th>
<th>Young child (5-7)</th>
<th>Child (8-12)</th>
<th>Adolescent (13-18)</th>
<th>Total sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean</td>
<td>CI</td>
<td>n</td>
<td>Mean</td>
</tr>
<tr>
<td>Child report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total fatigue</td>
<td>68</td>
<td>76.59</td>
<td>73.16-80.03</td>
<td>143</td>
<td>78.70</td>
</tr>
<tr>
<td>General fatigue</td>
<td>68</td>
<td>83.46</td>
<td>79.61-87.30</td>
<td>143</td>
<td>82.66</td>
</tr>
<tr>
<td>Sleep/rest fatigue</td>
<td>68</td>
<td>74.00</td>
<td>69.47-78.52</td>
<td>143</td>
<td>77.55</td>
</tr>
<tr>
<td>Cognitive fatigue</td>
<td>68</td>
<td>72.24</td>
<td>66.99-77.49</td>
<td>143</td>
<td>75.76</td>
</tr>
<tr>
<td>Parent report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total fatigue</td>
<td>104</td>
<td>82.87</td>
<td>80.77-84.96</td>
<td>83</td>
<td>83.01</td>
</tr>
<tr>
<td>General fatigue</td>
<td>104</td>
<td>82.80</td>
<td>80.49-85.10</td>
<td>83</td>
<td>84.46</td>
</tr>
<tr>
<td>Sleep/rest fatigue</td>
<td>104</td>
<td>82.92</td>
<td>80.36-85.49</td>
<td>83</td>
<td>87.77</td>
</tr>
<tr>
<td>Cognitive fatigue</td>
<td>104</td>
<td>82.77</td>
<td>79.84-85.70</td>
<td>83</td>
<td>76.71</td>
</tr>
</tbody>
</table>

CI: 95% confidence interval.
lems with SRF compared to children of parents born in the Netherlands (mean difference = 6.12, \(p = 0.017\) for parent reports). Children in a single-parent family had more symptoms of CF (mean difference = 8.29, \(p = 0.037\) for parent reports) and lower total fatigue scores (mean difference = 5.93, \(p = 0.039\) for parent reports) than children living in a two-parent household. Children of low educated parents had more problems with CF (mean difference = 9.35, \(p = 0.03\) for child reports) and lower total fatigue scores (mean difference = 7.17, \(p = 0.014\) for child reports) than children of high educated parents.

Missing responses for all items were rare; 0.2% in parent reports and 0.3% in child reports. No floor effects were detected. Ceiling effects ranged from 1.4% in child reports to 5.1% in parent reports. All child report and parent report scales approached or exceeded a Cronbach’s alpha of 0.70 (range 0.64-0.93) (Table 7.3).

Table 7.3 Internal consistency reliability for PedsQL Multidimensional Fatigue Scale

<table>
<thead>
<tr>
<th>Scale</th>
<th>Age group (years)</th>
<th>Toddler (2-4)</th>
<th>Young child (5-7)</th>
<th>Child (8-12)</th>
<th>Adolescent (13-18)</th>
<th>Total sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child report</td>
<td></td>
<td>(\alpha)</td>
<td>(\alpha)</td>
<td>(\alpha)</td>
<td>(\alpha)</td>
<td>(\alpha)</td>
</tr>
<tr>
<td>Total fatigue</td>
<td>NA</td>
<td>0.79</td>
<td>0.85</td>
<td>0.86</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>General fatigue</td>
<td>NA</td>
<td>0.67</td>
<td>0.71</td>
<td>0.80</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>Sleep/rest fatigue</td>
<td>NA</td>
<td>0.66</td>
<td>0.67</td>
<td>0.64</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Cognitive fatigue</td>
<td>NA</td>
<td>0.74</td>
<td>0.86</td>
<td>0.81</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>Parent report</td>
<td></td>
<td>0.88</td>
<td>0.89</td>
<td>0.91</td>
<td>0.93</td>
<td>0.91</td>
</tr>
<tr>
<td>Total fatigue</td>
<td>0.78</td>
<td>0.80</td>
<td>0.83</td>
<td>0.86</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>General fatigue</td>
<td>0.72</td>
<td>0.70</td>
<td>0.79</td>
<td>0.81</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>Cognitive fatigue</td>
<td>0.90</td>
<td>0.90</td>
<td>0.92</td>
<td>0.93</td>
<td>0.92</td>
<td></td>
</tr>
</tbody>
</table>

\(\alpha\): Cronbach’s coefficient alpha.

Forty-three children (12%) and 75 parents (15%) performed the retest. The retest response for the young child was too low for evaluation. Child report and parent report test-retest ICCs had moderate to excellent agreement (range 0.50 to 0.85) (Table 7.4). At group level, no significant differences emerged between the test- and retest, except for GF which was reported to be better after two weeks by the parents.

Fifty-five participants (11%) returned two questionnaires, completed by both parents. Inter-observer reliability ICCs had poor to excellent agreement (range 0.27-0.93) (Table 7.5). At group level, there were no significant differences between fathers (n = 169) and mothers (n = 440), except for lower SRF scores reported by the mothers.
Parent/child concordance ICCs had poor to good agreement (range 0.10-0.68) (Table 7.6). At group level, means of SRF, CF and total fatigue of the parent report were significantly higher compared to the child report.

Child report and the parent report total scores and most subscale scores demonstrated a significant difference between the healthy participants (89%) and the participants with an impaired health condition (11%) (Table 7.7). Effect sizes varied from small to medium, with children with an impaired health condition showing lower scores and thus more fatigue.

### Table 7.4 Test-retest reliability for PedsQL Multidimensional Fatigue Scale

<table>
<thead>
<tr>
<th>Scale</th>
<th>Age group (years)</th>
<th>Toddler (2-4)</th>
<th>Young child (5-7)</th>
<th>Child (8-12)</th>
<th>Adolescent (13-18)</th>
<th>Total sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>ICC</td>
<td>n</td>
<td>ICC</td>
<td>n</td>
<td>ICC</td>
</tr>
<tr>
<td>Child report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total fatigue</td>
<td>NA</td>
<td>NA</td>
<td>19</td>
<td>0.84</td>
<td>23</td>
<td>0.71</td>
</tr>
<tr>
<td>General fatigue</td>
<td>NA</td>
<td>NA</td>
<td>19</td>
<td>0.74</td>
<td>23</td>
<td>0.75</td>
</tr>
<tr>
<td>Sleep/rest fatigue</td>
<td>NA</td>
<td>NA</td>
<td>19</td>
<td>0.85</td>
<td>23</td>
<td>0.50</td>
</tr>
<tr>
<td>Cognitive fatigue</td>
<td>NA</td>
<td>NA</td>
<td>18</td>
<td>0.75</td>
<td>23</td>
<td>0.78</td>
</tr>
<tr>
<td>Parent report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total fatigue</td>
<td>24</td>
<td>0.68</td>
<td>NA</td>
<td>NA</td>
<td>24</td>
<td>0.82</td>
</tr>
<tr>
<td>General fatigue</td>
<td>24</td>
<td>0.57</td>
<td>NA</td>
<td>NA</td>
<td>24</td>
<td>0.69</td>
</tr>
<tr>
<td>Sleep/rest fatigue</td>
<td>24</td>
<td>0.59</td>
<td>NA</td>
<td>NA</td>
<td>24</td>
<td>0.83</td>
</tr>
<tr>
<td>Cognitive fatigue</td>
<td>24</td>
<td>0.69</td>
<td>NA</td>
<td>NA</td>
<td>24</td>
<td>0.73</td>
</tr>
</tbody>
</table>

ICC: internal consistency reliability.

* Only 1 child and 2 parents filled in the retest.

### Table 7.5 Inter-observer reliability for PedsQL Multidimensional Fatigue Scale

<table>
<thead>
<tr>
<th>Scale</th>
<th>Age group (years)</th>
<th>Toddler (2-4)</th>
<th>Young child (5-7)</th>
<th>Child (8-12)</th>
<th>Adolescent (13-18)</th>
<th>Total sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>ICC</td>
<td>n</td>
<td>ICC</td>
<td>n</td>
<td>ICC</td>
</tr>
<tr>
<td>Total fatigue</td>
<td>40</td>
<td>0.62</td>
<td>7</td>
<td>0.93</td>
<td>12</td>
<td>0.81</td>
</tr>
<tr>
<td>General fatigue</td>
<td>40</td>
<td>0.43</td>
<td>7</td>
<td>0.71</td>
<td>12</td>
<td>0.54</td>
</tr>
<tr>
<td>Sleep/rest fatigue</td>
<td>40</td>
<td>0.87</td>
<td>7</td>
<td>0.84</td>
<td>12</td>
<td>0.67</td>
</tr>
<tr>
<td>Cognitive fatigue</td>
<td>40</td>
<td>0.27</td>
<td>7</td>
<td>0.90</td>
<td>12</td>
<td>0.87</td>
</tr>
</tbody>
</table>

ICC: internal consistency reliability.
Table 7.6 Parent child concordance for PedsQL Multidimensional Fatigue Scale

<table>
<thead>
<tr>
<th>Scale</th>
<th>Toddler (2-4)</th>
<th>Young child (5-7)</th>
<th>Child (8-12)</th>
<th>Adolescent (13-18)</th>
<th>Total sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n  ICC</td>
<td>n  ICC</td>
<td>n  ICC</td>
<td>n  ICC</td>
<td>n  ICC</td>
</tr>
<tr>
<td>Total fatigue</td>
<td>NA 0.25</td>
<td>68 0.25</td>
<td>140 0.68</td>
<td>153 0.52</td>
<td>361 0.53</td>
</tr>
<tr>
<td>General fatigue</td>
<td>NA 0.22</td>
<td>68 0.22</td>
<td>140 0.48</td>
<td>153 0.50</td>
<td>361 0.46</td>
</tr>
<tr>
<td>Sleep/rest fatigue</td>
<td>NA 0.10</td>
<td>68 0.10</td>
<td>140 0.46</td>
<td>153 0.41</td>
<td>361 0.36</td>
</tr>
<tr>
<td>Cognitive fatigue</td>
<td>NA 0.22</td>
<td>68 0.22</td>
<td>140 0.66</td>
<td>153 0.45</td>
<td>361 0.49</td>
</tr>
</tbody>
</table>

ICC: internal consistency reliability.

Table 7.7 Construct validity for PedsQL Multidimensional Fatigue Scale

<table>
<thead>
<tr>
<th>Scale</th>
<th>Healthy sample</th>
<th>Impaired health condition sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n  Mean</td>
<td>CI 95% CI</td>
</tr>
<tr>
<td>Child report</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total fatigue</td>
<td>326 77.40</td>
<td>76.06-78.73</td>
</tr>
<tr>
<td>General fatigue</td>
<td>326 80.48</td>
<td>78.92-82.04</td>
</tr>
<tr>
<td>Sleep/rest fatigue</td>
<td>326 75.00</td>
<td>73.37-76.64</td>
</tr>
<tr>
<td>Cognitive fatigue</td>
<td>326 76.65</td>
<td>74.77-78.52</td>
</tr>
<tr>
<td>Parent report</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total fatigue</td>
<td>442 81.90</td>
<td>80.75-83.06</td>
</tr>
<tr>
<td>General fatigue</td>
<td>442 82.00</td>
<td>80.69-83.30</td>
</tr>
<tr>
<td>Sleep/rest fatigue</td>
<td>442 84.76</td>
<td>83.49-86.03</td>
</tr>
<tr>
<td>Cognitive fatigue</td>
<td>442 78.91</td>
<td>77.27-80.56</td>
</tr>
</tbody>
</table>

CI: 95% confidence interval.

**Discussion**

Our cohort showed that adolescence was associated with more fatigue, which might reflect the decrease in sleep duration at that age. Boys reported more fatigue than girls, indicating socially related gender differences. Fatigue was more common in children living in a single-parent family, which might be explained by the higher prevalence of sleep problems in these children. Being a child of an immigrant was associated with more fatigue, stressing the influence of the socio-cultural background on fatigue. Children of low educated parents reported more fatigue, corresponding with lower reported HRQOL in populations with low socioeconomic status.

The questionnaire was found to be feasible because of the minimal missing item responses. All scales approached or exceeded a Cronbach's alpha of 0.70, recommended for comparing groups. Parent report total scores approached or exceeded a Cronbach's alpha of 0.90, recommended for analyzing individual patient scales. Test-retest reliability was low in the tod-
The Dutch PedsQL™ Multidimensional Fatigue Scale

dlder version, suggesting that this version should be applied with caution. The low test-retest
accordance in adolescents might be a real variation due to the fluctuations in nocturnal sleep
duration. Inter-observer agreement was low in toddlers and adolescents. This observation
is difficult to explain and requires further exploration. The father/child agreement and the
mother/child agreement was low in the young child, the adolescent and the total sample.
This may be due to differences in reasoning and response reactions between parent and
child. Children scored lower than their parents. The low parent/child concordance (including
the weakest agreement for the young child and lower scores reported by the children)
has consistently been observed in HRQOL measurement, particularly for internalizing prob-
lems. These findings support the need to measure the perspectives of the child and the
parent, since both may influence healthcare utilization. The assumption was confirmed that
the questionnaire was able to distinguish between healthy children and children with an
impaired health condition. It is expected that a clinical sample with more severely ill children
will demonstrate worse fatigue scores and higher effect sizes.

