Supplementary files

Supplementary file 1. Example search strategy for the MEDLINE database

Example search strategy for the MEDLINE database.

S1	(MH "Child, Preschool")
S2	TI child* OR AB child*
S3	TI (boy* OR girl*) or AB (boy* OR girl*)
S4	TI toddler OR AB toddler
S5	TI young N1 child* OR AB young N1 child*
S6	TI early N1 child* OR AB early N1 child*
S7	TI early N1 year* OR AB early N1 year*
S8	TI “pre-primary” or AB “pre-primary”
S9	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8
S10	(MH "Schools, Nursery")
S11	TI nurser* OR AB nurser*
S12	(MH "Learning") OR TI early N1 learning OR AB early N1 learning
S13	TI (“preschool” or “pre-school”) OR AB (“preschool” or “pre-school”)
S14	TI kindergarten OR AB kindergarten
S15	TI (childcare OR child N1 care) OR AB (childcare OR child N1 care)
S16	TI (daycare OR day N1 care) OR AB (daycare OR day N1 care)
S17	(MH "Education") OR TI (education OR "preschool education" OR "outdoor education" OR "adventure education") OR AB (education OR "preschool education" OR "outdoor education" OR "adventure education")
S18	MM "Play and Playthings" OR TI (Play OR “play-based learning”) OR AB (Play OR “play-based learning”)
S19	TX (Waldkindergartens OR udeskole OR friluftsliv OR peuterspeelzaal OR kinderopvang OR bush N1 kinder*) OR TI (forest N1 kindergarten* OR forest N1 school*) OR AB (forest N1 kindergarten* OR forest N1 school*)
S20	S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19
S21	TI outdoor* OR AB outdoor*
S22	TI (nature OR “nature-based”) OR AB (“nature-based”)
S23	TI environment* OR TI outdoor N1 environment* OR AB outdoor N1 environment*

S24	TI (forest* OR wood* OR park* OR recreation* OR landscape* OR tree* OR hill* OR garden* OR beach* OR eco)
S25	AB (forest* OR wood* OR park* OR recreation* OR landscape* OR tree* OR hill* OR garden* OR beach* OR eco)
S26	TI (green OR greenspace or green N1 space) OR AB (green OR greenspace or green N1 space)
S27	TI (loose N1 parts OR “loose-parts”) OR AB (loose N1 parts OR “loose-parts”)
S28	TI (adventure* OR wild OR “open-air”) OR AB (adventure* OR wild OR “open-air”)
S29	S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28
S30	S9 AND S21 AND S30

Supplementary file 2. Modified quality appraisal tools for quantitative and qualitative studies

Modified Effective Public Health Practice Project (EPHPP) Quality Assessment Tool

Modifications highlighted in red text

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population? (i.e. children aged 2-7 years not in formal education yet)

1. Very likely
2. Somewhat likely
3. Not likely
4. Can't tell

(Q2) What percentage of selected individuals consented to the research?

1. 80 - 100% agreement
2. 60 – 79% agreement
3. less than 60% agreement
4. Not applicable
5. Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design:

1. Randomized controlled trial
2. Controlled clinical trial
3. Cohort analytic (two group pre + post)
4. Case-control
5. Cohort (one group pre + post (before and after))
6. Interrupted time series
7. Other specify _____
8. Can't tell

Was the study described as randomized? If NO, go to Component C.

No **Yes**

If Yes, was the method of randomization described? (See dictionary)

No **Yes**

If Yes, was the method appropriate? (See dictionary)

No **Yes**

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?

1. Yes
2. No
3. Can't tell

The following are examples of confounders:

1. Gender
2. Age
3. Socio economic status (SES – e.g. Parental education, deprivation status)

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?

1. All confounders
2. Two confounders
3. One confounder
4. Can't Tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) and/or analysists aware of the intervention or exposure status of participants?

1. Yes
2. No
3. Can't tell

(Q2) Were outcome assessors aware of the research question?

1. Yes
2. No
3. Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?

1. Yes
2. No
3. Can't tell

(Q2) Were data collection tools shown to be reliable?

1. Yes
2. No
3. Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK

See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?

1. Yes
2. No
3. Can't tell
4. Not Applicable (i.e. one time surveys or interviews)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).

1. 80 -100%
2. 60 - 79%
3. less than 60%
4. Can't tell
5. Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

COMPONENT RATINGS

Please transcribe the information from the grey boxes on pages 1-3 onto this page. See dictionary on how to rate this section.

A	**SELECTION BIAS**	**STRONG**	**MODERATE**	**WEAK**
		1	2	3
B	**STUDY DESIGN**	**STRONG**	**MODERATE**	**WEAK**
		1	2	3
C	**CONFOUNDERS**	**STRONG**	**MODERATE**	**WEAK**
		1	2	3
D	**BLINDING**	**STRONG**	**MODERATE**	**WEAK**
		1	2	3
E	**DATA COLLECTION METHOD**	**STRONG**	**MODERATE**	**WEAK**
		1	2	3
F	**WITHDRAWALS AND DROPOUTS**	**STRONG**	**MODERATE**	**WEAK**
		1	2	3

Overall Grade (based on above six criteria):

• Scored **1** for study design (i.e. controlled studies); AND
• Scored **1** or **2** in at least three other important components, including:
  ○ selection bias
  ○ confounders
  ○ blinding
	STRONG
	1

○ withdrawals and drop-outs.	
• Scored **1** for study design; AND • Scored **1** or **2** in two other important components, including: ○ selection bias ○ confounders ○ blinding ○ withdrawals and drop-outs. OR • Scored **2** for study design; AND • Scored **1** or **2** in at least three other important components, including: ○ selection bias ○ confounders ○ blinding ○ withdrawals and drop-outs.	**MODERATE** **2**
• Scored **1** for study design; AND • Scored **3** in more than two other important components, including: ○ selection bias ○ confounders ○ blinding ○ withdrawals and drop-outs. OR • Scored **2** for study design; AND • Scored **3** in more than one other important components, including: ○ selection bias ○ confounders ○ blinding ○ withdrawals and drop-outs. OR • Scored **3** for study design	**WEAK** **3**

Dixon-Woods (2004) checklist

Modifications highlighted in red text

Question 1	Are the research questions clear?
Question 2	Are the research questions suited to qualitative inquiry
Question 3	Are the following clearly described? - Sampling - Data collection - Analysis
Question 4	Are the following appropriate to the research question? - Sampling - Data collection - Analysis

Question 5	Are the claims made supported by sufficient evidence?
Question 6	Are the data, interpretations, and conclusions clearly integrated?
Question 7	Does the paper make a useful contribution to the review question?

Each question is answered either “yes”, “no”, or “can’t tell”.

Supplementary file 3. Characteristics of included studies

Table 1. Characteristics of included quantitative studies						
Author, year and country	**Study design**	**Age (range or mean ± SD), sex (n or % m/f), SES.**	**Exposure and comparison**	**Follow-up time point**	**Outcome(s)**	**Data analysis**
Nature-based ECE						
Agostini et al (2018), Italy. E: 41 children / 7 teachers / 1 school C: 52 children / 13 teachers / 1 school	Controlled before & after	E: Age: 47.2 months ± 6.52 Gender: 13m/28f C: Age: 46.75 months ± 6.95 Gender: 29m/23f SES not reported.	E: Teachers underwent special training in outdoor education over one year including (15 days). ECE consisted of a green park with some centuries-old trees (e.g., firs, willows, maples), plants and flowers, and without any play structures. C: ECE contained grass and cement without larger plants, trees, and play structures.	T1= Jan 2014 T2= May 2014 T3= Oct 2014 T4= May 2015 (16 months from baselines)	Motor competence	Mixed-Model Repeated Measures analysis of variance (ANOVA)
Choi et al (2014), South Korea. E: 18 children / 1 ECE C: 19 children / ECE	Controlled before & after	E: Age: 4.2 ± 1.1 Gender: 11m/7f SES: all middle class C: Age: 4.0 ± 1.1 Sex: 11m/8f SES: all middle class	E: Children attend forest kindergarten 5 days per week, year-round, regardless of weather conditions. Children are outdoors more than 80% of the day and usually play, walk, run, and observe various things in the forest. C: Regular kindergarten (not described)	8 months	Sleep	Wilcoxon signed rank test.
Ene-Voiculescu & Ene-Voiculescu (2015), Fjortoft (2004), Fjortoft (2001), Norway.	Controlled before & after	Age: 6.1 years Gender: 38m/37f SES not reported.	E: Children used the forest every day for 1-2 hours throughout the year when they attended kindergarten. Occasionally they used the outdoor playground inside the kindergarten fence. The small forest (7.7 hectares) consisted of mixed woodland vegetation,	10 months	Motor competence	T-test.

E: = 46 children / 1 kindergarten C: 29 children, / 2 kindergartens			some open spaces of rocks and open fields and meadows in between. C: Children used the traditional outdoor playground for 1-2 hours a day and visited natural sites only occasionally.			
Müller et al (2017), Canada. E: 43 children / 1 ECE C: 45 children / 1 ECE	Controlled before & after	Age: E: 63.56 months (3.33 SD) C: 64 months (3.56 SD) Gender not reported. SES not reported.	E: "nature kindergarten" C: "traditional kindergarten" Neither are described.	9 months Sep/Oct-May	PA Motor competence	Analyses of Covariance (ANCOVA)
Frenkel et al (2019), USA. E: 71 children / 5 ECE C: 70 children / 4 ECE	Controlled cross-sectional	Age: 4.3% = 2 years, 29.1% = 3 years, 50.4% = 4 years, 16.3% = 5 years Gender: 82m/59f SES: 103, 036 USD (median zip code Income)	E: All nature ECE sites were located in parks with distinct areas marked off with rocks and other natural features for daily activities. Children were encouraged to play in the natural environment, which included grassy areas, areas with dirt, and tree cover and to play with natural features such as sticks, rocks, and mud. C: Traditional ECE were primarily held indoors and had outdoor play areas built on concrete. children spending less than 1.5 hr outdoors each day.	N/A	Unintended consequences	Poisson regression models Covariates: age
Fyfe-Johnson et al (2019), USA. E: 20 children / 1 ECE C: 13 children (waitlist control or 2-hour nature-based, outdoor	Controlled cross-sectional	Age: 3-5 years Gender: E: 11m/9f C: 9m/4f SES: E: 18 > $90,000 C: 8> $90,000	E: The nature ECE occurs outdoors in a forested park where most children attend 5 days per week from 9 am to 1 pm; 2-day and 3-day per week options are available on a limited basis. The physical environment consists of dedicated classroom areas in the forested areas. Children use logs and tree stumps to sit; portable canopies are used during inclement weather. Most of the day is	N/A	Physical activity	Descriptives only.