Regarding the current study several limitations need to be mentioned. First, we had
low response rates, which might lead to non-response bias. Second, the ethnicity of the
participants was rather homogenous; only 2% of the children and 6% of the parents were
born outside the Netherlands, compared to 10% and 10%, respectively in the Dutch popula-
tion. A possible explanation is the language problems that immigrants experienced and
therefore decided not to participate in this study. Third, the educational background of our
parent respondents showed that 42% were highly educated, compared to 18% of the Dutch
population. Highly educated parents might have been more aware of the necessity of this
study wherefore more willing to participate. Children from immigrants as well as from low
educated parents may experience more fatigue; hence, the obtained norm-reference may
underestimate fatigue in the general population Information on nonparticipants was not
available, thus generalization of the norm-results should be made with caution.

In conclusion, the Dutch version of the PedsQL™ Multidimensional Fatigue Scale demon-
strates overall adequate psychometric properties in another sociocultural context. With the
obtained norm-references, it can be utilized as a tool to evaluate fatigue in children.
References


Chapter 8

Sleep in adolescents: preliminary psychometric properties of the Dutch version of the Adolescent Sleep Habits Questionnaire (ASHQ)

M. Suzanne Gordijn, Gertjan J.L. Kaspers, Reinoud J.B.J. Gemke
Abstract

Relatively little is known about sleep in adolescents. Instruments to measure sleep in adolescents are sparse. This study aims to report on the reliability, validity, and the norm-references of the Dutch version of the Adolescent Sleep Habits Questionnaire. In total, 665 adolescents aged 12-18 years and 668 parents completed this questionnaire. Internal consistency reliability was unsatisfactory for the subscale scores, but satisfactory for the total scores. The questionnaire was able to discriminate between healthy adolescents and chronically ill adolescents. These results demonstrate that the Dutch version of the Adolescent Sleep Habits Questionnaire has adequate reliability for the total scores and satisfactory validity. This version can be utilized for assessing sleep habits in adolescents, for which norm-references have been obtained.
Introduction

Adolescent sleep is typified by late bedtimes, resulting in insufficient total sleep time, which is particularly a problem on school nights. Inadequate sleep time in adolescence may lead to daytime sleepiness, attention problems, poor school performances, emotional problems and difficulties in social contacts. Because of these concerns, reliable and valid assessment of sleep problems in adolescence is important. In contrast with the emerging literature on sleep in children, relatively little is known about sleep in adolescence. Moreover, adolescents may have specific sleep problems due to late-night social activities and maturational changes in sleep regulatory processes. Therefore, instruments evaluating sleep in adolescence should be adapted to fit the adolescent way of life. However, reliable and valid instruments to measure sleep in adolescence are sparse. The Children’s Sleep Habits Questionnaire (CSHQ) is a reliable and valid screening tool for sleep difficulties for children aged 4–10 years, based on the International Classification of Sleep Disorders. Recently, an adolescent version of the CSHQ has been derived; the Adolescent Sleep Habits Questionnaire (ASHQ). This questionnaire has been translated with permission into Dutch, in accordance with internationally accepted methods. Substantial inter-cultural differences in sleep have been reported, precluding generalization of instruments before assessment in other social-cultural contexts has been performed. Accordingly, the aim of this study was to obtain a norm-reference and to evaluate the feasibility, reliability and validity of the Dutch version of the ASHQ.

Methods

Participants

The study was conducted in four high schools with a wide variety of educational levels in urban and suburban areas in the Netherlands, between February 2010 and April 2011. The ASHQ was distributed at the schools together with written information on the study and a stamped return envelope to be taken home by the children for their parents. Information regarding socio-demographic variables of the child (age, gender, country of birth) and the responding parent (age, gender, country of birth, marital status, education, employment) and the presence of a chronic health condition (physical or psychological) of the child was also collected. The ASHQ was self-administered for adolescents and parents. The instructions stated that the parent should complete the parent proxy-report separately from their child. The adolescents and their parents were asked to return the forms within two weeks. The study was approved by our Institutional Review Board.
Measurement
The ASHQ comprises a 50-item adolescent self-report and a 54-item parent proxy-report, allowing for a total score and six subscales: “Bedtime”, “Sleep behavior”, “Waking during the night”, “Morning waking”, “Sleep habits” and “Daytime sleepiness”. The participants rated how often a particular problem occurred in the most recent typical two weeks, using a four-point Likert scale, with an extra option “don’t know”. Responses are scored on a scale from zero to three. Higher scores indicate more sleep disturbances. Additional questions collect information regarding evening bedtime, morning wake-up time, the amount of sleep per night, the person that sets the rules about going to bed and the respondent’s opinion about the presence of sleep problems. Completion time of the ASHQ is about 10 minutes.

Statistical analysis
Feasibility was evaluated from the percentage of missing answers. Subscale scores were not calculated if more than 50% of the subscale answers were missing. Total scores were calculated if all subscales had less than 50% missing answers. For the remaining missing values, a series mean was imputed for that item. Range of measurement was based on the percentage of scores at extremes of the scaling range (floor and ceiling effects). Scale internal consistency reliability was assessed by calculating Cronbach’s coefficient alpha. Scales with a reliability of 0.70 or greater are considered as adequate for comparing patient groups. Construct validity was determined using the known-groups method. The ability of the ASHQ to discriminate between groups differing in known health condition (healthy participants and participants with a reported chronic health condition) was computed using unpaired t-tests. We assumed that healthy adolescents would report lower scores (fewer sleep disturbances) than adolescents with a chronic health condition. In order to determine the magnitude of the differences, effect sizes were calculated by dividing the difference in mean scores between both groups by the standard deviation of the healthy group. Effect sizes for differences in means up to 0.20 were considered to be small, effect sizes of about 0.50 moderate and effect sizes of about 0.80 large. Spearman correlations were calculated to examine correlations between total and subscale scores and age. Within-group differences were assessed by analysis of variance with post hoc Bonferroni correction for education, employment and marital status and t-tests for gender and country of birth. All statistical analyses were performed with SPSS (Version 15.0.1, Chicago, IL). A two-sided p value of <0.05 was accepted as statistically significant.
The Dutch Adolescent Sleep Habits Questionnaire

Results

Demographics
Altogether, 2393 questionnaires were distributed, of which 665 (28%) adolescent reports and 668 (28%) parent reports were returned. Four adolescent and parent reports were excluded because the respondents fell outside the age range of high school students: 12–18 years. Demographic data of the participants are presented in Table 8.1. The most common chronic health conditions were asthma, attention-deficit hyperactivity disorder and allergies. Mothers comprised the majority (84%) of the parent respondents. Analyses were performed on the total study group, including participants with a chronic health condition, to ensure applicability to a general adolescent population.

<table>
<thead>
<tr>
<th>Table 8.1 Demographic characteristics of sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sample</td>
</tr>
<tr>
<td>Adolescent report (n=661)</td>
</tr>
<tr>
<td>Age, years: mean (95% CI) 14.2 (14.1-14.4)</td>
</tr>
<tr>
<td>Gender, boys 277 (42%)</td>
</tr>
<tr>
<td>Impaired health condition 78 (12%)</td>
</tr>
<tr>
<td>Medication use 63 (10%)</td>
</tr>
<tr>
<td>Ethnicity, Dutch 631 (96%)</td>
</tr>
<tr>
<td>Parent report (n=664)</td>
</tr>
<tr>
<td>Ethnicity, Dutch 615 (93%)</td>
</tr>
<tr>
<td>Education, high 331 (50%)</td>
</tr>
<tr>
<td>Employment 563 (85%)</td>
</tr>
<tr>
<td>Single-parent family 93 (14%)</td>
</tr>
</tbody>
</table>

95% CI: 95% confidence interval.

Scale descriptives
On school days, the mean bedtime was 9.55 p.m. (95% CI 9.51–9.58 p.m.) reported by adolescents and 9.49 p.m. (95% CI 9.46–9.52 p.m.) reported by parents; the mean wake-up time was 6.57 a.m. (95% CI 6.55–6.59 a.m.) and 6.58 a.m. (95% CI 6.56–7.00 a.m.), respectively. On weekend days, the mean bedtime was 11.11 p.m. (95% CI 11.06–11.17 p.m.) reported by adolescents and 10.57 p.m. (95% CI 10.53–11.02 p.m.) reported by parents; the mean wake-up time was 9.36 a.m. (95% CI 9.30–9.42 a.m.) and 9.29 a.m. (95% CI 9.24–9.34 a.m.), respectively. Responses of the adolescents to additional questions demonstrated that 5% of them have a lot of trouble sleeping, 35% have some trouble sleeping and 59% have no trouble sleeping at all. Twenty-three percent of the adolescents reported to sleep too little on average school nights.
ASHQ scores are summarized in Table 8.2. Girls reported more sleep disturbances than boys (mean difference total score = 4.14, P < 0.001). Parents of children born outside the Netherlands reported fewer sleep problems in their children compared with parents of Dutch children (mean difference total score = 2.80, P = 0.027). An increasing age of the adolescent correlated with higher total scores on the self-reports and the proxy-reports (r 0.14, P < 0.001 and r 0.19, P < 0.001, respectively).

### Feasibility

Missing responses (including the answer option “don’t know”) for all items were 3.5% for the adolescent reports and 5.6% for the parent reports.

### Range of measurement

There were no significant floor or ceiling effects, with none of the respondents scoring at the minimum or the maximum. The lowest total score was 13.0 scored by adolescents and 7.0 scored by parents. The highest total score was 81.85 and 82.56, respectively.

### Table 8.2 Scale descriptives for the Adolescent Sleep Habits Questionnaire

<table>
<thead>
<tr>
<th>Scale</th>
<th>Total sample</th>
<th>n</th>
<th>Mean</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adolescent report</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>652</td>
<td>42.20</td>
<td>41.38-43.01</td>
<td></td>
</tr>
<tr>
<td>Subscale item</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bedtime</td>
<td>661</td>
<td>9.29</td>
<td>9.07-9.51</td>
<td></td>
</tr>
<tr>
<td>Sleep behavior</td>
<td>660</td>
<td>5.65</td>
<td>5.39-5.90</td>
<td></td>
</tr>
<tr>
<td>Waking during the night</td>
<td>661</td>
<td>2.66</td>
<td>2.50-2.82</td>
<td></td>
</tr>
<tr>
<td>Morning waking</td>
<td>658</td>
<td>9.10</td>
<td>8.89-9.32</td>
<td></td>
</tr>
<tr>
<td>Sleep habits</td>
<td>659</td>
<td>3.23</td>
<td>3.09-3.37</td>
<td></td>
</tr>
<tr>
<td>Daytime sleepiness</td>
<td>656</td>
<td>12.25</td>
<td>11.98-12.52</td>
<td></td>
</tr>
<tr>
<td><strong>Parent report</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>620</td>
<td>28.92</td>
<td>28.17-29.67</td>
<td></td>
</tr>
<tr>
<td>Subscale item</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bedtime</td>
<td>664</td>
<td>11.26</td>
<td>10.98-11.53</td>
<td></td>
</tr>
<tr>
<td>Sleep behavior</td>
<td>651</td>
<td>3.51</td>
<td>3.30-3.73</td>
<td></td>
</tr>
<tr>
<td>Waking during the night</td>
<td>634</td>
<td>1.71</td>
<td>1.59-1.82</td>
<td></td>
</tr>
<tr>
<td>Morning waking</td>
<td>662</td>
<td>7.81</td>
<td>7.54-8.08</td>
<td></td>
</tr>
<tr>
<td>Sleep habits</td>
<td>663</td>
<td>2.35</td>
<td>2.23-2.46</td>
<td></td>
</tr>
<tr>
<td>Daytime sleepiness</td>
<td>656</td>
<td>2.50</td>
<td>2.26-2.74</td>
<td></td>
</tr>
</tbody>
</table>

CI: 95% confidence interval.
Internal consistency

The total scores of the adolescent reports and the parent reports approached or exceeded a Cronbach's alpha of 0.70 (0.77 and 0.69, respectively), recommended for comparing groups (Table 8.3). The Cronbach alpha's of the subscales of the adolescent reports and the parent reports varied from low to adequate reliability (range < 0.01 - 0.79).

Table 8.3 Internal consistency reliability for the Adolescent Sleep Habits Questionnaire

<table>
<thead>
<tr>
<th>Scale</th>
<th>Total sample</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adolescent report</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td></td>
<td>0.77</td>
</tr>
<tr>
<td>Subscale item</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bedtime</td>
<td></td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Sleep behavior</td>
<td></td>
<td>0.53</td>
</tr>
<tr>
<td>Waking during the night</td>
<td></td>
<td>0.54</td>
</tr>
<tr>
<td>Morning waking</td>
<td></td>
<td>0.48</td>
</tr>
<tr>
<td>Sleep habits</td>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td>Daytime sleepiness</td>
<td></td>
<td>0.79</td>
</tr>
<tr>
<td><strong>Parent report</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td></td>
<td>0.69</td>
</tr>
<tr>
<td>Subscale item</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bedtime</td>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td>Sleep behavior</td>
<td></td>
<td>0.58</td>
</tr>
<tr>
<td>Waking during the night</td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>Morning waking</td>
<td></td>
<td>0.61</td>
</tr>
<tr>
<td>Sleep habits</td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>Daytime sleepiness</td>
<td></td>
<td>0.78</td>
</tr>
</tbody>
</table>

α: Cronbach's coefficient alpha.

Construct validity

Table 8.4 contains the comparisons between the healthy participants and the participants with a reported impaired health condition as a group. There were no significant differences in socio-demographic variables between both groups. For the adolescent report and the parent report, the total scores and certain subscale scores demonstrated a significant difference. The effect sizes varied from small to medium, with adolescents with an impaired health condition showing higher scores and thus more sleep disturbances.
Table 8.4 Construct validity for Adolescent Sleep Habits Questionnaire

<table>
<thead>
<tr>
<th>Scale</th>
<th>Healthy sample</th>
<th>Impaired health condition sample</th>
<th>Difference</th>
<th>Effect size</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean</td>
<td>CI</td>
<td>n</td>
<td>Mean</td>
</tr>
<tr>
<td><strong>Adolescent report</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>578</td>
<td>41.82</td>
<td>40.97-42.67</td>
<td>74</td>
<td>45.13</td>
</tr>
<tr>
<td>Subscale item</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bedtime</td>
<td>584</td>
<td>9.24</td>
<td>9.01-9.47</td>
<td>77</td>
<td>9.69</td>
</tr>
<tr>
<td>Sleep behavior</td>
<td>584</td>
<td>5.55</td>
<td>5.29-5.81</td>
<td>76</td>
<td>6.40</td>
</tr>
<tr>
<td>Waking during the night</td>
<td>584</td>
<td>2.65</td>
<td>2.48-2.81</td>
<td>77</td>
<td>2.81</td>
</tr>
<tr>
<td>Morning waking</td>
<td>582</td>
<td>9.04</td>
<td>8.81-9.27</td>
<td>76</td>
<td>9.58</td>
</tr>
<tr>
<td>Sleep habits</td>
<td>582</td>
<td>3.19</td>
<td>3.04-3.34</td>
<td>77</td>
<td>3.55</td>
</tr>
<tr>
<td>Daytime sleepiness</td>
<td>580</td>
<td>12.16</td>
<td>11.88-12.45</td>
<td>76</td>
<td>12.96</td>
</tr>
<tr>
<td><strong>Parent report</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>551</td>
<td>28.38</td>
<td>27.62-29.13</td>
<td>69</td>
<td>33.26</td>
</tr>
<tr>
<td>Subscale item</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bedtime</td>
<td>587</td>
<td>11.18</td>
<td>10.88-11.47</td>
<td>77</td>
<td>11.88</td>
</tr>
<tr>
<td>Sleep behavior</td>
<td>575</td>
<td>3.36</td>
<td>3.15-3.57</td>
<td>76</td>
<td>4.66</td>
</tr>
<tr>
<td>Waking during the night</td>
<td>562</td>
<td>1.66</td>
<td>1.54-1.78</td>
<td>72</td>
<td>2.05</td>
</tr>
<tr>
<td>Morning waking</td>
<td>586</td>
<td>7.65</td>
<td>7.37-7.94</td>
<td>76</td>
<td>9.01</td>
</tr>
<tr>
<td>Sleep habits</td>
<td>587</td>
<td>2.31</td>
<td>2.19-2.43</td>
<td>76</td>
<td>2.63</td>
</tr>
<tr>
<td>Daytime sleepiness</td>
<td>581</td>
<td>2.42</td>
<td>2.18-2.66</td>
<td>75</td>
<td>3.10</td>
</tr>
</tbody>
</table>

CI: 95% confidence interval.
Discussion

This paper reports on sleep habits in Dutch adolescents and on the psychometric properties of the Dutch version of the ASHQ. Our results show that the mean bedtime on weekdays was 9.52 p.m. and the mean wake-up time was 6.58 a.m. The average time spent in bed was 9 hours and 6 minutes. However, bedtime is not equal to sleep onset time, which has been reported to occur on average 16.8 minutes later than bedtime for adolescents.\(^{19}\) Moreover, one-third of our respondents answered that it usually takes longer than 30 minutes to fall asleep after “lights out”. Therefore, most adolescents did not obtain the optimal amount of sleep on school nights, which is defined, regardless of pubertal stage, as 9 hours and 12 minutes.\(^{20,21}\) As insufficient sleep in adolescence may lead to significant morbidity, adolescents need to be educated on these negative effects and on good sleep hygiene.\(^2,3\)

Besides the insufficient amount of sleep on school nights, approximately one-third of the adolescents have some trouble sleeping and 5% have a lot of trouble sleeping. Similar to previous studies, girls reported more sleep difficulties than boys, demonstrating gender-related differences.\(^{13,22}\) Children born outside the Netherlands had fewer sleep problems according to their parents in comparison with Dutch children, emphasizing the influence of sociocultural background on sleep.\(^9,13\) Sleep disturbances were age-related; with more problems at a higher age. A possible explanation is the fact that sleep onset time becomes increasingly later during adolescence whilst wake-up time does not change because of school, resulting in reduced sleep on school nights.\(^1,22\)

The Dutch version of the ASHQ seems to be feasible because of the minimal number of missing item responses. Floor and ceiling effects were minimal as well. Furthermore, the internal consistency coefficients of the total scores approached or exceeded a Cronbach's alpha of 0.70, recommended for comparing groups.\(^{16,17}\) The internal consistency coefficients of most subscales did not approach adequate reliability. These subscales require further exploration. Until then, we suggest only using the total scores.

All total scores and several subscale scores of the ASHQ were able to distinguish between healthy participants and participants with an impaired health condition, demonstrating adequate validity. In addition, the assumption was confirmed that healthy adolescents would have lower ASHQ scores (fewer sleep disturbances) than adolescents with an impaired health condition.

Some limitations of this study need to be mentioned. First, we had a low response rate, which might lead to non-response bias. This might be due to the fact that we combined the current study with the assessment of an additional questionnaire. Completion of both questionnaires might have meant too much effort and time for the adolescents and their parents. Moreover, we were not able to send reminders to the non-responders. Second, 50% of our parent respondents were highly educated, compared with 18% of the Dutch population.\(^{23}\) A possible explanation is that highly educated parents were more aware of the relevance of this
study and, therefore, more willing to respond. As children in higher socioeconomic classes are known to experience fewer difficulties sleeping, underestimation might be present. In summary, the Dutch version of the ASHQ adolescent report and parent report demonstrates adequate reliability for the total scores and satisfactory validity. With the obtained norm-references, it can be utilized as a tool in the evaluation of sleep in adolescents aged 12–18 years.
References

Chapter 9

Sleep, fatigue, depression and quality of life in survivors of childhood acute lymphoblastic leukemia

M. Suzanne Gordijn, Raphaële R.L. van Litsenburg, Reinoud J.B.J. Gemke, Jaap Huisman, Marc B. Bierings, Peter M. Hoogerbrugge, Gertjan J.L. Kaspers

Pediatr Blood Cancer 2013;60:479-85
Abstract

Background
With the improved survival of childhood acute lymphoblastic leukemia (ALL), the effect of treatment on psychosocial well-being becomes increasingly relevant. Literature on sleep and fatigue during treatment is emerging. However, information on these subjects after treatment is sparse. This cross-sectional study examined sleep and fatigue in relation to depression and quality of life (QoL) after treatment for childhood ALL.

Procedure
Sleep, fatigue, depression and QoL were evaluated by parent proxy and/or child self-reports of the Children's Sleep Habits Questionnaire, the PedsQL™ multidimensional fatigue scale, the Children's Depression Inventory and the Child Health Questionnaire. All total scores were compared to Dutch norm references.

Results
Sixty-two children were included, being 36 (interquartile range 22-62) months after finishing treatment. Parents rated the ALL survivors as having more disturbed sleep, more fatigue and poorer physical QoL compared to the Dutch norm. ALL survivors themselves reported less sleep problems, less depressive symptoms, and better psychosocial QoL than the Dutch norm. More sleep disturbances and fatigue correlated with more symptoms of depression and a worse QoL.

Conclusions
Differences in parental and self-reports, including worse parental ratings, might be explained by worried parents and/or the adaptive style of the children. Impaired sleep and fatigue correlated with more depressive symptoms and a worse QoL.
Introduction

Of all malignancies in children, acute lymphoblastic leukemia (ALL) is the most common type. Survival rates of ALL have significantly improved over time; currently, survival has reached 80-85%. Therefore, the adverse effects of treatment on physical and psychosocial well-being becomes increasingly relevant. Quantifying and understanding these undesired effects will enable clinicians in follow-up clinics to screen for these sequelae and to provide tailored care.

Literature on sleep and fatigue during treatment for childhood ALL is emerging. Several studies have demonstrated poor sleep quality secondary to the treatment for ALL. Zupanec et al. reported that sleep disturbances during maintenance therapy for childhood ALL are common (87%) and that these contribute to increased fatigue. In addition, Hinds et al. found that dexamethasone exposure is an important contributing factor to sleep disturbances and fatigue during treatment for ALL. Moreover, Van Litsenburg et al. identified impaired sleep as a contributing factor to poor quality of life (QoL) during treatment for childhood ALL. However, relatively little is known about the prevalence and impact of sleep problems and fatigue and its relation to depression and QoL after treatment for childhood ALL.

Sleep disturbances and fatigue have commonly been reported among adult cancer survivors and survivors of central nervous system tumors in childhood. Previous studies suggested that sleep problems and fatigue may originate from drug-induced neurotoxicities and a history of radiation. Sleep problems and fatigue in cancer survivors have been described to be associated with depression and an impaired QoL. Sleep problems and fatigue may lead to depression and a poor QoL and, likewise, the reverse. Therefore, the interpretation of the causality of this interrelationship is complex. As differing diagnoses and treatments have different levels of adverse effects, it is important to focus on specific diagnostic groups.

The aim of the present cross-sectional, multi-center study was to determine sleep, fatigue, depression and QoL in survivors of childhood ALL, compared to Dutch norm references. Based on the results conducted during treatment for childhood ALL and across survivors of other types of malignancies, we hypothesized that survivors of childhood ALL would have more sleep disturbances, more fatigue, more symptoms of depression and poorer QoL. To the best of our knowledge, this is the first study to explore the combination of sleep, fatigue and its relation with depression and QoL exclusively in survivors of childhood ALL.
Methods

Participants
The study population consisted of children between 5 and 17 years of age (at the time of the study) that had been successfully treated according to the Dutch Childhood Oncology Group (DCOG) ALL-9 or ALL-10 protocol between May 1997 and February 2008 in the VU University Medical Center Amsterdam, the University Medical Center Utrecht or the Radboud University Nijmegen Medical Center in the Netherlands. Based on clinical and biological factors and on the response to treatment, patients treated according to the ALL-9 protocol were classified in a non-high risk (NHR) or a high risk (HR) group and patients treated according to the ALL-10 protocol were classified in a standard risk (SR), a medium risk (MR) or a high risk (HR) group. Exclusion criteria were being under treatment for replased ALL and deficient Dutch language skills.

Procedure
All eligible children between 8 and 17 years of age and parents of eligible children between 5 and 17 years of age were invited by mail to participate. After receiving back the signed consent form, the questionnaires to be completed were mailed to the subjects. Information regarding socio-demographic variables (age, gender and country of birth) of the child and the responding parent was also collected. The instructions stated that one of the parents should complete the parent proxy-report separately from their child. Treatment variables were abstracted from medical records. This study was approved by our Institutional Review Board and written informed consent was obtained from the children (above 12 years of age) and their caregivers.

Measurement

Sleep
Sleep was evaluated by the validated Dutch version of the Children’s Sleep Habits Questionnaire (CSHQ), which has an internal consistency reliability (Cronbach’s alpha) varying from 0.47-0.68.20-21 Parents of children between 5 and 12 years of age completed the 33-item parent proxy-report which reflects eight key sleep domains; bedtime resistance, sleep onset delay, sleep duration, sleep anxiety, night wakening, parasomnias, sleep-disordered breathing and daytime sleepiness. Children of 8-12 years of age completed the 26-item child self-report which allows for a total score. Items were rated on a three-point scale. Dutch reference data were used.22 Participants between 13 and 17 years of age and their parents completed the adapted adolescent version of the CSHQ; the Adolescent Sleep Habits Questionnaire (ASHQ). The ASHQ comprises a 50-item adolescent self-report and a 54-item parent proxy-report. Items were rated on a four-point scale. Some items of both questionnaires were reversed in
order to consistently make a higher score indicative of more disturbed sleep. For the purpose of this study, our study group examined the psychometric properties of the ASHQ and collected Dutch reference data. The ASHQ total scores demonstrated satisfactory internal consistency reliability (Cronbach’s alpha of 0.77 for the self-report and 0.69 for the parent-report) and adequate construct validity determined using the known-groups method ($P = 0.029$ for the self-report and $P = 0.002$ for the parent-report). Reliability and validity of the subscale scores need further study and were therefore not reported.