enrichment class provided by experimental ECE			spent hiking and exploring the surrounding forest. No traditional play structures or pre-fabricated playgrounds are utilized. C: 2 hour nature-based outdoor enrichment class was offered once weekly by the same nature ECE the intervention group children attended. Classes were led by a teacher and attended by both child and caregiver. The classes consisted of science-based exploration through outdoor play in a forested park and involved: circle time, station time (learning stations that emphasize sensory and fine motor skills, creativity, and numerical and literacy skills), short stories, and hikes. Others were included in a wait-list control			
Lysklett et al (2019), Norway. E: 43 children / 4 ECE C: 49 children / 4 ECE	Controlled cross sectional	Age: 5.1-6 years Gender: 53m/39f SES not reported	Nature-based ECEs located close to a large recreational area, with woods, lakes and tracks just outside the city centre. Both types of preschools used the nearby nature area for hiking and playing every week E: nature ECE at least three times, per week C: traditional preschools once per week.	N/A	Motor competence	T-test
Meyer et al (2017), Canada. E: 46 children / 3 ECE C: n= 35 children / 2 ECE	Controlled cross-sectional	Age: 5-6 years Gender: 39m/42f SES: predominately middle-class children	E: Children spent every morning in nature participating in teacher-directed, nature-based learning activities. The nature kindergartens differed per site but included a beach, unmanaged wooded area, natural playground (trees and vegetation) and artificial playground. C: Children were assessed in their classrooms where they engaged some storytelling, singing, dancing, tai chi, reading, drawing, and art. They also took part in music and computer classes and science fair.	N/A	PA	Descriptives only.

Moen et al (2007), Norway. E: 267 children / 37 ECE C: 264 children / 32 ECE	Controlled cross-sectional	Age: 3-6 years. Gender not reported. SES not reported.	E: had ‘‘outdoor’’ or ‘‘nature’’ as part of their name, or emphasized outdoor pedagogy and children spent an average of 3.5–8 hours/day outdoors in winter. C: children spend on average spend 1.25–4.0 hours/day outdoors.	N/A	Unintended consequences	GLM
Scholz & Krombholz (2007), Germany E: 45 children / 10 forest kindergartens C: Rural = 42 children / 2 ECE; Urban = 42 children / 2 ECE	Controlled cross-sectional	Age: E: 5.5 (SD 0.4) C: Rural= 5.7 (0.4 SD); Urban= 5.7 (0.4 SD) Gender: 71 boys, 58 girls SES not reported.	E: forest kindergarten C: traditional rural and urban kindergarten	N/A	Motor competence	MANOVA Covariates: age
Weisshaar et al (2006), Germany. E: 506 children / 25 ECE C: 1201 children / 28 ECE	Controlled cross-sectional	Age: 4.9 (1.1 SD) Gender: 901m/803f SES not reported.	E: Forest kindergarten located in forested areas where children spend all-season full-time outdoors. C: Conventional kindergartens (not described)	N/A	Unintended consequences	Fisher test and logistic regression Covariates: age, sex, skin inspection, and recommended vaccination
Ernst (2014), USA. E: 46 educators	Cross-sectional	Not described.	Outdoor environments that range from relatively natural to wild spaces.	N/A	Motor competence	Multiple regression
Wright (2019), USA. 48 children / 2 ECE	Cross-sectional	Age: 3-5 years Gender not reported. SES not reported.	The 2 sites were located in a forested park/ They both consisted of large space (10,000Sq/ft), log borders, sloping areas, vegetation, large trees, natural loose parts.	N/A	Physical activity	Descriptives only

			Manufactured supplies such as shovels, wheelbarrows, books, magnifying glasses were brought in. 4 hours of the school day is spent outdoors.			
ECE natural playgrounds						
Brussoni et al (2017), Canada. E: 48 children / 2 ECE	Uncontrolled before & after (mixed methods)	Age: 4.28 (0.63 SD) Gender: 53% m/47%f SES not reported.	Playgrounds were improved using the Seven Cs which consists of 27 items, rated on a 5-point scale, for a maximum score of 135 Changes predominately involved inclusion of more natural elements such as, vegetation, boulders, rock, loose parts. Seven Cs scores increased from 44 to 97 in ECE A, and 35 to 125 in ECE B.	Data were collected at T2; May-July 2014) two-weeks after playground modification	Physical activity	Wilcoxon signed rank tests; General linear modelling. Covariates: age, gender, ECE
Cosco et al (2014), USA. E: 804 / 27 ECE	Uncontrolled before & after	Age: 2-5 years Gender not reported. SES not reported.	Preventing Obesity by Design is an ECE outdoor renovation intervention. Prior to the intervention the space had few structures (slides, swings etc.) in a rectangle space enclosed by a fence. Whereas, post intervention, the space had more natural elements, including trees, garden, vegetation etc.	Not described.	Physical activity	Logistic regression and bivariate correlations Covariates: gender
Luchs, & Fikus (2018), Germany. E: 17 children / 1 ECE	Cross-sectional	Age: 5.85 ± 0.49 years Gender: 9m/8f SES not reported.	E: the nature playground has large natural space featuring trees, grass, hills, vegetations, water C: the contemporary playground has traditional play structures such as slides and swings. It has some natural elements, including grass and trees.	N/A	Physical activity	Paired sample t-test
Storli et al (2010), Norway. E: 16 children / 1 ECE	Cross-sectional	Age: 3-5 years Gender: 9m/7f SES not reported	Nature - gathering loose nature materials, climbing running. Traditional - children engaged in activities such as cycling, digging, climbing	N/A	Physical activity	t-tests
Torkar & Rejc (2017), Slovenia.	Cross-sectional	Age: 4 and 5 years old	E: forest playground which contains a forest patch, river and bushes. The space is approx. 500 m^2	N/A	Physical activity	Mann Whitney

E: 25 children / 1 ECE		Gender: 16m/9f SES not reported.	C: Traditional playground which contains fixed equipment such as seesaw, roundabout, slide, climbers and playhouse. There is some nature surrounding the playground (trees, bushes). The space is approx. 500 m^2			
Natural elements within ECE						
Ng et al (2020), Australia. E: 159 children / 6 ECE C: 138 children / 5 ECE	Controlled before and after	Age: 2 years 10 months (0.82 SD) Gender: 49%m/51%f SES: No significant differences between intervention and control group reported.	Variable of interest was natural elements. Measured using the modified Environment and Policy Assessment and Observation (EPAO) physical environment domain. This tool assesses the prevalence of PA opportunities in the physical environment. There were 5 subscales: Fixed play equipment' and 'Portable play equipment' from the EPAO, 'Total size of playing area', 'Outdoor play spaces', and 'Natural elements'. A number of items per subscale were scored - 1 if present, 0 if not.	6 months	Physical activity	Mulitvariate linear regression Covariates: age, sex, parental education, accelerometer wear time.
Boldemann et al (2004), Sweden. E: 64 children / 2 ECE	Cross-sectional	Age: 1-6 years Gender:26m/38f SES not reported.	E: ECE 1 had play constructions surrounded by trees but exposed to the sun and ECE 2 had attractive play constructions positioned under a canopy of tree crowns. Average time spent outdoors was 207 min at site ECE 1, and 256 min at site 2.	N/A	UV exposure	t-tests
Boldemann et al (2006), Sweden. E: 199 children / 11 ECE	Cross-sectional	Age: 4.5-6.5 years Gender:114m/85f SES not reported.	ECE environment scores and averages dichotomized to (>2 high, <2 low) Outdoor environments were assessed on their play potential. They were scored 1, 2, and 3 with respect to size of outdoor area, overgrown surfaces (trees shrubbery) and integration of play structures or other defined play areas with vegetation.	N/A	Physical activity UV exposure	Bivariate analysis; Linear mixed-models.
Christian et al (2019), Australia.	Cross-sectional	Age: 3.4 ± 0.8 Sex: 53%m/47%f	ECE settings were dichotomized to vegetation < 3m in height or vegetation> 3m in height.	N/A	Physical activity UV exposure	Multilevel linear

E: 678 children / 48 ECE		SES: 32% = low, 34% = medium SES and 34% = high SES.	High-resolution airborne multispectral 4-band images and Geographic Information System (GIS) was used to identify the location, shape and size of ECE outdoor play spaces. Approximately 31% of centres' outdoor play space had vegetation with 23% (20.5 SD) having <3 m in height and 8% (13.7SD) with >3 m hight.			regression models. Covariates: age, gender, and ECE SES and size.
deWeger (2017), Australia. E: 274 children / 12 ECE	Cross-sectional	Age: 4.2 years (0.5 SD) Gender: 141m/133f SES not reported.	Variable = natural elements The quality of the outdoor learning environment in the ECE's was assessed for 3 hours per day over 2 days using the POEMS instrument. This is grouped into 5 domains: **Physical environment (13 questions),** Interactions (13 questions), Play and Learning Settings (13 questions), Program (9 questions), and Teacher/Caregiver role (8 questions). Scores are them summed to give a total score	N/A	Physical activity	Hierarchical linear modelling (HLM) Covariates: age, gender, BMI-z score and accelerometer wear time (level 1), outdoor environment quality (level 2)
Gubbels et al (2018), Netherlands. E: 151 children / 22 ECE	Cross-sectional	Age: 34.14 months (8.97 SD) Gender: 72m/79f SES not reported.	The SB and PA physical environment of each ECE was assessed using a standardized observation protocol, based on the updated Environment and Policy and Assessment Observation (EPAO). The following natural elements were assessed: large trees (2.5 m or taller), small trees (less than 2.5 m tall), trees that children can climb, shrubs, flowering plants, variation in ground (hills, mounds), grass, rocks large enough to climb, a hill for rolling down or climbing up.	N/A	Physical activity	Multivariate linear regression analyses