**Fatigue**

Parent proxy-report and child self-reports of the validated Dutch version of the PedsQL™ Multidimensional Fatigue Scale (PedsQL MFS) were used. This measure demonstrated adequate internal consistency reliability (Cronbach’s alpha varied from 0.77-0.92 [mean 0.84] for the parent-report and from 0.64-0.83 [all subscales >0.73 except for one, mean 0.73] for the child report).23-24 Children of 8-17 years of age and parents of children between 5 and 17 years of age completed this 18-item questionnaire which reflects three subscales; general fatigue, sleep/rest fatigue and cognitive fatigue. Each item is reverse-scored and rescaled to 0-100 scale, so that higher scores indicate fewer symptoms of fatigue. Dutch reference data were used that were collected and published by our group.24

**Depression**

The Dutch version of the Children’s Depression Inventory (CDI), completed by children between 8 and 17 years of age was included to evaluate symptoms of depression.25-26 This 27-item self-report inventory was developed as a downward extension of the adult-oriented Beck Depression Inventory and has excellent internal consistency reliability (Cronbach’s alpha of 0.85).27 Responses are scored on a scale from zero to two, with total CDI scores ranging between zero and 54. Higher scores indicate more symptoms of depression. Dutch norm scores (means) were used for comparison. CDI norms are provided by sex.25

**QoL**

Quality of life was evaluated by the validated Dutch version of the Child Health Questionnaire (CHQ). Cronbach’s alpha of this measure varied from 0.39-0.96 (mean 0.72) for the parent-report and from 0.56-0.90 (all subscales >0.70 except for one, mean 0.80) for the child report.28-30 Children of 8-17 years of age completed the 87-item child form and parents of children between 5 and 17 years of age completed the 50-item parent form. Per scale the items are summed up and transformed into a zero (worst possible score) to 100 (best possible score) scale. Physical and psychosocial summary scores were calculated.31 Dutch norm references (means) were used.28,30
Statistical analysis
Differences between participants and non-participants with respect to gender, treatment protocol and risk group stratification were examined with the chi-square test and differences in age and time since end of treatment were analyzed by means of t-tests. Questionnaire scores of the ALL survivors were compared to Dutch norm references for similar age categories. Differences between the ALL survivors and the norm references in questionnaire scores were computed using the independent samples t-test (CSHQ parent-reports and PedsQL MFS) and the one sample t-test (CHQ) or the Mann-Whitney U-test (CSHQ self-report and ASHQ) and the one-sample Kolmogorov-Smirnov test (CDI) when not normally distributed. In order to determine the magnitude of the differences between ALL survivors and the Dutch norm references, effect sizes were calculated by dividing the difference in mean scores by the standard deviation of the Dutch norm references (available for the CSHQ, the ASHQ and the PedsQL MFS). Effect sizes up to 0.20 were considered to be small, about 0.50 moderate and about 0.80 large. Differences in questionnaire scores between socio-demographic and treatment variables were assessed by t-tests for gender, country of birth (inside or outside the Netherlands) and treatment protocol (ALL-9 or ALL-10) and by analysis of variance with post hoc Bonferroni correction for ALL treatment protocol risk groups. Pearson’s correlations and linear regression analyses were calculated to examine the relation between questionnaire total scores on the one hand and socio-demographic variables (age) and treatment variables (time since end of treatment) on the other hand. Pearson’s correlations were also calculated to quantify the relationship between sleep, fatigue, depression and QoL questionnaire total scores. All statistical analyses were performed with SPSS (Version 15.0.1, Chicago, IL). To adjust for errors associated with multiple testing, a two-sided P-value of <0.01 was accepted as statistically significant.

Results
Demographics
We invited 146 eligible subjects to complete the questionnaires. Completed questionnaires concerning 62 subjects (overall response rate of 42%) were returned (35 pairs of self-reports and parent-reports, one separate self-report and 26 separate parent-reports). Parent respondents were mostly mothers (90%). No significant differences emerged among participants and non-participants with respect to age, gender, treatment protocol, risk group stratification and time since end of treatment. Descriptive data on all participating subjects are shown in Table 9.1.
Table 9.1 Sociodemographic and treatment characteristics of sample

<table>
<thead>
<tr>
<th>Age, years: mean (±SD)</th>
<th>9.7 (± 3.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age group</td>
<td></td>
</tr>
<tr>
<td>5-7 years: n (%)</td>
<td>18 (29%)</td>
</tr>
<tr>
<td>8-11 years: n (%)</td>
<td>25 (40%)</td>
</tr>
<tr>
<td>12-17 years: n (%)</td>
<td>19 (31%)</td>
</tr>
<tr>
<td>Time since end of treatment, months: median (±IQR)</td>
<td>36 (22-62)</td>
</tr>
<tr>
<td>Sex, boys (%)</td>
<td>50%</td>
</tr>
<tr>
<td>Country of birth, the Netherlands: n (%)</td>
<td>58 (93.5%)</td>
</tr>
<tr>
<td>Sex responding parent, male (%)</td>
<td>9.8%</td>
</tr>
<tr>
<td>Age responding parent, years: mean (±SD)</td>
<td>41.8 (± 5.9)</td>
</tr>
<tr>
<td>Country of birth responding parent, the Netherlands: n (%)</td>
<td>55 (90.2%)</td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
</tr>
<tr>
<td>ALL-9 protocol: n (%)</td>
<td>21 (34%)</td>
</tr>
<tr>
<td>Non-high risk: n (%)</td>
<td>13 (62%)</td>
</tr>
<tr>
<td>High risk: N (%)</td>
<td>8 (38%)</td>
</tr>
<tr>
<td>ALL-10 protocol: n (%)</td>
<td>41 (66%)</td>
</tr>
<tr>
<td>Standard risk: n (%)</td>
<td>15 (37%)</td>
</tr>
<tr>
<td>Medium risk: n (%)</td>
<td>23 (56%)</td>
</tr>
<tr>
<td>High risk: n (%)</td>
<td>3 (7%)</td>
</tr>
</tbody>
</table>

IQR: interquartile range.

Sleep

CSHQ and ASHQ scores are summarized in Table 9.2. Parents rated the ALL survivors under the age of 12 years as having more bedtime resistance, daytime sleep and total sleep disturbances than the norm, with moderate effect sizes. In addition, there was a non-significant trend for the parent-rated ASHQ scores to be higher than the Dutch norm. In turn, ALL survivors under the age of 12 years themselves reported less sleep problems compared to the Dutch norm, with a moderate effect size. Except for the subscale bedtime, this trend could also be identified for adolescents. Socio-demographic and treatment variables had no effect on total sleep scores.
Table 9.3 shows PedsQL MFS scores. Effect sizes varied from moderate to large, with parents rating the ALL survivors as having more general fatigue and total fatigue than the norm. Fatigue reported by survivors themselves did not differ from the Dutch norm. Socio-demographic and treatment variables had no effect on total fatigue scores.

Depression

Table 9.4 shows CDI scores. Female ALL survivors reported less depressive symptoms compared to the Dutch norm. This trend could also be detected in male ALL survivors. Socio-demographic and treatment variables did not influence total depression scores.

QoL

CHQ scores are presented in Table 9.5. Parents rated the ALL survivors as having poorer physical QoL than the norm. Except for the general health perception, ALL survivors themselves reported better psychosocial QoL compared to the Dutch norm. Socio-demographic and treatment variables did not affect total QoL scores.
Table 9.3 Scale descriptives for the PedsQL™ Multidimensional Fatigue Scale (PedsQL MFS)

<table>
<thead>
<tr>
<th></th>
<th>ALL survivors</th>
<th>Dutch norm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PedSQL MFS</strong></td>
<td>n</td>
<td>Mean</td>
</tr>
<tr>
<td><strong>Child report</strong></td>
<td>35</td>
<td>82.38</td>
</tr>
<tr>
<td>General fatigue</td>
<td></td>
<td>82.38</td>
</tr>
<tr>
<td>Sleep/rest fatigue</td>
<td></td>
<td>77.50</td>
</tr>
<tr>
<td>Cognitive fatigue</td>
<td></td>
<td>76.31</td>
</tr>
<tr>
<td>Total fatigue</td>
<td></td>
<td>78.73</td>
</tr>
</tbody>
</table>

| **Parent report**    | 61            | 70.38      | 19.31 | 497   | 81.27 | 14.17 | <.001    | .77         |
| General fatigue      |               | 70.38      | 19.31 |       | 81.27 | 14.17 | <.001    | .77         |
| Sleep/rest fatigue   |               | 79.56      | 17.40 |       | 83.84 | 13.86 | .028     | .31         |
| Cognitive fatigue    |               | 72.58      | 24.00 |       | 78.48 | 17.93 | .067     | .33         |
| Total fatigue        |               | 74.25      | 17.94 |       | 81.21 | 12.62 | .004     | .55         |

Higher scores indicate fewer symptoms of fatigue.
* Independent samples t-tests.

Table 9.4 Scale descriptives for the Children's Depression Inventory (CDI)

<table>
<thead>
<tr>
<th></th>
<th>ALL survivors</th>
<th>Dutch norm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CDI</strong></td>
<td>n</td>
<td>Mean</td>
</tr>
<tr>
<td>Child report (girls)</td>
<td>18</td>
<td>4.94</td>
</tr>
<tr>
<td>Total score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child report (boys)</td>
<td>18</td>
<td>5.22</td>
</tr>
<tr>
<td>Total score</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Higher scores indicate more symptoms of depression.
*One-sample Kolmogorov-Smirnov test.

Table 9.5 Scale descriptives for the Child Health Questionnaire (CHQ)

<table>
<thead>
<tr>
<th></th>
<th>ALL survivors</th>
<th>Dutch norm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CHQ</strong></td>
<td>n</td>
<td>Mean</td>
</tr>
<tr>
<td>Child report</td>
<td>35</td>
<td>96.1</td>
</tr>
<tr>
<td>Physical functioning</td>
<td></td>
<td>95.9</td>
</tr>
<tr>
<td>Role functioning (emotional)</td>
<td></td>
<td>96.8</td>
</tr>
<tr>
<td>Role functioning (behavior)</td>
<td></td>
<td>97.1</td>
</tr>
<tr>
<td>Bodily pain</td>
<td></td>
<td>76.9</td>
</tr>
<tr>
<td>General behavior</td>
<td></td>
<td>83.1</td>
</tr>
<tr>
<td>Mental health</td>
<td></td>
<td>84.9</td>
</tr>
<tr>
<td>Self-esteem</td>
<td></td>
<td>80.9</td>
</tr>
</tbody>
</table>
Table 9.5 Scale descriptives for the Child Health Questionnaire (CHQ) (Continued)

<table>
<thead>
<tr>
<th></th>
<th>ALL survivors</th>
<th>Dutch norm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean</td>
</tr>
<tr>
<td>General health perception</td>
<td>66.8</td>
<td>11.2</td>
</tr>
<tr>
<td>Family activities</td>
<td>93.2</td>
<td>10.4</td>
</tr>
<tr>
<td>Family cohesion</td>
<td>83.0</td>
<td>16.0</td>
</tr>
<tr>
<td>Physical summary score</td>
<td>84.2</td>
<td>7.2</td>
</tr>
<tr>
<td>Psychosocial summary score</td>
<td>88.3</td>
<td>6.6</td>
</tr>
<tr>
<td>Parent report</td>
<td>61</td>
<td>14.04</td>
</tr>
<tr>
<td>Role functioning (emotional/behavior)</td>
<td>89.1</td>
<td>23.96</td>
</tr>
<tr>
<td>Role functioning (physical)</td>
<td>91.8</td>
<td>19.63</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>77.2</td>
<td>22.07</td>
</tr>
<tr>
<td>General behavior</td>
<td>78.1</td>
<td>13.35</td>
</tr>
<tr>
<td>Mental health</td>
<td>78.1</td>
<td>12.95</td>
</tr>
<tr>
<td>Self-esteem</td>
<td>77.1</td>
<td>14.10</td>
</tr>
<tr>
<td>General health perception</td>
<td>54.0</td>
<td>16.63</td>
</tr>
<tr>
<td>Parental impact (emotional)</td>
<td>73.0</td>
<td>20.45</td>
</tr>
<tr>
<td>Parental impact (time)</td>
<td>88.5</td>
<td>15.45</td>
</tr>
<tr>
<td>Family activities</td>
<td>87.0</td>
<td>16.58</td>
</tr>
<tr>
<td>Family cohesion</td>
<td>77.1</td>
<td>15.38</td>
</tr>
<tr>
<td>Physical summary score</td>
<td>47.7</td>
<td>9.41</td>
</tr>
<tr>
<td>Psychosocial summary score</td>
<td>51.4</td>
<td>8.32</td>
</tr>
</tbody>
</table>

Higher scores indicate better QoL.

<sup>a</sup>One sample t-tests.

Relation between sleep, fatigue, depression and QoL

Correlations between the sleep, fatigue, depression and QoL questionnaires are presented in Table 9.6. The CHQ physical summary score negatively correlated with CSHQ (both self-reports and parent-reports) and the CHQ psychosocial summary score negatively correlated with the CSHQ (self-reports) and the CDI, indicating poorer QoL in children with more sleep difficulties and in children with more symptoms of depression. In addition, the CHQ physical and psychosocial summary scores positively correlated with the PedsQL MFS (parent-reports), indicating poorer QoL in fatigued children. Furthermore, the PedsQL MFS negatively correlated with the CSHQ and the ASHQ (both parent-reports) and the CDI, indicating more fatigue in children with sleep difficulties and in children with more symptoms of depression.
This paper reports on sleep and fatigue in relation to depression and QoL after treatment for childhood ALL from the perspective of both the parents and the ALL survivors. According to parents, sleep was impaired after treatment for childhood ALL. However, survivors themselves reported less sleep problems compared to the Dutch norm. Information on sleep in childhood ALL survivors is scarce and the available information, which is solely based on self-reports, is conflicting. Mulrooney et al. reported no clinically meaningful differences in sleep between ALL survivors and a sibling control group. However, according to Meeske et al. nearly 50% of the ALL survivors reported sleep problems compared to 15-35% of the general population.6

Our results show that parents rated the ALL survivors as being more fatigued than the Dutch norm whilst fatigue reported by survivors themselves did not differ. Our self-report results correspond with previous studies that found no more self-reported fatigue in ALL survivors compared to siblings and the general population, respectively.6-8

Female ALL survivors reported less depressive symptoms in comparison to the Dutch norm references. This trend could also be detected in male ALL survivors. These findings are in accordance with Harila et al. who found less depressive symptoms among childhood ALL survivors in comparison to healthy controls and with Pemberger et al. who showed lesser
extents of depression in survivors of childhood cancer in comparison to the general population.\textsuperscript{33,34}

Further results demonstrate that parents rated the ALL survivors as having poorer physical QoL. ALL survivors themselves reported better psychosocial QoL. Reports in the literature on QoL among survivors of childhood ALL are inconsistent. Certain studies, mainly based on self-reports, demonstrated equal or better QoL in survivors of childhood ALL compared to controls or norm references\textsuperscript{35,36} while other studies demonstrated lower QoL\textsuperscript{37,38}. This discrepancy may at least partly be attributed to the use of self-reports versus proxy-reports.

Similar to certain previous studies regarding QoL in childhood cancer, parental ratings were significantly worse than childhood ALL survivor ratings.\textsuperscript{39,40} Survivors reported equal or even better scores compared to the Dutch norm. There are several possible explanations for these findings. First, although the child may be cured of cancer, the parents still have concerns about the child’s future which may be reflected in the measurement scores.\textsuperscript{41} Another clarification might be the coping mechanism of the child. Due to changes in the internal standard, the child may minimize physical or psychological symptoms which are considered irrelevant compared to the life-threatening illness that they experienced. This phenomenon has also been described as “response shift” or “framing”.\textsuperscript{42-44} Another coping style which is described in children with cancer is the “repressive adaptive style” involving denial, defensiveness, minimization and avoidance.\textsuperscript{43,45,46} Furthermore, it has been suggested that childhood cancer survivors are more willing to give social desirable answers.\textsuperscript{35,46} Finally, the experience of a life-threatening condition such as ALL is an opportunity for personal growth which may lead to an enhanced appreciation of life, also named “benefit finding”.\textsuperscript{46,47} Accordingly, professionals should be aware of the possibility that parents may underestimate the child’s QoL while childhood ALL survivors themselves may minimize their psychological and physical symptoms. Therefore, this study emphasizes the need for including both child self-reports and parent proxy-reports. In addition, it would be helpful to objectively measure sleep and fatigue after childhood ALL, for example, using actigraphy, polysomnography, or multiple sleep latency tests. Nevertheless, an essential outcome of the present study is that, despite the experience of a life-threatening disease and the long period of intensive treatment, survivors of childhood ALL seem to appreciate their lives well.

Another important finding of the current study is that proxy-reported sleep disturbances and fatigue inversely correlated with QoL and although ALL survivors reported less sleep problems and less fatigue than the Dutch norm, self-reported sleep disturbances negatively correlated with QoL and self-reported fatigue correlated with depression. This is in line with literature regarding sleep, fatigue, depression and QoL in survivors of other types of malignancies.\textsuperscript{10,12,16,17} Although we are not able to make causal inferences, improving sleep quality by, for example, behavioral and/or pharmacotherapeutic interventions and thereby managing fatigue, may improve QoL as well.
Some limitations of this study need to be mentioned. First, we had moderate response rates. However, these correspond with previous studies that sent postal invitations for questionnaire surveys without incentives in childhood cancer survivor. Moreover, participants and non-participants did not differ in socio-demographic and disease variables. Additionally, although we restricted our inclusion criteria to childhood ALL survivors under the age of 18, the age range of the children included in this study was rather wide. The CSHQ demonstrated moderate internal consistency reliability, yet it is the best available validated Dutch questionnaire regarding sleep disturbances in children. Further, the psychometric properties of the original version of the ASHQ have not yet been published. Hence, the present study reports on the reliability and validity of the Dutch version of this questionnaire. In addition, the number of subjects who completed the ASHQ was relatively small. Finally, the Dutch norm references of the parent form of the CHQ comprises children of 5-13 years of age, whereas our study group included children between 5 and 17 years of age. However, this was the best available reference group.

In conclusion, in line with our hypothesis, this study demonstrates that according to parents sleep, fatigue and QoL were substantially impaired in survivors after treatment for childhood ALL. Survivors of childhood ALL themselves, however, reported less sleep problems, less depressive symptoms, better psychosocial QoL and similar fatigue and physical QoL compared to the Dutch norm references. Discrepancies between parental and self-ratings might be explained by worried parents and/or the adaptive style of the children. Impaired sleep and fatigue were associated with more symptoms of depression and an impaired QoL, indicating a possibility for intervention to improve psychosocial health after surviving childhood ALL.
References


Chapter 10

Hypothalamic-pituitary-adrenal axis function in survivors of childhood acute lymphoblastic leukemia and healthy controls

M. Suzanne Gordijn, Raphaële R.L. van Litsenburg, Reinoud J.B.J. Gemke, Marc B. Bierings, Peter M. Hoogerbrugge, Peter M. van de Ven, Cobi J. Heijnen, Gertjan J.L. Kaspers

Psychoneuroendocrinology 2012;37:1448-56
Summary

Of all malignancies in children, acute lymphoblastic leukemia (ALL) is the most common type. Since survival significantly improves over time, treatment-related side effects become increasingly important. Glucocorticoids play an important role in the treatment of ALL, but they may suppress the hypothalamic-pituitary-adrenal (HPA) axis. The duration of HPA axis suppression is not yet well defined. The present study aimed at assessing the function of the HPA axis by determining the cortisol awakening response (CAR) and the dexamethasone (DEX) suppression test in children that were treated for childhood ALL, compared to a healthy age and sex matched reference group. In addition, questionnaires regarding sleep, fatigue, depression and quality of life were completed by the children and their parents. Forty-three survivors who finished their treatment for childhood ALL 37 (interquartile range 22-75) months before and 57 healthy controls were included. No differences in CAR were observed between ALL survivors and the reference group, but survivors of ALL had higher morning cortisol levels and an increased cortisol suppression in response to oral dexamethasone. Higher cortisol levels in childhood ALL survivors were associated with more fatigue and poorer quality of life. We conclude that the experience of a stressful life event in the past may have caused a long-term dysregulation of the HPA axis in childhood ALL survivors, as reflected in an increased cortisol production and an enhanced negative feedback mechanism.
Introduction

Of all malignancies in children, acute lymphoblastic leukemia is the most common type. Since treatment and survival rates of ALL significantly improve over time, treatment-related side effects become more and more important. Glucocorticoids, like prednisolone and dexamethasone, play an important role in the treatment of ALL. However, supraphysiologic doses of glucocorticoids may suppress the hypothalamic-pituitary-adrenal (HPA) axis function. Suppression of the HPA axis, resulting in reduced cortisol levels, may cause life-threatening hypoglycemia and/or hypotension, an increased pro-inflammatory response and an inadequate host-response against infections.

Long term suppression of the HPA axis has been reported after courses of high-dose glucocorticoids during treatment for ALL. However, the duration of the suppression has not yet been established accurately. Animal studies on long-term effects of exposure to dexamethasone during early life have reported a reduction of HPA axis activity later in life. In addition, neonatal treatment with dexamethasone for chronic lung disease of prematurity has been associated with a blunted HPA axis activity in children at school age. Since treatment with high-dose glucocorticoids early in life may induce long-lasting side effects, emerging concern has risen about the long-term neuroendocrine sequelae of high-dose glucocorticoid treatment in childhood ALL. Nevertheless, HPA axis activity in survivors of childhood ALL has not been studied before.

As the survival rates of children with ALL have dramatically increased over time, it is essential to gain additional insights about the possible health risks associated with treatment. A reliable marker for HPA axis activity is the cortisol increase occurring within 30-45 min after awakening. This sharp rise in cortisol upon awakening is present in 75% of healthy adults. The cortisol awakening response (CAR) has the advantage of stress-free sampling and a high intra-individual stability and it is not influenced by sleep duration or time of awakening. Due to the negative feedback of dexamethasone, the CAR is strongly inhibited after intake of this synthetic glucocorticoid. Therefore, the low dose (0.5 mg) dexamethasone (DEX) suppression test can be used to test the integrity of the HPA axis, enabling differentiation between normal and enhanced suppression.

The aim of the present study was to assess the HPA axis regulation in children that received high-dose glucocorticoids for the treatment of childhood ALL, compared to a healthy age and sex matched reference group. We hypothesized that the HPA axis activity of childhood ALL survivors was reduced, based on previous studies regarding the long-term effects of high-dose glucocorticoids given early in life.

Considering the frequently reported problems with sleep, fatigue and quality of life in childhood cancer survivors and the increasing evidence for associations between these impairments and cortisol levels, we also collected information regarding sleep, fatigue, depression and quality of life using validated questionnaires.
Methods

Participants
The study population consisted of children that were treated according to the Dutch Childhood Oncology Group (DCOG) ALL-9 or ALL-10 protocol between May 1997 and February 2008 in the VU University Medical Center Amsterdam, the University Medical Center Utrecht or the Radboud University Nijmegen Medical Center in the Netherlands. The ALL-9 protocol comprised of a four-week induction course of high-dose dexamethasone (6 mg/m²) and consecutive cyclic courses of dexamethasone (6 mg/m²) during maintenance treatment for both Non-high risk (NHR) and High risk (HR) patients.²⁵ The ALL-10 protocol consists of a four-week induction course of prednisolone (60 mg/m²) for all patients. Standard Risk (SR) patients received an additional two-week course of dexamethasone (10 mg/m²), Medium Risk (MR) patients received 84 weeks of additional cyclic courses of dexamethasone (6 mg/m²) and High Risk (HR) patients received an additional three-week course of dexamethasone (10 mg/m²) during maintenance therapy.

The reference group consisted of age (in years) and sex matched friends of the ALL survivors and children without a chronic health condition that were recruited from the orthopedic surgery and the ear, nose and throat outpatient clinic of the VU University Medical Center, Amsterdam, the Netherlands. Exclusion criteria were treatment with glucocorticoids for other reasons than ALL, dysfunction of the hypothalamus, the pituitary and/or the adrenals and being under treatment for relapsed ALL. This study was approved by our Institutional Review Board and written informed consent was obtained from the children and their caregivers.

Cortisol assessment and analysis
Cortisol was determined in saliva. Children and their parents were provided with detailed written instructions for collecting saliva at home. Subjects were instructed to collect saliva samples over one day (study day 1) immediately after awakening and 15 and 30 min thereafter (while lying in bed) and at 2100 h. Children and their parents filled in a daily log reporting bedtimes, wake-up times and the precise times at which they collected their samples. Subjects were instructed not to eat or drink, except for water, and not to brush their teeth before completing saliva sampling in the morning to avoid contamination of saliva with food or blood caused by micro-injuries of the oral cavity. Fifty-six of the 100 subjects also performed a dexamethasone (DEX) suppression test. They collected saliva over the consecutive day (study day 2) following the same saliva collection protocol after oral administration of 0.25 mg dexamethasone at 2100 h at the night before.

Due to non-compliance to the sampling instructions, 92% of the participants on the first study day and 95% of the participants on the second study day reported a delay in saliva collection in relation to waking, which resulted in a third morning sample of one minute to
80 min too late. As preliminary results indicated a tendency of the cortisol levels of our participants to further increase between 30 and 40 min after awakening, we decided to analyze all cortisol samples within the first 40 min after awakening. By expanding this time period, the number of cortisol samples to analyze increased with 43 on the first study day and with 30 on the second study day.

Subjects stored the saliva samples in their freezer until completing the sampling and then returned the samples to the hospital. Samples were stored at -20°C until analysis with a commercially available competitive luminescence immunoassay (IBL-International, Hamburg, Germany) of all samples in one batch. The intra- and inter-assay coefficients of variance were both below 8%.

**Questionnaires**

Sleep, fatigue and quality of life were evaluated by parent-proxy and child-self reports of the following valid and reliable questionnaires: the Children's Sleep Habits Questionnaire (CSHQ) or the Adolescents Sleep Habits Questionnaire (ASHQ) for children above the age of 12 years, the PedsQL™ multidimensional fatigue scale (PedsQL MFS) and the Child Health Questionnaire (CHQ). The reliable and valid Children's Depression Inventory (CDI), only available as a self-report, was included for the potential modifying effect of depression. All total scores were compared to Dutch norms.

Because of the potential modifying effect of the socioeconomic status on the cortisol levels, we collected additional information regarding the educational level of both parents of the participants.

**Statistical analysis**

CAR was analyzed by generalized estimating equations (GEE) analysis. This analysis is an extension of ordinary linear regression that can take into account intra-individual correlation between measurements. The GEE analysis allows for the inclusion of collection time as a continuous variable in the model, thereby allowing different collection times for the subjects. To assess whether the CAR on the first and the second study day varied between both study groups (ALL survivors vs reference group), we included a linear and quadratic term for time since awakening plus their interaction with study group in the fixed effects regression equation. To examine whether the influence of the dexamethasone suppression test at the second day on the CAR varied between both study groups, we included for all terms (time since awakening and study group) the interaction with factor day in the regression equation.

Previous studies on the CAR did not investigate the actual sampling times and used the one-way ANOVA and the area under the curve for analysis of the CAR instead of the generalized estimating equations or the linear mixed model analysis. In order to facilitate comparison of our results with previous studies, CAR was also analyzed by one-way ANOVA and by area under the curve with respect to the ground (AUCg), which is a measure of the
total cortisol secretion, and with respect to the increase (AUCi), which is a measure of the dynamics of cortisol secretion.\textsuperscript{42}

Within-group differences were assessed by t-tests for continuous variables, and by chi-square test for categorical variables. Educational level of the parents was classified as low education (no education, primary school and primary vocational education), moderate education (secondary school and secondary vocational education) and high education (higher vocational education and university). Correlations between cortisol levels and continuous variables were determined using Pearson correlation.

All statistical analyses were performed with SPSS (Version 15.0.1, Chicago, IL). A two-sided $p$ value of <0.05 was accepted as statistically significant.