			A sum score of all the types of natural elements that were present was calculated.			
Määttä et al (2019), Finland. E: 864 children / 66 ECE	Cross-sectional	Age: 4 years 4 months (10 SD) Gender: 48% girls SES: 29% had mother with high educational background (at least masters)	Observation instrument was designed for the study and consisted of items from the EPAO. ECE physical environments were assessed, of which, surfaces in the preschool grounds (9 items) and terrain in the playground, related to the natural environment (grass, forest, trees, rocks).	N/A	Physical activity	Multilevel linear regressions models Covariates: age, gender, season, municipality, pre-school group cluster
Määttä et al (2019b), Finland. E: 655 children / 66 ECE	Cross-sectional	Age: 4.7years (0.89 SD) Gender: As above SES: As above	Frequency of nature trips (mean/per week): Teachers completed weekly diary of activities which were categorised into 5 groups (1=outdoors, 2=teacher-led sessions, 3=free play, 4=organised PA lessons and 5=mixed sessions). Daily number of each activity was calculated and summed for the week level and then divided by the number of the days (from 3 to 5) to form the average daily amount of each activity. A questionnaire was then completed to determine activities that are close to the ECE and occur regularly (nature visits). Visits were recorded for mean times per week	N/A	Physical activity	Multilevel linear regressions models. Covariates: age, gender, average attendance at preschool and study season
Olesen et al (2013), Denmark. E: 441 children / 42 ECE	Cross-sectional	Age: 5.8 years Gender: 49.5%m/50.5%f SES not reported.	Researchers collected a range of environmental correlates, of which, vegetation and hilly landscape related to nature	N/A	Physical activity	Univariate analyses and multi-level modelling Covariates: Gender, rain, preschool

						type, afternoon hours, location, indoor area, Playground area, playground time, parent education
Sando (2019), Norway. E: 80 children / 8 ECE	Cross-sectional	Age:3.5 (SD=0.5) Gender: 41m/39f SES not reported.	The places and materials in the playground were categorised into nature, pathways, open area and fixed functional equipment. Nature was coded in four of the institutions and ranged from large forest areas (1500 m^2) to smaller areas with trees and natural surfaces.	N/A	Physical activity	A random intercept multilevel model Covariates: age, gender
Sando & Sandseter (2019), Norway. E: 73 / 8 ECE	Cross sectional (mixed-methods)	Age: 4.2 years (0.7 SD) Gender: 36m/37f SES not reported.	ECE settings featuring nature were coded (places). For objects, these were coded when a child was holding, using or interacting with an object and included: sand, water, mud and nature materials The variables for places and objects describe the percentage of time the child is at a place or in which the object was used during each observation.	N/A	Physical activity	Generalized linear latent and mixed models
Söderström at al (2013), Sweden. E: 172 children / 9 ECE	Cross-sectional	Presented per ECE Age: S1: 4.6 (1.0 SD) S2. 4.1 (0.5 SD) S3: 4.3 (0.7 SD) S4: 4.4 (0.8 SD) S5: 4.7 (0.8 SD) S6: 4.6 (0.9 SD) S7: 4.3 (0.9 SD) S8: 4.6 (0.6 SD)	Outdoor Play Environment Categories (OPEC) scoring tool was used to assess playgrounds on (i) total outdoor area, (ii) amount of trees, shrubbery and hilly terrain and (iii) integration between vegetation, open areas and play structures, each component with a score range of 1–3 (high score = high quality).	N/A	Sleep Unintended consequences Weight status	ANOVA and MANOVA Covariates: Age, gender, birth Weight, mother SES.

		S9: 4.8 (0.7 SD) Gender: % f S1: 29% S2. 41% S3: 50 % S4: 42% S5: 50% S6: 56% S7: 61% S8: 41% S9: 63%	The OPEC scores were then dichotomized (low OPEC value< 2, high OPEC value >2)			
Sugiyama et al (2012), Australia. E: 89 children / 10 ECE	Cross-sectional	Age: 4.1 (0.6 SD) Gender: 54%m/46%f SES not reported	Questionnaire assessing characteristics of the ECE's was completed by the centre Director. Outdoor characteristics of relevance were gradient shade, vegetation, surface material (grass).	N/A	Physical activity	Multilevel linear regression Covariate: age, gender and time spent outdoors

Table 2. Characteristics of included qualitative studies

Author, year and country	Age (range or mean ± SD), sex (n or % m/f), SES.	Exposure and comparison	Research aims	Data collection method	Details of analysis
Nature-based ECE					
Bjørgen (2016), Norway.	Age: 3-5 years Gender: 10m/14f	Children played in the ECE outdoor play space for 3 hr/day, and each week would go on trips (1 or 2x) to natural environments.	What is the relation between environmental affordances and	Observations were made with video recording the different seasons of the year for 20 days, 10 days on trips in a	Thematic analysis - the first phases of coding were assessing and identifying the children's level of PA in

24 children / 1 ECE	SES not reported.	The large outdoor area consists of outdoor toys (buckets, shovels, trucks, balls), swings, sandboxes, climbing racks, natural materials, small trees, a varied surface of grass, sand, asphalt, and small hills. The destination for excursions in diverse natural landscape environment is approximately 300–700m from the centre. One type of natural environment was open fields suitable for tobogganing, running and playing on skis. Another natural environment consisted of woods. Trips were made to the natural environments all year round.	PA levels among 3–5 year olds?	natural environment and 10 days in the centres play space. A total of 50 h of direct observation was conducted. Coding of the physical activity levels of children was assessed and adapted using the Observational System for Recording Physical Activity in Children-Preschool Version (OSRAC-P) manual.	different play situations. Figures were used as an analytical tool helped to discern patterns, differences and similarities in the data material, which laid foundations for the qualitative analysis of the affordances. Thereafter themes of affordances are identified within the data. The theory of affordances and criteria from the 7Sc were used in the analysis process.
Dowdell et al (2011), Australia. E: 6 children / 1 ECE C: E: 6 children / 1 ECE	Age: 2-6 years Gender: 6m/6f SES not reported.	E: Has an emphasis on nature and sustainable education. The space is large and consists of sandpit, fairy garden, play equipment, grass area and vegetable garden. C: Located in a warehouse this centre has an entirely artificial indoor play area. It consists of a bike track, home corner (playhouse etc), climbing structures, quiet play area, sandpit and obstacle course.	How are children's play behaviours and social interactions influenced by the opportunities and materials present in their outdoor play environment?	Play behaviours were recorded using a behaviour mapping schedule. Each child was observed individually and every 10 seconds an observation based on social interaction and play behaviour was recorded.	Once all the observations were made for each child at each centre they were then tallied up. Play behaviours were then categorised into four different groups: social activities, cognitive activities, physical and motor skill activities and other activities.
Liu (2020), USA Nature interaction: E: 29 children / 1 ECE	Age: 4-5 years Gender: 30m/ 25f SES: E: 48,000	E: contains high levels of nature with a variety of perceived affordances. Outdoor time = 1.5 hours/day. 32 types (categories- vegetation (tress, shrubs, flowers, grasses), natural ground surface (wood chips, meadow,	How does the designed nature-based outdoor play environment in ECE impact	RQ 1. Field observation, behaviour mapping, semi-structured interview with teachers. RQ2. Field observation,	Content analysis was used for: children's frequent play locations, types of play behaviors, frequency and diversity of different ways of interaction with natural

C: 26/ 1 ECE Restorative experiences: E: 10 children / 1 ECE C: 9 children/ 1 ECE	US (household income); C: 59,000 (household income) of children attending each centre	multipurpose lawns), natural materials, natural play structures (e.g. wood, stick, water, sand logs, ice, leaves), animals, experiential elements (rain, snow, sky view, light, air) of natural elements and play settings and 4 types of non-nature-based play settings (concrete track, bicycles, concrete hall, concrete sq.) were identified C: low levels of nature and perceived affordances. Outdoor time = 1.5 hours/day. 13 types of natural elements and 11 (vegetation, natural ground, animals) types of non-nature-based play settings (examples include: play structure, playhouse, outdoor kitchen, bicycles) were identified.	children's interaction with natural elements? How does the designed nature-based outdoor play environment in ECE impact children's restorative experience?	structured Interview with children, semi-structured interview with teachers.	elements, as well as restorative experience from semi-structured interviews with teacher and structured interview with children. Themes (coding categories) were drawn from the theoretical framework. Specifically, children's types of play behaviors and their ways of interacting with natural elements were coded using function taxonomy of affordance (Heft, 1988; Kyttä, 2002) and Gibson's affordance theory.
Sandseter (2009), Norway. 29 children from both experimental and control groups E: 1 ECE C: 1 ECE	Age: 4-5 years Gender: 21f/8m SES not reported.	E: Located in a forest with no fixed play equipment and fencing and children spent most of their time outdoors. C: fixed equipment, such as swings, climbing tower, play hut and a few trees.	To explore affordances for risky play in two different play environments: an ordinary ECE playground and a nature playground.	7 days were spent on each of the ECE playgrounds. Video recordings and field notes of risky play situations were collected based on categories of risky play; a) great heights, b) high speed c) dangerous tools, d) dangerous elements, e) rough-and-tumble play, f) where the children can disappear/get lost. Both the children's play and the staff's supervision were observed. The field notes and the video recordings were transcribed into an electronic word file.	A content analysis was performed on the data. The analysis was theory-driven. Firstly, each of the play environments' potential affordances for risky play, as categorized by Sandseter (2007), were analysed in relation to the most relevant affordance categories to evaluate their potential affordances for risky play. Secondly, the transcriptions of the video observations, field notes, and interviews were examined to determine the types of risky play

				12 children in the ordinary preschool and 11 children in the nature and outdoor preschool participated in a one-to-one qualitative interview with the researcher. Each interview was approximately 20-30 minutes and was recorded on audiotape. The interviews were semi-structured, using an interview guide list of questions and issues. The interview guide was based on the six categories of risky play and aimed to explore the types of risky play that the children engaged in within the different play environments and whether the staff constrained or intervened in their actions. Upon completion of the interviews, the audiotapes were professionally transcribed verbatim into an electronic word file.	children engaged in within different environments. Thirdly, the observations and the interviews were analysed to determine the degree to which children experienced mobility license while engaging in risky play. The transcriptions of the video observations were examined to determine the extent to which, and in which situations, the staff had children under surveillance while they engaged in risky play or was taking initiative to or constrained risky play.
Streelasky (2019), Canada. 15 children / 1 ECE	Age: 5-6 years Gender not reported. SES not reported.	The ECE setting had an outdoor, nature-based focus where children spent afternoons in the forested area. The teacher who was involved in an Outdoor Environmental Leadership Programme engaged the students in an integrated learning approach where key curriculum areas were addressed (e.g. language arts, social studies, science	What learning experiences do kindergarten children value at school? and what modes are they choosing to express and represent their	Qualitative interpretative approach involving (i) group discussions, (ii) participant observations, (iii) anecdotal notes, (iv) artefact collection and (v) individual semi-structured interviews (children's narratives).	Data were analysed and grouped into themes. Image based analysis was used to develop deeper understanding of children's interests and knowledge. Thematic analysis was used

		and physical education). Children also had time to freely explore the forest.	valued school learning experiences?		to gain insight into children's practices which followed 6 phases: (i) familiarising oneself with the data and identifying items of potential interest, (ii) generating initial codes, (iii) searching for themes, (iv) reviewing potential themes, (v) defining and naming themes and (vi) reporting the themes.
ECE natural playgrounds					
Herrington & Studtmann (1998), USA. 36 children / 1 ECE (2 "labs")	Age: 2-6 years Gender: 16m/20f SES not reported.	Pre-modification: Lab A: consisted of a patio area, grass lawn, play structures, swing set, doll house, trees and vegetation. Lab C consisted of a porch area, grass lawn, play areas, swing set, trees and vegetation. Post-modification: Playground were naturalised with increased natural elements: ice sculptures, wind chimes, canopy, chalk, buckets, playhouse, water pay, vegetation and trees were added to the labs. Lab A received more natural elements than lab C but both were more natural post intervention.	What natural materials and conditions of the outdoor environment can contribute to the development of young children ranging from 2 to 6 years old?	Phase 1: sequence sampling of children during free-play. Children were video-taped interacting with the site for 1 month. Once the modifications were made, data collection began a week later. Data collection involved video-taping, sound recording, and field notes. Videotaping involved following a child for 20 minutes as they moved throughout the yard in free play. Voice recordings of the children were made of one of the two selected children from each Lab. Voice recordings were transcribed	20 hours of videotapes were analysed. During analysis, notes were made. For Phase 1 the notes were: (1) interaction with an intervention (2) duration of interaction (3) children's behavioural modification made between pre and post intervention (4) children's movement changes made between pre and post intervention. For Phase 2 the criteria were: (1) which children were engaged in the intervention; (2) how many children were engaged (3) the duration and nature of their engagement with the intervention (4) how behavior and paths of

				into text documents. Field notes (weather, teacher and children present, anecdotal observations etc.) were made daily by researchers. Notes were recorded by researchers on a pre-printed notation sheet that displayed a plan view of both yards. Phase 2: Video documentation and anecdotal notes were employed to record event sampling. Event sampling allowed subjects to be taped if they interacted with the plant interventions. The specific intervention sites were recorded on a rotating basis. Children were video-taped using the same schedule as in Phase 1 and fieldnotes were made in the same manner as in Phase I	movement changed between pre and post intervention. Video clips were selected that illustrated the notes. These clips were put together on one VCR tape using a television and VCR recorder. The conversations of the children participating in Phase 1 were transcribed at 10 second intervals. The anecdotal notes were reviewed and complied.
Puhakka et al (2019), Finland. 12-24 children (not clear) / 6 ECE	Age: 3-5 years Gender not reported. SES not reported.	Playground yards were transformed through enhancing the biodiversity by incorporating more greenspace and vegetation. For example, replacing areas covered in gravel with forest floor. Children spent time outdoors every day (0.5–2 h in the morning and in the afternoon) as well as participating in teacher led activities 4-5 days/ week.	Does biodiversity exposure and greening playgrounds affect 3–5 years-old children's physical activity and play, their environmental relationships, and their well-	Educators and child nurses completed interviews and surveys respectively. 49 parents completed surveys. Surveys were completed one month after the playground was modified. Surveys included both structured and open ended questions which related to children's play activities, and enthusiasm.	Interviews were recorded and transcribed verbatim. Survey and interview data were analysed using qualitative content analysis to identify different affordances. The affordances were then classified into 6 themes which emerged from analysis and coding. How these affordances

			being in the urban environment in Finland.	Interviews with parents focussed on children perception of modifications. The educator thematic interviews focused on possible changes in children's play and other activities in the yard, in children's and educators' interest in and knowledge of nature, their well-being, attitudes towards outdoor activities, and practices and atmosphere in the ECE setting	supported children's relationship with the modified playground were then mapped. Finally, these two elements were brought together to form three perspectives.
Wishart et al (2019), Australia. 75 children / 1 ECE	Age:4-5 years Gender not reported. SES not reported.	The two playgrounds were located on different sides of the building, each extending to the back of the building where a connecting gate was sometimes opened to allow free-flow of children between the two spaces. E: Traditional equipment was replaced with terraces, inclines, logs and rocks designed to afford physical activities and gross motor skills such as climbing and balancing. other elements included: Natural gardens with fruit trees; herb garden and small plants; logs; stepping-stones; log enclosure; small tree forest; sandpit with pebbles and medium-size rocks. C: standard equipment: slide, ladders, swings, climbing frames, sand-pit, surfaces open area. This area also	Does the naturalised design of the new space provide equivalent actualisable affordances for different types of physical activity to those provided by the more traditional playspace, with its conventional equipment and resources	Behaviour mapping using a time-sampling observation tool. Observations were conducted between 10:30–15:30 during sessions. The two playscapes were divided into zones and children were observed in 3 minute cycles. For each observation, the tool also noted: number of boys and girls (no further count of children was taken); presence of educators; whether play was solitary or group; location and general contextual information. 40 observations in the naturalised space and 42 observations in the traditional space were made.	Behaviour mapping tracked the incidence of different categories of movement across different areas of the two playscapes, to investigate if different categories of movement were more likely to occur in specific areas or in relation to specific features.

		included a grass area, veg garden, trees and shrubs.			
Abbreviations: E= experimental; C= control; n= number; m=male; f= female; ECE = early childhood education (includes preschools, day care, kindergarten etc.); SES= socioeconomic status; PA= physical activity.					

Supplementary file 4. Quality of included quantitative studies as assessed by the EPHPP tool

Study ID	Selection bias	Study design	Confounders	Blinding	Data collection methods	Withdrawals and drop-outs	Final Grade
Agostini et al (2018)	3 = Weak	1 = Strong	3 = Weak	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
Boldemann et al (2004)	3 = Weak	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Boldemann et al (2006)	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Brussoni et al (2017)	2 = Moderate	2 = Moderate	2 = Moderate	3 = Weak	1 = Strong	1 = Strong	**2 = Moderate**
Choi et al (2014)	3 = Weak	1 = Strong	2 = Moderate	3 = Weak	3 = Weak	1 = Strong	**2 = Moderate**
Christian et al (2019)	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Cosco et al (2014)	1 = Strong	2 = Moderate	1 = Strong	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
deWeger (2017)	2 = Moderate	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	N/A	**3 = Weak**
Ene-Voiculescu & Ene-Voiculescu (2015), Fjortoft (2004), Fjortoft (2001)	3 = Weak	1 = Strong	1 = Strong	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
Ernst (2014)	2 = Moderate	3 = Weak	3 = Weak	3 = Weak	3 = Weak	N/A	**3 = Weak**
Frenkel et al (2019)	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Fyfe-Johnson et al (2019)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	1 = Strong	N/A	**3 = Weak**
Gubbels et al (2018)	3 = Weak	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Luchs, & Fikus (2018)	3 = Weak	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Lysklett et al (2019)	3 = Weak	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	N/A	**3 = Weak**
Määttä at al (2019)	3 = Weak	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	N/A	**3 = Weak**
Määttä et al (2019)	3 = Weak	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	N/A	**3 = Weak**
Meyer et al (2017)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	3 = Weak	N/A	**3 = Weak**
Moen et al (2007)	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	3 = Weak	N/A	**3 = Weak**
Müller et al (2017)	3 = Weak	1 = Strong	3 = Weak	3 = Weak	1 = Strong	1 = Strong	**3 = Weak**
Ng et al (2020)	3 = Weak	1 = Strong	1 = Strong	3 = Weak	1 = Strong	1 = Strong	**2 = Moderate**
Olesen et al (2013)	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Sando (2019)	2 = Moderate	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
Sando & Sandseter (2019)	3 = Weak	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	N/A	**3 = Weak**
Scholz & Krombholz (2007)	3 = Weak	3 = Weak	1 = Strong	3 = Weak	3 = Weak	N/A	**3 = Weak**
Söderström at al (2013)	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Storli et al (2010)	3 = Weak	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**

Sugiyama et al (2012)	3 = Weak	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	N/A	**3 = Weak**
Torkar & Rejc (2017)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	3 = Weak	N/A	**3 = Weak**
Weisshaar et al (2006)	2 = Moderate	3 = Weak	2 = Moderate	4 = Weak	1 = Strong	N/A	**3 = Weak**
Wright (2019)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	3 = Weak	N/A	**3 = Weak**

Supplementary file 5. Findings per eligible study
Quantitative

Table 1. Nature-based ECE on physical activity								
Study details (Author, year and country) **Sample size** (n children / n ECE)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Accelerometer								
Nature-based ECE								
Müller et al (2017), Canada. E: 43 children / 1 ECEs C: 45 children / 1 ECEs	Controlled before & after	SB and MVPA ActiGraph GT1M measured for 5 consecutive school days on three separate occasions: Oct (start of school yr), Jan and Apr (end of school yr). Cut points not described	SB (mins/ ECE day)	E: Oct= 167 Jan= 174 C: Oct= 178 Jan= 178	Apr= 151 Apr= 152	Within-group seasonal differences, but no between-group differences. (inferential statistics not provided)	▲	Weak
			MVPA (mins/ ECE day)	E: Oct= 74 Jan= 79 C: Oct = 79 Jan= 79	Apr = 68 Apr= 62	As above.	▲	
Fyfe-Johnson et al (2019), USA. E: 20 children / 1 ECEs	Controlled cross-sectional	PA and SB ActiGraph GT3X+ accelerometer worn for a minimum if 5 days (inc 1 weekend).	Habitual PA (mins/ day) SB Light	 E: 467 (60 SD) C: 453 (51 SD) E: 91.6 (13 SD) C: 102 (10 SD)	Mead diff: 14.4, (95% CI: -29.1, 58.0) -10.1 (95% CI: -19.2, -1.0)	Children who attended nature-based ECE engaged in more SB, and less light PA and MVPA.	▼	Weak

C: 13 children (waitlist control or 2-hour nature-based, outdoor enrichment class provided by experimental ECEs		Weartime for total PA was 656 (59 SD), C= 667 (59 SD) Pate et al. (2006) cut points	MVPA	E: 97.4 (16 SD) C: 113 (24 SD)	-15.5 (95% CI: -31.9, 0.87)			
			Habitual Weekday PA (mins/day) SB Light MVPA	 E: 468 (66 SD) C: 461 (54 SD) E: 93.5 (18 SD) C: 101 (15 SD) E: 97.1 (21 SD) C: 112 (30 SD)	 6.9 (95% CI: -40.1, 54.0) -7.3 (95% CI: -20.1, 5.4) -14.9 (95% CI: -36.3, 6.5)	As above.	 ▼	
			Habitual Weekend PA (mins/day) SB Light MVPA	 E: 486 (65 SD) C: 453 (51 SD) E: 88.7 (14 SD) C: 103 (15 SD) E: 95.8 (16 SD) C: 113 (22 SD)	 33.0 (95% CI: -14.8, 80.9) -14.2 (95% CI: -25.9, -2.4) -17.7 (95% CI: -33.8, -1.5)	As above.	 ▼	
			PA (mins/ ECE day – 9.00-13.00) SB Light MVPA	 E: 153 (19 SD) C: 166 (13 SD) E: 31.8 (11 SD) C: 32.7 (5 SD) E: 33.2 (15 SD) C: 34.7 (7 SD)	 -13.5 (95% CI: 63.3, 54.2) -0.9 (95% CI: -2.1, 0.64) -1.5 (95% CI: -2.8, 1.2)	Children who attended nature-based ECE engaged in lower SB compared to traditional ECE during the ECE day, but traditional ECE engaged in higher light and MVPA.	 ▲ ▼ ▼	