**Results**

**Demographics**

A total of 43 ALL survivors and 57 healthy controls were included in the study (Figure 10.1). The most common reasons for ALL survivors for declining participation were resistance against the DEX suppression test due to negative associations with dexamethasone therapy

![Figure 10.1 Overview of inclusion of ALL survivors and healthy controls.](attachment:image.png)
during treatment for childhood ALL and reluctance to be confronted again with the disease period. Participants and non-participants did not differ significantly in age, gender, treatment protocol, risk group stratification and time since end of treatment. Sociodemographic characteristics of all participating subjects are shown in Table 10.1. Of all participating ALL survivors, 35% were treated according to the ALL-9 protocol, of which 53% was classified as non-high risk and 47% as high risk, and 65% of the participating ALL survivors were treated according to the ALL-10 protocol of which 43% was classified as standard risk, 46% as medium risk and

<table>
<thead>
<tr>
<th>Table 10.1 Sociodemographic characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Study day 1</td>
</tr>
<tr>
<td>n = 42</td>
</tr>
<tr>
<td>Age, years (mean ± SD)</td>
</tr>
<tr>
<td>10.6 ± 3.9</td>
</tr>
<tr>
<td>Gender, boys (%)</td>
</tr>
<tr>
<td>52%</td>
</tr>
<tr>
<td>Education mother, high (%)</td>
</tr>
<tr>
<td>20 (48%)</td>
</tr>
<tr>
<td>Education father, high (%)</td>
</tr>
<tr>
<td>18 (45%)</td>
</tr>
<tr>
<td>Time off-treatment, months (median (IQR))</td>
</tr>
<tr>
<td>37.0 (22-75)</td>
</tr>
<tr>
<td>Treatment:</td>
</tr>
<tr>
<td>ALL-9 protocol, n (%)</td>
</tr>
<tr>
<td>14 (67%)</td>
</tr>
<tr>
<td>Non-high risk, n (%)</td>
</tr>
<tr>
<td>7 (50%)</td>
</tr>
<tr>
<td>High risk, n (%)</td>
</tr>
<tr>
<td>7 (50%)</td>
</tr>
<tr>
<td>ALL-10 protocol, n (%)</td>
</tr>
<tr>
<td>28 (33%)</td>
</tr>
<tr>
<td>Standard risk, n (%)</td>
</tr>
<tr>
<td>12 (43%)</td>
</tr>
<tr>
<td>Medium risk, n (%)</td>
</tr>
<tr>
<td>13 (46%)</td>
</tr>
<tr>
<td>High risk, n (%)</td>
</tr>
<tr>
<td>2 (11%)</td>
</tr>
<tr>
<td>Study day 2</td>
</tr>
<tr>
<td>N = 37</td>
</tr>
<tr>
<td>Age, years (mean ± SD)</td>
</tr>
<tr>
<td>11.0 ± 3.9</td>
</tr>
<tr>
<td>Gender, boys (%)</td>
</tr>
<tr>
<td>51%</td>
</tr>
<tr>
<td>Education mother, high (%)</td>
</tr>
<tr>
<td>16 (43%)</td>
</tr>
<tr>
<td>Education father, high (%)</td>
</tr>
<tr>
<td>15 (43%)</td>
</tr>
<tr>
<td>Time off-treatment, months (median (IQR))</td>
</tr>
<tr>
<td>39.0 (25.0-80.0)</td>
</tr>
<tr>
<td>Treatment:</td>
</tr>
<tr>
<td>ALL-9 protocol, n (%)</td>
</tr>
<tr>
<td>15 (41%)</td>
</tr>
<tr>
<td>Non-high risk, n (%)</td>
</tr>
<tr>
<td>8 (53%)</td>
</tr>
<tr>
<td>High risk, n (%)</td>
</tr>
<tr>
<td>7 (47%)</td>
</tr>
<tr>
<td>ALL-10 protocol, n (%)</td>
</tr>
<tr>
<td>22 (59%)</td>
</tr>
<tr>
<td>Standard risk, n (%)</td>
</tr>
<tr>
<td>9 (41%)</td>
</tr>
<tr>
<td>Medium risk, n (%)</td>
</tr>
<tr>
<td>10 (45%)</td>
</tr>
<tr>
<td>High risk, n (%)</td>
</tr>
<tr>
<td>3 (14%)</td>
</tr>
</tbody>
</table>

IQR: 25% to 75% interquartile range
11% as high risk. The median time since end of treatment was 37 (interquartile range 22-75) months. Groups were age and sex matched on both study days.

**CAR**

No differences in CAR were observed between ALL survivors and the reference group (group \( x \) duration \( x \) duration, \( P = 0.388 \) and \( P = 0.661 \), on study day 1 and 2, respectively) (Figure 10.2 and Figure 10.3). Nevertheless, from approximately 28 min after awakening on the first study day mean cortisol levels appeared to differ between both study groups as confidence intervals did not overlap (Figure 10.2).

The potential confounders age, gender, educational level of the parents, treatment protocol and time after treatment did not affect the regression equation. Analysis of all the cortisol data, not restricted to the first 40 min after awakening, revealed the same results.

In order to compare our results with previous studies, CAR was also analyzed without correction for the actual sampling time. The ALL survivors showed a higher AUCg, reflecting the overall cortisol output, in comparison to the reference group (mean difference 85.39, \( P = 0.014 \)). The AUCi, demonstrating cortisol dynamics, did not show group differences (mean difference 38.26, \( P = 0.187 \)). One way ANOVA at the different time points showed significant group differences at 15 min (ANOVA: \( F_{1,99} = 6.02, P = 0.016 \)) and 30 min (ANOVA: \( F_{1,99} = 6.48, P = 0.012 \)) after awakening, with higher cortisol levels for the ALL survivor group. In addition, ALL survivors had higher morning cortisol levels in comparison to the reference group (mean dif-

![Figure 10.2](image-url)

**Figure 10.2** Cortisol awakening response (CAR) on study day 1.

--- represents the ALL survivor group and ----- represents the reference group.

The fat lines represent the mean cortisol levels and the thin lines represent the 95% confidence intervals.
ference in mean morning cortisol level: 2.65, \( P = 0.013 \); mean difference in maximum morning cortisol level: 3.36, \( P = 0.014 \); mean difference in the peak minus the nadir morning cortisol level: 2.20, \( P = 0.043 \). Time since end of treatment for ALL was not associated with mean or maximum morning cortisol levels. No associations were found between the educational level of the parents of both study groups and morning cortisol levels of the participants.

**CAR after the DEX suppression test**

Survivors of ALL showed significantly more cortisol suppression in response to the DEX suppression test compared to the reference group (group x day, \( P < 0.001 \)) (Figure 10.3). Analysis of all the cortisol data, not restricted to the first 40 min after awaking, revealed the same results.

![Cortisol awakening response (CAR) after the DEX suppression test on study day 2.](image)

---

**Relation between cortisol and fatigue, sleep and health-related quality of life**

In ALL survivors, peak cortisol levels negatively correlated with all subscale scores and the total score of the parent-reports of the PedsQL MFS (“general fatigue”: \( r = -0.38, P = 0.012 \), “sleep/rest fatigue”: \( r = -0.31, P = 0.046 \), “cognitive fatigue”: \( r = -0.45, P = 0.03 \), “total fatigue”: \( r = -0.46, P = 0.002 \)), indicating more symptoms of fatigue in ALL survivors with higher cortisol levels. In addition, peak cortisol levels negatively correlated (indicating worse scores) with the subscales “physical functioning” (\( r = -0.49, P = 0.001 \), “behaviour” (\( r = -0.31, P = 0.047 \), “mental
health” ($r$ -0.32, $P = 0.040$), “general health” ($r$ -0.40, $P = 0.009$), “parent impact” ($r$ -0.41, $P = 0.007$), “family activities” ($r$ -0.33, $P = 0.030$), “physical total score” ($r$ -0.39, $P = 0.012$) and “psychosocial total score” ($r$ -0.33, $P = 0.036$) of the parent-reports of the CHQ. Peak cortisol levels were not associated with self-report scores of the PedsQL MFS, CHQ, CSHQ or ASHQ and CDI and not with the parent-report scores of the CSHQ or ASHQ.

In healthy controls, peak cortisol levels did not correlate with parent-report or self-report scores of the fatigue, sleep and health-related quality of life questionnaires.

Discussion

This study identified significant differences in HPA axis function between survivors of childhood ALL and healthy controls. Contrary to our hypothesis, survivors of ALL had higher morning cortisol levels than the reference group. The CAR did not differ between both study groups. A possible explanation for the increased morning cortisol production in survivors of childhood ALL may be that ALL survivors suffered from stress. In this context it is of interest that distress has been found to contribute to higher cortisol levels in adult cancer patients. In addition, cortisol levels have been reported to decrease in adult breast cancer patients after interventions to reduce distress, i.e. an integrated yoga program. To the best of our knowledge, no information regarding the occurrence of stress in survivors of childhood ALL is available, except for the fact that pediatric cancer patients of distressed parents are more vulnerable to stress than pediatric cancer patients of low stressed parents, although non-environmental factors such as genetic predisposition cannot be ruled out. Moreover, research has shown that stressful life events in the past may result in long-term dysregulation of the HPA axis. Although the majority of studies in adults with posttraumatic stress disorder reported decreased cortisol levels, most studies in children who experienced a stressful life event demonstrated increased cortisol levels. This difference in HPA axis activity may be due to a differential response to severe stress between children and adults.

It is conceivable that the experience of a life-threatening disease as ALL and the intensive treatment with chemotherapy may account for a stressful life event as well. Hyperresponsiveness of the HPA axis may reflect an adaptive biological alteration that enables children to better cope with the stressful experiences.

The clinical implications of the detected increased morning cortisol production in survivors of childhood ALL need to be elucidated. Studies in trauma exposed children have not only shown enhanced cortisol levels, but also increased sympathetic nervous system activity. Sympathetic overactivity in turn, is associated with the development of the metabolic syndrome. As survivors of childhood ALL are at increased risk for the metabolic syndrome, the relation between enhanced cortisol levels and sympathetic nervous system activation in ALL survivors may be of specific interest for future studies.
The present study also showed an enhanced cortisol suppression in response to oral dexamethasone in the ALL survivor group. This observation is in line with previous research regarding traumatized veterans with or without a posttraumatic stress disorder.\textsuperscript{40} The detected enhanced negative feedback mechanism further supports the suggestion of trauma related alteration in HPA axis function in survivors of childhood ALL. A potential explanation is an up-regulation of the pituitary glucocorticoid receptors or of the glucocorticoid receptor sensitivity, resulting from stress.\textsuperscript{40, 64}

This study also showed that higher cortisol levels in childhood ALL survivors, but not in healthy controls, were associated with more fatigue and poorer quality of life as rated by the parents. The observed association between cortisol and fatigue is in contrast to previous studies which described decreased cortisol levels in fatigued adults.\textsuperscript{21-23} This incongruity is difficult to explain and requires further exploration. The sparse evidence in the literature regarding the association between cortisol levels and quality of life is conflicting.\textsuperscript{65-67} Cortisol levels were not related to self reported and proxy reported indicators of sleep. Reports in the literature about the association between sleep and cortisol levels are inconsistent.\textsuperscript{20, 68-70} Moreover, no associations emerged between cortisol levels and self-reported sleep, fatigue, quality of life and depression. However, it should be noted that survivors of childhood ALL did report similar or even better scores on these questionnaires in comparison to Dutch norm scores. Differences in parental and self ratings have frequently been reported and might be explained by worried parents and/or an adaptive style of the children.\textsuperscript{71-73} Therefore, the clinical relevance of the associations that were found between cortisol levels and fatigue and quality of life rated by the parents remains unclear.

Our results demonstrated low participant adherence to the saliva collection protocol. Previous studies did not report on the actual cortisol collection times, but might have equally suffered from a similar low adherence.\textsuperscript{38-41} Future studies would benefit from collecting information about the actual saliva sampling times in order to determine the amount of adherence to the saliva collection protocol and to decide whether AUC’s can be estimated. If reliable estimation of AUC’s is not possible or if sampling times differ between the groups that are compared, statistical methods that model the response over time may be preferred.

Some limitations of this study need to be mentioned. First, the degree of stress perceived by the participants during both study days was not evaluated. Survivors of childhood ALL might have experienced more stress in comparison to the healthy controls resulting in higher morning cortisol levels. Second, we had low response rates, which might lead to non-response bias. Although participants and non-participants did not differ in demographic and disease specific variables, non-participants might have been more distressed and therefore decided not to participate in this study. In addition, it would have been interesting if survivors of childhood cancer who were not treated with high-dose glucocorticoid therapy were also included. This would be helpful to explore the long-term influence of high-dose
glucocorticoid therapy, apart from the stressful life-event of childhood cancer, on the HPA axis regulation.