			Sedentary bouts (ECE day) Bout, total number Bouts, number per day Bouts, total length Bout, average length	 E: 6.3 (3 SD) C: 6.4 (4 SD) E: 1.9 (1 SD) C: 2.0 (1 SD) E: 88.9 (47 SD) C: 100 (59 SD) E: 12.8 (5 SD) C: 16.1 (3 SD)	 -0.05 (95% CI: -2.9, 2.8) -0.11 (95% CI: -0.94, 0.73) -11.3 (95% CI: -54.4, 31.7) -3.3 (95% CI: -6.7, 0.13)	Children who attended nature-based ECE had similar total bouts and number of bouts per day to the control group. The bout total and average length were also higher in the control group.	▼	
ECE natural playgrounds								
Brussoni et al (2017), Canada. E: 48 children / 2 ECE	Uncontrolled before & after (mixed methods)	MVPA ActiGraph GT3X/GT3X+ worn during scheduled outdoor time (20 mins). Pate et al. (2006) cut points	MVPA (mins/ outdoor time)	Not presented.	- 1.32 min, 0.37 SE, p< 0.001	There was a significant decrease in time spent in MVPA from T1 to T2 across ECE's.	▼	Moderate
Luchs, & Fikus (2018), Germany. E: 17 children / 1 ECE	Cross-sectional	Gait cycles Microprocessor-based pedometer (StepWatch, Orthocare Innovations, Washington DC, USA) Worn twice for 45 minutes, once on the nature playground and once on the traditional playground.	Gait cycles/mins at playground	E: 25 (4.99 SD) C: 28.55 (9.60 SD) p= 0.109, d = 0.54)		No significant difference in mean gait cycles/min between the nature and traditional playground.	▼	Weak

Storli et al (2010), Norway. E: 16 children / 1 ECE	Cross-sectional	CPM ActiGraph (model not described) Worn for three separate days over 6 months, including 2 days of outdoor activity on the pre-school playground (winter and spring) and one day in nature (spring). Wear time varied between 102–136 minutes Cut points not described,	Mean CPM	E: (spring) 1292 (307 SD) C: (spring) 1261 (426 SD) C: (winter) 1496 (475 SD) (p= 0.01)		There is an association between the levels of PA for the natural environment and traditional (spring and winter) playgrounds meaning PA levels are similar across the environments.	►	Weak
Torkar & Rejc (2017), Slovenia. E: 25 children / 1 ECE	Cross-sectional	Distance (km) Measured using GPS for 20 mins.	Distance (km)	E: 0.72 (0.49 SD) C: 0.49 (0.19 SD) (p= 0.132, r= 0.21)		There were no significant differences between the forest and traditional playground.	▲	Weak
Natural element within ECE								
Ng et al (2020), Australia. E: 159 children / 6 ECE C: 138 children / 5 ECE	Controlled before and after	PA ActiGraph GTX3+ worn during ECE days ECE monitoring days were considered valid based on at least 1 day at ECE with 75% wear time Pate et al. (2006) cut points	Total PA min/ ECE day) MVPA min/ ECE day)	β= 14.46, p< 0.01 β= 10.04, p< 0.01		Natural grassed area was positively associated with Total PA and MVPA. Non-significant time x group interaction for natural elements on Total PA and MVPA (regression coefficients not presented)	►	Weak

Boldemann et al (2006), Sweden. E: 199 children / 11 ECE	Cross-sectional	Step counts Yamax Digiwalker SW-200, MLS 2000 pedometer. Wear time not detailed.	Step counts/ min ECE day	High environment = 21.6 (95% CI: 20.6–22.5) Low environment = 17.7 (95% CI: 16.8–18.6) p<0.001		High environment score increased step count	▲	Weak
Christian et al (2019), Australia. E: 678 children / 48 ECE	Cross-sectional	Total PA Actigraph GT3TX+ Valid data included at least 1 day at ECE with 75% wear time. Data was averaged for children who attended more than 1 day during the 7-day monitoring period. Pate et al. (2006) cut points	Total PA (min/ ECE day)	% < 3m vegetation: β<-0.01 (95% CI: -0.22,0.21), p= 0.96) % > 3m vegetation: β = 0.02 (95%CI: -0.28, 0.32), p=0.89		Shade-related variables (vegetation < 3 metres in height and vegetation > 3 metres in height) were not significantly associated with minutes/day of total PA.	▲	Weak
			MVPA (min/ ECE day)	% < 3m vegetation: β = -0.01 (95% CI: -0.18, 0.16), p=0.91 % > 3m vegetation: β = 0.08 (95%CI: -0.16, 0.32), p=0.52		As above for MVPA	▲	
deWeger (2017), Australia.	Cross-sectional	Total PA and MVPA (min/day at ECE), cpm and step counts	Total PA (min/ ECE day)	intercept= 59.5, coefficient= 3.5, 1.8 SE, t= 1.89, p= 0.060		No significant association between setting with natural elements on total PA.	▲	Weak

E: 274 children / 12 ECE		Actigraph GT3X+ Accelerometers were worn for one ECE week (range of 1-5 days). Mean wear time was 390 minutes (87.4) or for 6.5 hours (1.5). Pate et al. (2006) cut points	MVPA (min/ ECE day)	intercept= 10.3, coefficient= 1.7, 1.2 SE, t= 1.37, p= 0.17		As above for MVPA.	▲	
			Mean CPM / ECE day	intercept= 102000.5, coefficient= 4511.9, 5683.5 SE, t= 0.79, p= 0.43		As above for CPM.	▲	
			Step counts / ECE day	intercept= 2889.9, coefficient= 199.5, 89.8 SE, t= 2.22, p= 0.027		There was a positive association between settings with natural elements and step counts.	▲	
Gubbels et al (2018), Netherlands. E: 151 children / 22 ECE	Cross-sectional	SB, MVPA and CPM Actigraph GT3X+ Children were asked to wear the monitor for 7 consecutive days during their waking hours. Minimal wear time per day was 360 minutes and children had to have at least one valid ECE day to be included. Pate et al. (2006) cut points	Habitual SB %	β= −0.31, p< 0.001		Natural elements were significantly and positively associated with a reduction in percent time spent in SB	▲	Weak
			Habitual MVPA %	β= 0.27, p< 0.01		Natural elements were significantly and positively associated with an increased percent time spent in MVPA	▲	
			Habitual Mean CPM	β= 0.21, p< 0.01		Natural elements were significantly and positively associated with increased CPM.	▲	

Määttä et al (2019), Finland. E: 864 children / 66 ECE	Cross-sectional	Total PA Actigraph GT3X Worn for 7 days, 24-hours/day. A minimum wear time of 240 min during preschool hours was set. Evenson et al. (2008) cut points.	Total PA (min/hour in ECE)	Grass: β= 0.31, (95%CI: −0.84 - 1.46) Forest: β= −0.59, (95%CI: −1.87 - 0.69) Trees: β= -0.34, (95%CI: −2.13 - 1.45) Rocks: β= 0.01, (95%CI: −1.21 - 1.24)		There were no significant main or effect for grass, forest, trees or rocks	▲ ▼ ▼ ▲	Weak
Määttä et al (2019b), Finland. E: 655 children / 66 ECE	Cross-sectional	Sedentary Time As above.	Sedentary time (min/hour in ECE)	Frequency of nature trips β= −1.026 (95%CI: −1.804, −0.248), p= 0.010		Frequency of nature trips was associated with children's lower sedentary time.	▲	Weak
Olesen et al (2013), Denmark. E: 441 children / 42 ECE	Cross-sectional	MVPA ActiGraph accelerometer Children wore the monitors for 1 week. Minimum wear time was	MVPA (percent/ ECE day)	Vegetation: - 0.7; 95% CI: - 1.3 to -0.0, p= 0.04)		The multilevel analysis showed that the daily percentage of MVPA was significantly negatively associated with vegetation	▼	Weak

		3 pre-school days, with at least 3 hours of measurement. Median wear-time was 4 weekdays, 7.15 hours per day. Evenson et al. (2008) cut points.		Hilly landscape - 0.4; 95% CI: -1.1 to 0.2, p= 0.18.		The multilevel analysis showed that the daily percentage of MVPA was no association with hilly landscape.	▼	
Sugiyama et al (2012), Australia. E: 89 children / 10 ECE	Cross-sectional	MVPA and SB ActiGraph GT1M a Worn for 3 days at ECE. Minimum wear time was 2 days for at least 4 hours during the ECE day. Average wear time was 6 hours 40 minutes per ECE day. Sirard et al. (2005) cut points.	MVPA (min/ outdoor time)	Mostly natural surface: β= −5.8, (95% CI: −9.9, −1.7), p<0.01		Children attending ECE's with mostly natural surfaces were found to engage in significantly less MVPA compared with ECE with mostly "built" surfaces.	▼	Weak
				More vegetation: β= −1.2, (95% CI: −5.9, 3.5)		No association.	▼	
				Some gradient: β= 1.3, (95%CI: −4.5, 7.0)		As above.	▲	
				Much shade: β= 2.3, (95%CI: −3.5, 8.0)		As above.	▲	
			SB (min/ outdoor time)	Mostly natural surface: β= 8.0, (95% CI: −1.4, 17.4)		Natural surfaces, vegetation, gradient, and shade were not associated with SB.	▼	
				More vegetation: β= 2.3, (95% CI: −7.0, 11.6)			▼	
				Some gradient: β= −2.4, (95% CI: −13.7, 8.9)			▲	

				Much shade: β= −0.9, (95% CI:−12.6, 10.8)			▲	
Observational								
Nature-based ECE								
Meyer et al (2017), Canada. E: 46 children / 3 ECE C: 35 children / 2 ECE	Controlled cross-sectional	PA and PA types OSRAC-P Sampling Observation System which includes coding for body movements (stationary, slow-easy, moderate, and vigorous movements) and specific activity types (including climb, crawl, jump/skip, push/pull, rough and tumble, run, sit/squat, stand, throw, walk, and other). 2 students were observed at a time for 30-second intervals (5 sec observation, 25 sec coding). Observations occurred every 30 seconds for a period of 5 minutes which resulted in 20 observations. This was then repeated.	PA frequencies: Stationary Slow-easy Moderate Vigorous	 E:0.56 (0.15 SD) C: 0.84 (0.02 SD) E:0.30 (0.08 SD) C: 0.16 (0.02 SD) E:0.12 (0.08 SD) C: 0 (0 SD) E: 0.02 (0 SD) C: 0 (0 SD)		Children in the nature kindergarten were less stationary and engaged in more slow-easy and moderate physical activity compared to the control ECE.	N/A	Weak
			PA types: (frequencies) Sit/Squat Walk Stand Fine Motor Eat	 E: 0.19 (0.13 SD) C: 0.53 (0.09 SD) E: 0.17 (0.02 SD) C: 0.06 (0.01 SD) E: 0.14 (0.08 SD) C: 0.16 (0 SD) E: 0.14 (0.06 SD) C: 0.12 (0.09) E: 0.08 (0.03 SD) C: 0 (0 SD)			N/A	

			Lie Down Push/Pull Rough & Tumble Run Climb Jump Throw Crawl Balance Other	E: 0.01 (0.01 SD) C: 0 (0 SD) E: 0.01 (0.01 SD) C: E: 0 (0 SD) C: 0 (0 SD) E: 0.04 (0.02 SD) C: 0 (0 SD) E: 0.10 (0.07 SD) C: 0 (0 SD) E: 0 (0 SD) C: 0 (0 SD) E: 0.01 (0.01 SD) C: 0 (0 SD) E: 0.01 (0.01 SD) C: 0.01 (0.01 SD) E: 0.05 (0.04 SD) C: 0.01 (0.01 SD) E: 0.05 (0.02 SD) C: 0.10 (0 SD)				
Wright (2019), USA. 48 children / 2 ECE	Cross-sectional	PA Children were observed and recorded over 2 school years. A randomised time sampling protocol was used with 10 min intervals at five zones.	overall frequency / relative frequency (% each type of activity was out of total instances of all PA)			"manipulation" was the most frequent PA type observed. balance, run, sit stand and squat were less frequent.	N/A	Weak

		A sub-sample of the recordings was taken and coded at the 0:00, 1:00 and 2:00 mark for 20-second intervals. An adapted version of (OSRAC-P) was used to code the PA types.	Balance: Climb: Dig/Rake: Jump/Skip: Lie Down: Manipulation: Push/Pull: Resistive: Run: Sit: Stand: Squat: Throw: Walk:	34 / 7% 22 / 5% 19 / 4% 29 / 6% 9 / 2% 107 / 23% 21 / 4% 28 / 6% 34 / 7% 33 / 7% 38 / 8% 44 / 9% 16 / 3% 16 / 3%				
ECE natural playgrounds								
Cosco et al (2014), USA. E: not clear / 27 ECE	Uncontrolled before & after	PA Children's Activity Rating Scale (CARS) CARS allows trained observers to record children's PA on a five-point scale: 1) stationary or motionless, 2) stationary with limb or trunk movements, 3) slow-easy, 4) moderate, and 5) fast.	PA		Unstandardised (standardised effects) 0.113 (0.067), p= 0.001	At post-intervention there was an effect on children's PA.	▲	Weak
			Non sedentary PA		0.202 (1.22), p= 0.001	As above for non-sedentary PA.	▲	
			MVPA		0.061 (1.063), Non-sig	Non-significant	▲	
Natural elements within ECE								
Sando (2019), Norway.	Cross-sectional	PA Observational System for Recording PA in Children-Preschool (OSRAC-P)	PA (1-5)	3.2 (0.9 SD), (regression coefficient= 0.004)		Nature was not a statistically significant predictor of PA.	▲	Weak

E: 80 children / 8 ECE		PA is coded from 1 (stationary) to 5 (fast movement). 2 children were filmed per day. The 1st for 2 minutes followed by a 6-minute break, then the 2nd child. Filming alternated between each child until 6 video observations of each child were recorded. 480 video clips in the outdoor environment constituted a full sample. There was a total of 471 video clips in the final analysis.						
Sando & Sandseter (2019), Norway. E: 73 / 8 ECE	Cross sectional (mixed-methods)	PA and wellbeing (combined outcome) Wellbeing - Leuven Wellbing Scale measures wellbeing on a scale 1 (extremely low) -5 (extremely high). A score of 1 is when children exhibit high levels of discomfort (whining, screaming, sadness) and 5 is clear signs of happiness, relaxed and lively. Physical activity: see above, OSRAC-P which codes PA from 1	PA and wellbeing	Nature: No association Sand: b =−0.027, (95% CI =−0.043–0.011), p= 0.001. Nature materials: b =−0.008, (95% CI =−0.015–0.001), p= 0.028. Water: no association Mud: no association		Nature is not associated with observations with high wellbeing and PA.	▼	Weak

		(stationary) to 5 (fast-movement).						

Abbreviations: E= experimental; C= control; n= number; ECE = early childhood education (includes preschools, day care, kindergarten etc.); PA= physical activity; MVPA= moderate to vigorous PA; SB= sedentary behaviour; CPM= counts per minute; Yr= Year; min = minutes; SD= standard deviation; SE= standard error; CI= confidence intervals.

Effect direction explained:
▲: positive association
►: no change/ conflicting findings
▼: negative health association
▲: positive association and statistical significance ($p<0.05$)
▼: negative association and statistical significance ($p<0.05$)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Table 2. Nature-based ECE on motor competence

Study details (Author, year and country) **Sample size** (n of children / n ECE settings for exp and con)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Nature-based ECE								
Agostini et al (2018), Italy. E: 41 children / 7 teachers / 1 school	Controlled Before & After	Body function, gross motor skills and fine motor skills Kuno Beller Developmental Tables completed by educators	Body Function	T1 (Jan 2014) E:11.02 (0.81 SD) C:10.15 (1.03 SD)	T4 (May 2015) 12.81 (0.71 SD) 12.39 (1.24 SD) p= 0.010; ηp^2= 0.27	There was a significant time x group interaction on children's body function.	▲	Weak

C: 52 children / 13 teachers / 1 school		which assesses development in 8 developmental areas: Body Function, Awareness of the Surrounding Environment, Social and Emotional Development, Play, Language, Cognitive Development, Gross and Fine Motor Skills.				There were no significant differences between groups at T4.		
			Gross Motor Skills	E:11.79 (1.01 SD) C:10.87 (0.91 SD)	13.32 (0.80 SD) 12.96 (1.07 SD) p= 0.021; ηp2= 0.24	As above.	▲	
			Fine Motor Skills	E:10.86 (0.76 SD) C:10.01 (1.34 SD)	12.73 (0.88 SD) 12.56 (1.28 SD) p= 0.000; ηp2= 0.15.	As above.	▲	
Ene-Voiculescu & Ene-Voiculescu (2015), Fjortoft (2004), Fjortoft (2001), Norway. E: = 46 children / 1 kindergarten C: 29 children, / 2 kindergartens	Controlled Before & After	Motor fitness The EUROFIT Physical Fitness Test which consists of: flamingo balance test (standing on 1 foot - balancing); plate tapping (tapping of 2 plates alternatively-speed of limb movement); sit and reach (flexibility); standing broad jump (jumping for distance from a standing start – explosive strength); sit-ups (max n of sit-ups in 30 secs); bent arm hang (from a bar- functional strength); shuttle run (running and turning,	Flamingo balance test / n of instabilities in 30 secs	E: 4.7 (0.8 SE) C: 4.0 (0.6 SE)	E: 1.5 (0.3 SE), p<0.001 C: 3.3 (0.7 SE)	At post-test, there were significant differences in the intervention group compared to the control group in the Flamingo balance test (p< 0.001).	▲	Weak
			Plate tapping / time in secs for 50 taps	E: 35.0 (1.9 SE) C: 29.9 (1.1 SE)	E: 28.1 (1.2 SE), p<0.001 C: 27.4 (2.6 SE)	No significant differences at post-test.	▼	
			Sit and reach / cm	E: 24.9 (0.8 SE) C: 25.3 (1.0 SE)	E: 24.4 (0.8 SE) C: 25.5 (0.9 SE)	As above.	▼	
			Standing broad jump / cm	E: 102.8 (2.9 SE) C: 103.1 (4.3 SE)	E: 113.1 (3.6 SE), p<0.001 C: 111.3 (3.8 SE), p<0.01	As above.	▲	

		shuttle - speed and agility) Beam walking to test dynamic balance and Indian skip (clapping right knee with left hand and vice versa - coordination), which were added.	Sit-ups / reps.30 secs	E: 5.3 (0.6 SE) C: 5.9 (0.8 SE)	E: 6.5 (0.6 SE) p<0.01 C: 7.0 (1.1 SE)	As above.	▼	
			Bent arm hang / sec	E: 2.6 (0.4 SE) C: 2.6 (0.6 SE)	C: 7.0 (1.0 SE), p<0.001 C: 5.4 (1.1 SE), p<0.001	As above.	▲	
			Beam walking / sec	E: 11.4 (1.4 SE) C: 7.7 (0.8)	E: 7.5 (0.7 SE), p<0.01 C: 7.2 (1.1 SD)	As above.	▼	
			Indian skip / reps.30 secs	E: 21.8 (2.2 SE) C: 27.8 (2.4 SE)	E: 43.6 (1.9 SE), p<0.001 C: 37.2 (1.8 SE), p<0.001	At post-test, there were significant differences in the intervention group compared to the control group in the Indian skip co-ordination test (p< 0.01).	▲	
			Shuttle run run/sec	E: 31.9 (0.7 SE) C: 30.7 (0.8 SE)	E: 29.7 (0.5 SE), p<.01 C: 30.3 (0.7 SE)	No significant differences at post-test.	▲	
Müller et al (2017), Canada. E: 43 children / 1 nature-kindergarten	Controlled before & after	Perceived physical competence, and locomotor and object control skills. Subscale of the Pictorial Scale of	Perceived Physical Competence	E: 18.72 (0.47 SE) C: 18.58 (0.44 SE)	E: 19.03 (0.48 SE) C: 19.47 (0.44 SE) p= 0.45, η2= 0.01	At post-test there was a small and non-significant effect	▼	Weak