In conclusion, the present study showed that survivors of childhood ALL had higher morning cortisol levels and an enhanced cortisol suppression in response to dexamethasone compared to healthy controls. These findings may indicate that the experience of stressful life events in the past, i.e. childhood ALL, may lead to long-term HPA axis alterations.
References


Chapter 11

Adequate endocrine and cardiovascular response to social stress in survivors of childhood acute lymphoblastic leukemia

M. Suzanne Gordijn, Reinoud J.B.J. Gemke, Marc B. Bierings, Peter M. Hoogerbrugge, Marijke D.J. Tersteeg-Kamperman, Cobi J. Heijnen, Joost Rotteveel, Gertjan J.L. Kaspers

Summary

Survivors of childhood ALL have been demonstrated to have increased morning cortisol levels compared to healthy controls. Information regarding the response of the HPA axis and the sympathetic nervous system to stress in childhood ALL survivors is not available. The present study aimed at assessing the endocrine and cardiovascular stress response in childhood ALL survivors and healthy controls by evaluating perceived stress on visual analogue scales, by determining saliva cortisol, blood pressure and heart rate in response to the Trier Social Stress Test for Children (TSST-C). Fifty survivors who had completed their treatment for childhood ALL 57 (IQR 47.0-72.3) months before and 50 healthy age and sex matched controls were included. Exposure to the TSST-C induced a significant response of perceived stress, saliva cortisol and cardiovascular outcome variables in the total study group. These responses did not significantly differ between survivors of childhood ALL and healthy controls. We conclude that the endocrine and cardiovascular response to social stress are intact in survivors of childhood ALL.
Introduction

Acute lymphoblastic leukemia (ALL) is the most common malignancy in childhood. Significant alterations in HPA axis function of survivors of childhood ALL have been identified. We have shown before that survivors of ALL have higher morning cortisol levels and an enhanced cortisol suppression in response to oral dexamethasone compared to healthy controls. In addition, a recent study demonstrated higher plasma cortisol levels as well as a higher cortisol amplitude in the response to the Trier Social Stress Test (TSST) in adult survivors of childhood cancer in comparison to healthy controls. Since stressful life events early in life may result in long-term dysregulation of the HPA axis, it is conceivable that the experience of a life-threatening disease such as cancer and the intensive treatment with chemotherapy may account for a stressful life event as well. Studies in trauma exposed children have not only shown enhanced cortisol levels, but also increased sympathetic nervous system activity. Increased sympathetic outflow may have clinical implications, since it is associated with development of metabolic syndrome. In addition, an altered cardiovascular response to stress is an early predictor of an increased risk for cardiovascular disease in later life. To date, no studies have investigated sympathetic nervous system activity in survivors of childhood ALL. Therefore, we now assessed the endocrine and cardiovascular response to a psychosocial stressor in survivors of childhood ALL by the TSST for children (TSST-C). Referring to previous studies regarding the stress response in traumatized subjects, we hypothesized that survivors of childhood ALL show an enhanced endocrine and cardiovascular response in comparison to healthy controls.

Methods

Participants

As this concerns a follow-up study, all 45 survivors of childhood ALL who participated in our previous study on HPA axis function in childhood ALL survivors were invited to participate in the present study and six of them declined. Eleven additional patients who initially responded positively to our previous study but ultimately refused to participate were invited for the present study and agreed to participate. The study population consisted of children that had been treated according to the Dutch Childhood Oncology Group (DCOG) ALL-9 or ALL-10 protocol between May 1997 and February 2008 in the VU University Medical Center Amsterdam, the University Medical Center Utrecht or the Radboud University Nijmegen Medical Center in the Netherlands. The control group consisted of age (in years) and sex matched friends of the ALL survivors. Exclusion criteria were treatment with glucocorticoids at the time of the study, dysfunction of the hypothalamus, the pituitary and/or the adrenals and
being under treatment for relapsed ALL. This study was approved by our Institutional Review Board and written informed consent was obtained from the children and their caregivers.

**Procedure**

For stress induction children were exposed to the TSST-C which was described in detail by Buske-Kirschbaum et al. After a clinical interview and a physical examination of 30 minutes, the TSST-C was performed which consisted of a 30-minute relaxation period watching a video, a 10-minute preparation period, a 5-minute public speech task, and a 5-minute age-appropriate mental arithmetic task. The latter two tasks were carried out in front of a “jury” judging the child’s performance. This was followed by a 10-minute debriefing period, during which the child was praised for his/her excellent performance, and by another 50-minute relaxation period watching a video (Figure 11.1). All TSST-C sessions started at 1300h and were performed by the same research nurse.

The perceived stress of the TSST-C was analyzed by a visual analogue scale (VAS). At repeated time intervals (t=0, t=60, t=70, t=80, t=90 and t=140 minutes) (Figure 11.1), participants indicated their perceived stress, by placing a mark on a 10-cm horizontal line, the left end of the line being labelled “not at all” (0) and the right end “extremely” (10). In order to evaluate the endocrine stress response to the TSST-C, saliva samples for cortisol were collected at repeated time intervals (t=60, t=70, t=80, t=90 and t=140 minutes) (Figure 11.1). Cardiovascular determinants were continuously monitored by the Nexfin (BMEYE B.V., Amsterdam, the Netherlands). The Nexfin device non-invasively registers the pressure waveform in the finger, and provides measurements of the beat-to-beat brachial blood pressure and heart rate and the pulse contour analysis-based estimated stroke volume, cardiac output and systemic

---

**Schedule for the TSST-C**

![Schedule of the TSST-C](image)

**Figure 11.1** Schedule of the TSST-C.
vascular resistance. The Nexfin monitor has been validated clinically and is considered to be a reliable method for noninvasive blood pressure measurement.\textsuperscript{13}

**Cortisol assessment and analysis**

Saliva cortisol samples were stored at -20\(^\circ\)C until analysis with a commercially available competitive luminescence immunoassay (IBL-International, Hamburg, Germany) of all samples in one batch.\textsuperscript{14} The intra- and inter-assay coefficients of variance were both below 8%.

**Statistical analysis**

Within-group differences were assessed by t-tests for continuous variables, and by chi-square test for categorical variables. For analysis of the effect of the stressor on the VAS scores and the endocrine and cardiovascular outcome variables, repeated-measurement analysis of variance (ANOVA) was used. VAS scores, cortisol data and cardiovascular outcome variables were log-transformed before statistical analysis to obtain normal distribution, whilst untransformed values are presented in Figure 11.2. Data regarding cortisol levels were also analyzed by area under the curve with respect to the ground (AUCg), which is a measure of the total cortisol secretion, and with respect to the increase (AUCi), which is a measure of the dynamics of cortisol secretion.\textsuperscript{15} In addition, the stress-induced cortisol increase was assessed by peak level minus nadir level, peak level minus baseline level and peak level minus return level. Cardiovascular data were analyzed by Nexfin software (BMEYE B.V., Amsterdam, the Netherlands), which provided measurements of systolic (SYS), diastolic (DIA), and mean arterial pressure (MAP), heart rate (HR), stroke volume (SV), cardiac output (CO), and systemic vascular resistance (SVR). Mean cardiovascular outcome variables were calculated for the different phases of the TSST-C: (1) the 30-minute relaxation period, (2) the 10-minute explanation and preparation phase, (3) the 5-minute speech task, (4) the 5-minute arithmetic task, (5) the 10-minute debriefing period, and (6) the 50-minute relaxation period. All statistical analyses were performed with SPSS (Version 20, Chicago, IL). A two-sided p value of <.05 was accepted as statistically significant.

**Results**

**Demographics**

A total of 50 ALL survivors and 50 healthy age and sex matched controls were included in the study. No differences in age, gender, treatment protocol, risk group stratification and time since end of treatment were observed between survivors of childhood ALL who participated in the current study and the six non-participants. Mean age of both study groups was 12.3 (± 3.9) years. Fifty percent of the ALL survivors and 48% of the healthy controls were boys. Sociodemographic characteristics (age, gender and education of the parent), anthropometrics
(weight, height and body surface area), blood pressure and heart rate did not differ between the study groups. For childhood ALL survivors, the median time after treatment was 57 (IQR 47.0-72.3) months.

**VAS score**

Exposure to the TSST-C induced a significant VAS score response in the total study group ($P <0.001$) (Figure 11.2). VAS scores at the six different evaluation time points did not show significant differences between ALL survivors and healthy controls. There was no significant group effect ($P = 0.766$) and no time x group interaction ($P = 0.846$).

**Endocrine response**

In the total study group exposure to the TSST-C induced a significant cortisol response ($P <0.001$) (Figure 11.3). Cortisol levels at the five different sampling time points were not different between ALL survivors and controls. There was no group effect ($P = 0.704$) and no time x group interaction ($P = 0.169$). Analysis of the AUCg and the AUCi during the entire period of cortisol measurement revealed no significant differences between both study groups ($P = 0.423$ and $P = 0.106$, respectively). The stress-induced cortisol increase (peak cortisol level at $t = 90$ minus nadir cortisol level at $t = 70$ and peak cortisol level at $t = 90$ minus baseline cortisol level at $t = 60$) and the recovery of cortisol levels after stress (peak cortisol level at $t = 90$ minus return cortisol level at $t = 140$) did not differ significantly between both study groups ($P = 0.275$, $P = 0.134$ and $P = 0.353$, respectively).
No associations were found between age, gender, time since end of treatment, treatment protocol and educational level of the parent on the one hand and cortisol dynamics on the other hand.

**Cardiovascular response**

In the total study group exposure to the TSST-C induced a significant blood pressure (SYS, DIA and MAP) response (all three $P < 0.001$). Mean blood pressure variables at the six time points of the TSST did not differ between ALL survivors and healthy controls. There was no group effect ($P = 0.363$, $P = 0.668$, $P = 0.280$, respectively) and no time x group interaction ($P = 0.903$, $P = 0.998$, $P = 0.582$, respectively).

Exposure to the TSST-C induced a significant cardiovascular response (HR, CO, SV and SVR) in the total study group (HR, CO and SV $P < 0.001$ and SVR $P = 0.014$). No differences in the mean cardiovascular outcome variables during the six different phases of the TSST-C were found between both study groups. There was no group effect ($P = 0.479$, $P = 0.955$, $P = 0.695$, $P = 0.998$, respectively) and no time x group interaction ($P = 0.311$, $P = 0.191$, $P = 0.456$, $P = 0.207$, respectively).
Discussion

The current study shows that the endocrine and cardiovascular response to the TSST-C do not differ between survivors of childhood ALL and healthy controls. Our previous study demonstrated increased morning cortisol production in childhood ALL survivors in comparison to healthy controls.\(^1\) Research has shown that traumatic life events in the past, such as childhood malignancy, may lead to alterations in basal as well as in stress-induced HPA axis activity later in life.\(^2\)\(^-\)\(^6\) We performed this follow-up study to further elucidate the HPA axis responsiveness in childhood ALL survivors. The current study demonstrates that the HPA axis reactivity to social stress is not altered in survivors of childhood ALL. This is in contrast to Laufer et al. who found an enhanced endocrine response to the TSST in cancer survivors.\(^2\) However, Laufer et al. included adult survivors of childhood cancer, whilst the mean age of our study group was 12.3 years. In addition, Laufer et al. included a heterogeneous study group comprising survivors of different types of childhood cancer, whilst our study exclusively included survivors of childhood ALL.

As previous studies in trauma exposed children did not only show enhanced cortisol levels, but also increased sympathetic nervous system activity which in turn may lead to metabolic and cardiovascular diseases later in life,\(^7\)\(^-\)\(^9\) we also studied the cardiovascular response to social stress in childhood ALL survivors. The present study demonstrates that the cardiovascular response to social stress is not altered in childhood ALL survivors. This suggests that the risk to develop metabolic and cardiovascular diseases related to an increased sympathetic nervous system activity is not enhanced in survivors of childhood ALL. As stated by Claessens et al. the impact of early life adversity on HPA axis responsiveness is complex due to environmental and genetic influences. In addition, it is timing and dose-dependent and therefore difficult to predict.\(^16\) Moreover, HPA axis responsiveness can be evaluated on different levels. Our previous study indicates that the experience of childhood ALL alters basal morning cortisol production.\(^1\) The current study demonstrates that the experience of childhood ALL does not alter the cortisol response to social stress.

Some limitations of our study need to be mentioned. First, as depression has been described to be associated with increased global cortisol levels in cancer survivors,\(^2\) it would have been interesting if data on depression was collected. However, in our previous study survivors of childhood ALL reported less symptoms of depression compared to Dutch norms.\(^1\) Furthermore, the number of participants was relatively small. However, because of the lack of information on the endocrine and cardiovascular stress response in children that were treated for ALL, our results provide important information.

In conclusion, although previous research demonstrated enhanced basal morning cortisol production in childhood ALL survivors, the findings of the current study indicate that survivors of childhood ALL have an adequate endocrine and cardiovascular response to social stress.
References


Chapter 12

General discussion and future perspectives
Short term neuroendocrine consequences of treatment for childhood ALL and lymphoma

In this thesis, it has become increasingly clear that adrenal insufficiency frequently occurs during treatment for childhood ALL and lymphoma. As reflected in our systematic literature review (Chapter 3) and our clinical studies (Chapter 4 and 5), the majority of children treated for ALL or lymphoma suffer from adrenal insufficiency in the first days to weeks after cessation of glucocorticoid therapy.

In the study described in this thesis, we showed that the median time to recover from adrenal insufficiency after a four-week induction course with prednisolone for childhood ALL was 28 days. The majority of patients recovered within five weeks after cessation of induction therapy. However, one-third of the patients suffered from ongoing adrenal insufficiency at the end of the follow-up period of 96 days. Moreover, ongoing adrenal insufficiency after induction therapy was associated with adrenal insufficiency during maintenance therapy (Chapter 4). Furthermore, lymphoma patients with adrenal insufficiency in between consecutive glucocorticoid courses, suffered from prolonged adrenal insufficiency after cessation of the final glucocorticoid course (Chapter 5).