C: 45 children / 1 traditional kindergarten		Perceived Competence and Social Acceptance for Young Children (six items) - children were asked to indicate who they are more like based on two descriptions of children (one competent and one not). Each item was scored on a four-point scale, where 4 indicates a high degree of perceived competence and 1 indicates a low score. TGMD-2 - assesses 6 locomotor and 6 object control skills. Scored either 1 or 0 depending on whether component was performed correctly.	Locomotor skills	E: 24.68 (1.01 SE) C: 24.61 (0.94 SE)	E: 28.03 (0.82 SE) C: 25.72 (0.80 SE) p= 0.03, η2= 0.06	At post-test there was a moderate and significant effect	▲	
			Object control skills	E: 21.71 (0.98 SE) C: 23.05 (0.91 SE)	E: 23.97 (0.89 SE) C: 23.05 (0.91 SE) p= 0.15, η2= 0.03	At post-test there was a small and non-significant effect	▲	
Lysklett et al (2019), Norway. E: 43 children / 4 preschools C: 49 children / 4 preschools	Controlled cross sectional	Motor competence Assessed using the Movement Assessment Battery (MABC-2). The test includes 8 subtests divided into 3 categories: 1) manual dexterity (posting coins, threading beads and drawing a line into a trail), 2) ball skills (catching beanbag and rolling ball into goal),			Mean difference			Weak
			Manual dexterity	E: 3.72 (2.99 SD) C: 3.29 (2.67 SD)	0.43 (95% CI: −0.74–1.59), p= 0.498	No significant differences in scores between the nature and traditional preschools for total and subtest scores.		
			Ball	E: 2.60 (2.34 SD) C: 2.41 (1.67 SD)	0.20 (95% CI: −0.64–1.03), p= 0.641		▼	
			Static and dynamic balance	E: 1.08 (1.71 SD) C: 0.94 (1.58 SD)	0.14 (95% CI: −0.53–0.82), p= 0.678			

		and 3) static and dynamic balance (one-leg balance, walking heel raised and jumping over cord). Children are scored from 0-5. The total score sums the eight tests with a score of 0 the best and 40 the poorest.	Total	E: 7.41 (4.91 SD) C: 6.64 (3.72 SD)	0.76 (95% CI: −1.03–2.56), p= 0.399			
		The assessment for fitness consisted of 9 subtests: standing broad jump, Jumping on two feet, Jumping on one foot, Throwing a tennis ball (m), Putting a medicine ball, Climbing wall bars, Shuttle run, 20 m sprint, Reduced Cooper test. A total test score was calculated and transformed into z-scores (standardized scores).	Standing broad jump (cm) Jumping on two feet (s) Jumping on one foot (s) Throwing a tennis ball (m) Putting a medicine ball (m)	E: 94.78 (14.07 SD) C: 97.63 (15.59 SD) E: 6.16 (3.58 SD) C: 5.18 (1.61 SD) E: 5.48 (2.19 SD) C: 4.85 (1.19 SD) E: 6.00 (2.17 SD) C: 6.21 (1.88 SD) E: 1.88 (0.49 SD) C: 1.96 (0.43 SD)	Mean difference −2.86 (95% CI: −9.26–3.55), p= 0.378 0.98 (95% CI: −0.22–2.18), p= 0.108 0.63 (95% CI: −0.22–1.49), p= 0.144 −0.21 (95% CI: −1.06–0.64), p= 0.623 −0.08 (95% CI: −0.27–0.11), p= 0.379	Children attending the traditional preschools performed better in the shuttle run, reduced Cooper test and the total score compared to the nature playground. The rest were non-significant.	▼ ▼ ▼ ▼ ▼	

			Climbing wall bars (s)	E: 32.32 (14.60 SD) C: 31.21 (11.38 SD)	1.11 (95% CI:−4.37–6.59), p= 0.688		▼	
			Shuttle run (s)	E: 31.40 (3.96 SD) C: 30.00 (2.45 SD).	1.40, 95% CI: 0.05–2.74, p= 0.043		▼	
			20 m sprint (s)	E: 5.66 (0.48 SD) C: 5.53 (0.57 SD)	0.13 (95% CI: 0.13 – −0.08), p= 0.232		▼	
			Reduced Cooper test (m)	E: 740.09 (120.44 SD) C: 817.56 (105.32 SD)	77.47, 95% CI: −124.22– −30.71, p= 0.001),		▼	
			Total test score (z)	E: −0.12 (0.65 SD) C: 0.17 (0.57 SD)	−0.29, 95% CI: −0.55– −0.04, p= 0.025		▼	
Scholz & Krombholz (2007), Germany E: 45 children / 10 forest kindergartens C: Rural = 42 children / 2 ECE; Urban = 42 children / 2 ECE	Controlled cross-sectional	Fundamental movement skills (test not described) Consisted of the following domains: balancing forward (balance); balancing backward (balance); jumping left and right; (coordination, speed); long jump; (coordination, speed); jumping forwards on one leg (coordination, endurance); hanging on	Balancing forward (n of correct steps)	E:22.5 (1.7 SD) C (R): 20.5 (3.5 SD) C (U): 19.4 (3.6 SD) p<0.000		There was a significant higher performance in forest nurseries vs conventional rural and urban nurseries for balancing forwards and backwards, hanging on pull up bar, and one-leg jump forward on left.	▲	Weak
			Balancing backward (n of correct steps)	E: 51.5 (10.1 SD) C (R): 39.9 (10.9 SD) C (U): 35.5 (14.3 SD) p<0.000			▲	

		pull up bar (strength endurance); shuttle run (speed, coordination)	Jumping left and right (n of jumps)	E: 29.9 (6.0 SD) C (R): 31.1 (7.3 SD) C (U): 27.0 (7.1 SD) p=0.012			►	
			Long jump (distance in cm)	E: 94.0 (16.1 SD) C (R): 102.4 (18.4 SD) C (U): 94.0 (18.7 SD)			►	
			Hanging on pull up bar (time in seconds - max 30 sec)	E: 25.6 (6.2 SD) C (R): 20.7 (7.7 SD) C (U): 19.7 (7.0 SD) p<0.000			▲	
			Shuttle run (time in seconds)	E: 9.6 (1.2 SD) C (R): 9.1 (0.8 SD) C (U): 10.2 (1.5) p<0.000			▼	
			Jumping forwards on one leg (n of jumps on each leg – max 20)	Right: E: 17.5 (4.4 SD) C (R): 17.2 (4.9 SD) C (U): 16.0 (6.0 SD)			▲	
				Left: E: 17.8 (4.5) C (R): 16.8 (5.3), C (U): 14.1 (6.8) p=0.007			▲	

Ernst (2014), USA. E: 46 educators	Cross-sectional	Physical development Questionnaire (not described) on importance of natural outdoor settings on children's cognitive, social, and physical development and their appreciation for the environment. Responses were provided on a five-point scale, ranging from one (strongly disagree) to five (strongly agree)	Physical development (1-5)	4.39 (1.31 SD), r= 0.05		Educators agreed that experiences in natural settings were important for children's physical development. There was no association between frequency of nature experiences and belief regarding importance of outdoor settings for physical development.	▲	Weak

Abbreviations: E= experimental; C= control; n= number; ECE = early childhood education (includes preschools, day care, kindergarten etc.); SD= standard deviation; SE= standard error; CI= confidence intervals; cm= centimetres; sec= seconds; R= rural; U= urban

Effect direction explained:
▲: positive association
►: no change/ conflicting findings
▼: negative health association
▲: positive association and statistical significance ($p<0.05$)
▼: negative association and statistical significance ($p<0.05$)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Table 3. Nature-based ECE on weight status

Study details (Author, year and country)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**

Sample size (n of children / n ECE settings for exp and con)								
Natural elements within ECE								
Söderström at al (2013), Sweden. E: 172 children / 9 ECEs	Cross-sectional	BMI Weight = digital scale, height = measuring tape	BMI	Low OPEC Overweight= 16% Normal weight= 82% High OPEC Overweight= 7% Normal weight= 87% p= - 0.07		Outdoor environment quality was not significantly associated with BMI or waist.	▲	Weak
		Waist Measuring tape	Waist (cm)	Low OPEC: 52.6 (3.5 SD) High OPEC: 52.2 (3.5 SD) p= 0.25			▲	

Abbreviations: E= experimental; C= control; n= number; ECE = early childhood education (includes preschools, day care, kindergarten etc.); SD= standard deviation; BMI= body mass index; cm= centimetres; OPEC= outdoor Play Environmental Categories

Effect direction explained:
▲: positive association
►: no change/ conflicting findings
▼: negative health association
▲: positive association and statistical significance ($p<0.05$)

▼: negative association and statistical significance ($p<0.05$)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Table 4. Nature-based ECE on Sleep

Study details (Author, year and country) **Sample size** (n of children / n ECE settings for exp and con)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Nature-based ECE								
Choi et al (2014), South Korea. E: 18 children / 1 ECE C: 19 children / ECE	Controlled Before & After study	Sleep Parents competed the CSHQ which consists of 33 items with a 3 point scale, "usually (5–7 times a week)", "sometimes (2–4 times a week)", and "rarely (0–1 time a week)". This questionnaire consists of 8 domains: bedtime resistance, sleep onset delay, sleep duration, sleep anxiety, night wakings, parasomnia, sleep-	Total score of CSHQ	E: 51.6 ± 8.2 C: 55.6 ± 6.6	E: 47.7 ± 5.7, p= 0.02 C: 55.8 ± 6.5, p= 0.92 Between group: p< 0.01	After post-test, the CSHQ total score, sleep disordered breathing and daytime sleepiness were significantly lower in children from the forest kindergarten program compared with the regular kindergarten program. There was no significant difference in total sleep time or other sub-scales.	▲	Moderate
			Total sleep time (hours)	E: 10.5 ± 1.1 C: 10.7 ± 1.1	E: 10.5 ± 1.0, p= 0.68 C: 10.4 ± 0.9, p= 0.21		▲	
			Bedtime resistance	E: 11.8 ± 2.6 C: 12.7 ± 2.5	E: 11.3 ± 2.4, p= 0.34 C: 12.8 ± 2.2, p= 0.98		▲	

		disordered breathing, and daytime sleepiness. These domain scores are accumulated for a total CSHQ Score. Total sleep time was also reported.	Sleep onset delay	E: 1.3 ± 0.6 C: 1.2 ± 0.5	E: 1.2 ± 0.4, p= 0.08 C: 1.4 ± 0.7, p= 0.36		▲	
			Sleep duration	E: 3.7 ± 1.1 C: 4.1 ± 1.4	E: 3.3 ± 0.6, p= 0.13 C: 3.7 ± 1.3, p= 0.37		▲	
			Sleep anxiety	E: 7.1 ± 2.0 C: 7.4 ± 1.8	E: 6.5 ± 2.0, p= 0.28 C: 7.5 ± 1.5, p= 0.84		▲	
			Night wakings	E: 3.6 ± 0.8 C: 3.6 ± 0.8	E: 3.5 ± 0.4, p= 0.71 C: 3.6 ± 1.0, p= 0.99		▲	
			Parasomnia	E: 9.2 ± 2.0 C: 10.0 ± 1.8	E: 8.6 ± 1.5, p= 0.11 C: 9.3 ± 1.9, p= 0.12		▲	
			Sleep disordered breathing	E: 3.3 ± 0.6 C:3.4 ± 0.8	E: 3.1 ± 0.5, p= 0.16 C: 3.7 ± 1.0, p= 0.10 Between group: p= 0.04		▲	
			Daytime sleepiness	E: 11.6 ± 2.5 C: 13.3 ± 2.9	E: 9.8 ± 1.0, p= 0.02 C: 13.7 ± 3.5, p= 0.52 Between group: p< 0.01		▲	