Similar to previous studies, the recovery time of adrenal function after cessation of glucocorticoids for childhood ALL or lymphoma showed considerable inter-individual variation. Mahachoklertwattana et al. showed that three out of 24 children with ALL had adrenal insufficiency at the end of the follow-up period of 24 weeks after induction therapy. In addition, Einaudi et al. demonstrated that five out of 40 children with ALL had an insufficient cortisol response 28 days after induction therapy with prednisolone and three out of 24 children had an insufficient cortisol response 28 days after induction therapy with dexamethasone. Previous research identified that certain polymorphisms of the glucocorticoid receptor (GR) gene (NR3C1) result in alterations of sensitivity to glucocorticoids in vivo. We therefore explored a possible association between the duration of adrenal insufficiency after induction therapy in childhood ALL patients and the GR polymorphisms with previously reported altered sensitivity for glucocorticoids and sufficient prevalence in the general population (ER22/23EK, GR-β, N363S and BclI) (Chapter 6). We demonstrated that the duration of adrenal insufficiency in carriers of the ER22/23EK polymorphism, which is associated with glucocorticoid resistance, was significantly shorter in comparison to noncarriers. This suggests that ER22/23EK carriers have a lower susceptibility for developing adrenal insufficiency after high-dose glucocorticoid therapy for childhood ALL. In addition, we demonstrated that adrenal insufficiency in patients homozygous for BclI, which is associated with increased glucocorticoid sensitivity, was significantly longer in comparison to patients that were not homozygous for BclI. This indicates that patients homozygous for BclI have an increased susceptibility for developing prolonged adrenal insufficiency during treatment for childhood ALL. Another factor which was associated with differences in the duration of adrenal insufficiency was gender. We found
that the duration of adrenal insufficiency was significantly longer in boys in comparison to girls with ALL, irrespective of menarcheal state (Chapter 4). A possible explanation for this finding is the interference of sex hormones with the HPA axis. Estrogens have been described to enhance the HPA axis response\textsuperscript{11-14}, whilst androgens inhibit the HPA axis response\textsuperscript{15-17}. Another potential explanation for the difference in duration of adrenal insufficiency between boys and girls might be gender specific effects of GR polymorphisms on adrenal function.\textsuperscript{3} Because this is the first study to demonstrate prolonged duration of adrenal insufficiency in boys, this finding needs to be interpreted with great caution. Mahachoklertwattana et al. examined the influence of gender on adrenal function as well, but they found no difference in the duration of adrenal insufficiency between both sexes.\textsuperscript{1} Future research should focus on this potentially clinically relevant finding.

Clinical signs of adrenal insufficiency are non-specific and difficult to distinguish from symptoms of disease or chemotherapy related side-effects and include malaise, dizziness, nausea, fatigue, anorexia, hypoglycemia and hypotensive crisis (Chapter 2).\textsuperscript{18} Moreover, due to inadequate host defence against infections, adrenal insufficiency may contribute to the infection-related morbidity and mortality during treatment for childhood ALL and lymphoma.\textsuperscript{19-21} Rapid and effective management of complications of adrenal insufficiency can be achieved by glucocorticoid coverage (e.g. hydrocortisone) (Chapter 2).\textsuperscript{22-24} However, glucocorticoid coverage during periods of stress after cessation of glucocorticoid therapy in childhood ALL and lymphoma is not standard practice. In addition, there is no consensus on the optimal treatment of this complication. We therefore propose guidelines to treat adrenal insufficiency in childhood ALL or lymphoma.

**Recommendations**

The current supportive care guidelines recommend glucocorticoid coverage during periods of stress in the first 12 months after cessation of substantial glucocorticoid use (> 15 mg/m$^2$/day hydrocortisone equivalent, >14 days). A stress dose is defined as hydrocortisone 30-50 mg/m$^2$/day, according to the level of stress, divided into four daily doses.\textsuperscript{22-24}

Because our study demonstrated that certain patients with an adequate adrenal function developed adrenal insufficiency after a dexamethasone course of only five days, we recommend stress dose therapy after cessation of each high-dose glucocorticoid course for childhood ALL or lymphoma, irrespective of the duration of the course. We recommend to perform a low dose ACTH test five weeks after the last (tapering) dose of glucocorticoid therapy in all children treated for childhood ALL or lymphoma. In the period between cessation of glucocorticoid therapy and the adrenal function test, as well as in the period between consecutive glucocorticoid courses, every patient should receive glucocorticoid coverage during periods of stress to avoid life-threatening complications. During the first low dose ACTH test, the majority of patients will show sufficient cortisol responses (≥550 nmol/L) and
stress dose schedules can be discontinued. On the other hand, a minority of the patients will show insufficient cortisol responses and they need to be followed up until recovery. Low dose ACTH tests should be repeated, for example once every three months. In the mean time, these patients need stress dose therapy (Figure 12.1).

As illustrated in Chapter 2, some patients treated for childhood ALL or lymphoma may present with clinical symptoms of adrenal insufficiency in the absence of stress, i.e. during basal circumstances. In these patients, we advise to determine a basal morning cortisol level (between 0800 h and 1000 h). In case of an inadequate basal morning cortisol level <80 nmol/L, we recommend to add glucocorticoid substitution therapy to the stress dose scheme and to consult a pediatric endocrinologist.23,25 A substitution dose is defined as hydrocortisone 8-12 mg/m²/day divided into three daily doses (½:¼:¼).22,26,27 After two weeks, gradual cessation of the substitution therapy can be performed. In case of suboptimal basal morning cortisol levels of 80-270 nmol/L, we advise to repeat basal morning cortisol samples, e.g. once every three weeks, until basal morning cortisol levels >270 nmol/L have been reached (Figure 12.2). Continuation of the stress dose therapy depends on the results of the low dose ACTH test (Figure 12.1).

![Figure 12.1](image-url) Recommendation on stress dose therapy and adrenal function testing after cessation of glucocorticoid therapy for childhood ALL or lymphoma.
Future research objectives

Because this is the first study to describe an association between the ER22/23EK and the BclI polymorphisms of the GR and the duration of adrenal insufficiency after cessation of high-dose glucocorticoid therapy in childhood ALL, further studies are required to validate this association. Moreover, it needs to be elucidated whether early identification of patients homozygous for the BclI genotype and thus more susceptible for prolonged adrenal insufficiency after cessation of high-dose glucocorticoid therapy reduces the risk of life-threatening complications in children treated for ALL.

In addition, this is the first study that showed prolonged duration of adrenal insufficiency in boys compared to girls. Therefore, this potentially clinically relevant finding needs to be confirmed in larger studies. We recommend to include the evaluation of the adrenal and gonadal axis in further studies to examine their influence on the duration of adrenal insufficiency.

Moreover, a reliable predictor of the development of adrenal insufficiency in children treated for ALL or lymphoma would be helpful in order to identify patients with a high risk on prolonged adrenal insufficiency. A recent study demonstrated that cortisol levels after a low dose dexamethasone suppression test are reliable predictors for suppressed adrenal function on the seventh day after cessation of prednisone in healthy adult volunteers undergoing a 14-day course of prednisone treatment (0.5 mg/kg). Patients with a cortisol level

![Figure 12.2](image-url) Management of clinical symptoms of adrenal insufficiency after cessation of glucocorticoid therapy for childhood ALL or lymphoma.
after administration of dexamethasone in the lowest quartile had a 44% risk of developing an impaired adrenal function (demonstrated by a LD ACTH test), whilst participants in the highest quartile of cortisol levels had a low risk of adrenal failure (no risk in the cohort of that study). It would be of interest to study whether the low dose dexamethasone suppression test is predictive of the development of adrenal insufficiency resulting from glucocorticoid treatment in childhood ALL or lymphoma patients as well.

**Long term neuroendocrine and psychosocial consequences of treatment for childhood ALL**

Despite previous reports on prolonged adrenal insufficiency in children treated for ALL, cortisol levels in survivors of ALL have never been studied before. Both animal and human studies on long-term effects of exposure to dexamethasone during early life demonstrated hyporesponsivity of the HPA axis in later life. We hypothesized that the HPA axis activity in survivors of childhood ALL would be impaired. Interestingly, however, we found higher morning cortisol levels in ALL survivors compared to healthy controls. Moreover, suppression of cortisol levels in response to oral dexamethasone was higher in ALL survivors (Chapter 10). Previous research has shown that the exposure to a stressful life event in the past may lead to long-term dysregulation of the HPA axis. De Bellis et al. found that maltreatment experiences in childhood were associated with higher urinary free cortisol levels. In addition, Carrion et al. demonstrated that children with a history of trauma were characterized by increased adrenal activity. Furthermore, the results of a study by Saltzman et al. indicated that children exposed to marital violence have elevated salivary cortisol levels. The hyperresponsiveness of the HPA axis which was found in these studies may reflect an adaptive alteration that enables the traumatized child to better cope with the stressful experiences. It is conceivable that the experience of a life-threatening disease as childhood ALL and the intensive treatment with chemotherapy may account for a stressful life event as well. Interestingly, a recent study demonstrated enhanced plasma cortisol levels in survivors of childhood or adolescence cancer. In line with our study, the authors suggested that the exposure to a life-threatening experience in childhood or adolescence increases the endocrine response. This study indicates that enhanced HPA axis activity as a result of trauma exposure is not limited to survivors of childhood ALL, but might also be present in survivors of other types of childhood cancer.

Previous studies in trauma exposed children have not only shown enhanced cortisol levels but also increased sympathetic nervous system activity. Increased sympathetic outflow may have clinical implications, since it may contribute to the development of the metabolic syndrome and of cardiovascular disease in later life. Since no information on the sympathetic nervous system activity in survivors of childhood ALL was available, we
Chapter 12

examined the combination of HPA axis regulation and sympathetic nervous system activity in survivors of childhood ALL by evaluating the endocrine and cardiovascular response to social stress (Chapter 11). No significant differences were observed in the perceived stress, the saliva cortisol response or the cardiovascular response (systolic, diastolic, and mean arterial pressure, heart rate, stroke volume, cardiac output, and systemic vascular resistance) to the Trier Social Stress Test for Children between survivors of childhood ALL and healthy controls. We conclude that, despite the detected enhanced morning cortisol levels, the response of the HPA axis and the sympathetic nervous system to social stress is not altered in survivors of childhood ALL.

With increased survival rates of childhood ALL, issues concerning psychosocial well-being of survivors become increasingly important. This is reflected in the growing amount of instruments to measure various aspects of psychosocial health in childhood cancer survivors (Chapter 7 and 8). Previous studies including work of our group demonstrated an increased incidence of sleep problems and fatigue in childhood ALL and that these problems contributed to an impaired quality of life.45-47 As there is increasing evidence that sleep problems, fatigue and impaired QoL are associated with cortisol levels, we therefore studied sleep, fatigue and QoL in relation to cortisol levels in survivors of childhood ALL (Chapter 9 and 10).48-52 As illustrated in Chapter 9, according to parents of childhood ALL survivors, sleep, fatigue and physical QoL were substantially impaired after finishing treatment. Survivors of childhood ALL themselves, however, reported less sleep problems, less depressive symptoms and better psychosocial QoL than the Dutch norm. Discrepancies between parental and self-rating might be explained by worried parents and/or the adaptive style of the children. Furthermore, both parental and self-reports demonstrated that impaired sleep and fatigue were associated with more symptoms of depression and an impaired QoL. As illustrated in Chapter 10, higher cortisol levels in childhood ALL survivors were associated with more fatigue and poorer physical and psychosocial QoL. Accordingly, enhanced cortisol levels in childhood ALL survivors might be associated with clinically relevant problems.

Recommendations

At present, the clinical implications of the detected increased morning cortisol levels in childhood ALL survivors and its psychosocial consequences are unclear. For this reason, firm recommendations can not be made and further research is warranted.

As illustrated by Chapter 9, parents of childhood ALL survivors may underestimate the child's QoL, whilst childhood ALL survivors themselves may minimize their psychological and physical symptoms. We therefore recommend to include both child self-reports and parent proxy-reports when evaluating psychosocial well-being.
Future research objectives

As survival rates of childhood ALL increase over time, a better knowledge on the neuroendocrine and psychosocial consequences of childhood ALL is important. It is unclear whether the increased morning cortisol levels in childhood ALL survivors in comparison to healthy controls are of clinical relevance. However, since increased cortisol levels may lead to cardiovascular morbidity\textsuperscript{53-55}, it is essential to explore the clinical relevance of the increased cortisol production in survivors of childhood ALL in future studies, including longitudinal ones with long-term follow-up.

In addition, this was the first study to relate cortisol levels to psychosocial health in childhood ALL survivors. Therefore, the associations between increased cortisol levels and an impaired psychosocial health in childhood ALL survivors need to be confirmed in future studies.

Furthermore, since our study demonstrated that impaired sleep and fatigue in childhood ALL survivors were associated with more symptoms of depression and an impaired QoL, we hypothesize that improving sleep quality, and thereby managing fatigue, will improve QoL. However, before we can implement effective intervention strategies, it is necessary to obtain more information on the types and risk factors of sleep disturbances and the development of sleep during and after treatment for childhood ALL by conducting more detailed studies. In addition, because of the discrepancies between parental and self-ratings regarding sleep, it will be helpful to objectively measure sleep, for example using actigraphy.

In this thesis, we have gained more insight in the short term neuroendocrine consequences of treatment with high-dose glucocorticoid therapy for childhood ALL and lymphoma. Therefore, we were able to suggest guidelines regarding glucocorticoid coverage in order to avoid life-threatening complications in children with ALL or lymphoma who suffer from adrenal insufficiency. In addition, we identified long term neuroendocrine and related psychosocial consequences of treatment for childhood ALL and lymphoma. However, further research on the clinical implications of these findings is needed.
References