Natural elements within ECE								
Söderström at al (2013), Sweden. E: 172 children / 9 ECE	Cross-sectional	Sleep A sleep diary was completed for one week by the children's parents. Parents recorded the time the children woke up and the time they went to sleep. Sleep time was calculated as a mean of the seven days.	Mean sleep time (minutes)	Low OPEC (n= 103): 642 (32 SD) High OPEC (n= 66): 658 (44 SD) p= 0.03		Outdoor environment quality was significantly associated with night sleep (p = 0.03)	▲	Weak

Abbreviations: E= experimental; C= control; n= number; ECE = early childhood education (includes preschools, day care, kindergarten etc.); CSHQ= Children's Sleep Habits Questionnaire; OPEC= outdoor Play Environmental Categories

Effect direction explained:
▲: positive association
►: no change/ conflicting findings
▼: negative health association
▲: positive association and statistical significance ($p<0.05$)
▼: negative association and statistical significance ($p<0.05$)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Table 5. Nature-based ECE on UV Exposure

Study details (Author, year and country) **Sample size** (n of children / n ECE settings	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**

for exp and con)								
Natural elements within ECE								
Boldemann et al (2004), Sweden. E: 64 children / 2 ECE	Cross-sectional	UV Exposure Measured using a Dosimeter (Biosense VioSpor blue line, type III 0.8–33 MED). Each child wore 2 Dosimeters attached to each shoulder using safety pins. They were worn during the school day.	UV exposure per day (JCIE/m^2)	Site 1: 222 JCIE/m^2, 15.3 % (95% CI 14.3–17.5, $p<0.05$) Site 2: 175 JCIE/m^2, 13.3 % (95% CI 9.9–14.6, $p<0.05$)		The was a statistically significant difference in UVR exposure between site 1 and site 2.	▲	Weak
Boldemann et al (2006), Sweden. E: 199 children / 11 ECE	Cross-sectional	UV Exposure Measured using a Polysulphone dosimeter (Diffey, 1984; Herlihy et al., 1994) The Dosimeter was pinned to the right shoulder and worn during school hours.	UV Exposure (J/m^2)	Low environment: ECE 3: 160 (95%CI:130–190) ECE 4: 241 (95%CI:200–281) ECE 6: 156 (95%CI:115–196) ECE 7: 83 (95%CI: 67–98) ECE 8: 269 (95%CI:214–324) ECE 10: 243 (95%CI:217–268) High environment: ECE 1: 104 (95%CI: 95–113) ECE 2: 129 (95%CI:104–154) ECE 5: 289 (95%CI:230–348)	Daily UV exposures ranged between 74 and 292 J/m	Outdoor environment quality was significantly associated with UV Exposure.	▲	Weak

				ECE 9: 292 (95%CI:232–351) ECE 11: 196 95%CI: 177–215)				
Christian et al (2019), Australia. E: 678 children / 48 ECE	Cross-sectional	UV Exposure Measured using a Polysulphone film mounted cardboard holders (UV badge) The UV badge was attached to the child's left shoulder and worn each day whilst at ECE for up to 3 days.	UV exposure (J/m^2) per average day of ECE.	% <3 m vegetation: β = -2.26 (95%CI -3.03, -1.49); p<0.01 % >3m vegetation: β = 0.91 (95%CI -12.46, 14.28), p= 0.89		ECE centre vegetation was significantly negatively associated with children's UVR exposure. For every 1% increase in centre vegetation, children's UVR exposure decreased by 2.3 J/m2 per day at ECE (p <0.01).	▲	Weak

Abbreviations: E= experimental; C= control; n= number; ECE = early childhood education (includes preschools, day care, kindergarten etc.); SD= standard deviation; SE= standard error; CI= confidence intervals.

Effect direction explained:
▲: positive association
►: no change/ conflicting findings
▼: negative health association
▲: positive association and statistical significance (p<0.05)
▼: negative association and statistical significance (p<0.05)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Table 6. Nature-based ECE on unintended consequences								
Study details (Author, year and country) **Sample size** (n of children / n ECE settings for exp and con)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Nature-based ECE								
Frenkel et al (2019), USA. E: 71 children / 5 ECE C: 70 children / 4 ECE	Controlled cross-sectional	Illness and injury Educators completed a standardised weekly illness and injury tracking log developed for this study. An illness episode was when a child was absent for at least 1 day due to illness (fever, respiratory, stomach, other).	Illness total Fever Respiratory Stomach Other	E: 1.49 C: 1.62 (age adjusted IRR: 0.93, 95% CI: 0.64, 1.34). E: 0.25 C: 0.47 E: 0.92 C: 1.01 E: 0.29 C: 0.37 E: 0.18 C: 0.07		No significant difference in the incidence of total illness between nature ECE and traditional ECE	▲	Weak
		An injury was counted if it required first-aid attention from teachers	Total injury	E: boys= 0.94 girls= 1.87 C: boys= 0.96 girls= 0.34		No significant difference in minor injury was found between boys at nature and traditional ECE. Girls at nature ECE had a significantly higher incidence of minor injury compared	▲ (boys) ▼ (girls)	

				boys: (age-adjusted IRR: 1.46, 95% CI: 0.59, 3.6) Girls: (age-adjusted IRR: 5.91, 95% CI: 1.98, 17.7).		with girls at traditional ECE.		
			Open wound/cut	E: boys= 0.60 girls= 1.31 C: boys= 0.48 girls= 0.23				
			Sprain	E: boys= 0 girls= 0 C: boys= 0 girls= 0				
			Child Bite	E: boys= 0.17 girls= 0 C: boys= 0 girls= 0				
			Other	E: boys= 0.17 girls= 0.56 C: boys= 0.48 girls= 0.11				

Moen et al (2007), Norway. E: 267 children / 37 ECE C: 264 children / 32 ECE	Controlled cross-sectional	Sickness absenteeism Parent noted daily reports of sickness absenteeism Absenteeism refers to the ratio of the total number of sickness absenteeism days to the sum of the number of sickness absenteeism days and the number of days the child was attending the day care centre during the study period.	Sickness absenteeism	estimate = - 0.0083, SE= 0.1830, t= 20.045, p> 0.05		No statistically significant difference in sickness absenteeism between the outdoor ECE and regular day ECE.	▼	Weak
Weisshaar et al (2006) E: 506 children / 25 ECE C: 1201 children / 28 ECE	Controlled cross-sectional	Tick bites and borreliosis Self- report questionnaire. Presence of at least 1 tick bite (yes/no). Presence of borreliosis (yes/no)	Tick bite % (presence – yes/no) Risk	Yes: E: 73.2% C: 26.6% No: E: 26.8% C: 73.4% p=0.0001 Adj OR= 6.74, 95% CI: 5.29–8.60		Children attending forest kindergartens reported a significantly higher prevalence of tick bites compared to the traditional kindergartens. Attending a forest kindergarten was a risk factor for having at least one tick bite when adjusting for age, sex, skin inspection and recommended vaccination.	▼	Weak
			Borreliosis % (presence – yes/no)	Yes: E: 2.0%		As above	▼	

				C:0.4% No: E: 98.0% C: 99.6% (p= 0.004)				
			Risk	Adj OR= 4.61, 95% CI: 1.50–14.17				
Natural elements within ECE								
Söderström at al (2013), Sweden. E: 172 children / 9 ECE	Cross-sectional	Symptoms (illness) The sum of days with symptoms of illness (runny nose, cough, fever, respiratory problems/asthma, itchy skin, diarrhoea, stomach ache, ear pain, body ache, sticky eyes, any medicine taken and days where parents had worries for their child). High score = less healthy.		p= 0.12 (descriptive statistics not presented)		Outdoor environment quality was not significantly associated with symptoms	N/A	Weak

Abbreviations: E= experimental; C= control; n= number; ECE = early childhood education (includes preschools, day care, kindergarten etc.); SD= standard deviation; SE= standard error; CI= confidence intervals.

Effect direction explained:
▲: positive association
►: no change/ conflicting findings
▼: negative health association
▲: positive association and statistical significance ($p<0.05$)
▼: negative association and statistical significance ($p<0.05$)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Qualitative

Table 7. Findings from eligible qualitative studies

Theme	Sub-theme	Studies	Quotes derived from participants or authors conclusions
Natural settings provide more affordances compared to traditional settings	Natural settings **enable** children to **diversify** their **play** (inc. risky, free and active play)	Dowdell et al (2011) Herrington & Studtmann (1998) Liu (2020) Puhakka et al (2019) Sandseter (2009) Wishart et al (2019)	*"You ask them what they are playing, they are going to tell you something different every day.", "we have some tree stems. they are in different sizes. they will line them up according size, they are smart kids. They line them up according to size and walk across, kind of like a balancing beam and they jump off them." (Liu, 2020).*
	Natural settings **enable** children to engage in **high intensity physical activity**	Bjørgen (2016) Puhakka et al (2019).	*"High physical-motor levels are created, the children jump down and run back up. They talk, shout and laugh. Three of the girls jump together and try to land in differing ways. They hold hands and try to jump together from the small knoll. There is laughter. They are eager and enduring. The small knoll has many opportunities for variation, in height and width, which invite challenges suitable for each child's resources. The children have visual, verbal and physical contact with each other. The top of the knoll provides an overview. Some find it scary the first time they try, but together they challenge each other, supporting and encouraging each other. The children decide how much they will participate and how they jump, and how they wish to solve the challenges offered by the knoll" (Bjørgen, 2016).*

	Natural settings **afford** children with **higher levels of risk** compared to traditional settings	Sandseter (2009) Streelasky (2019)	*I like playing in the fallen logs and trees on the playground; it is so much fun, but a bit scary too! I like the big pile of sticks and logs that we made – it is for another fort that is going to be really high off the ground." (Streelasky, 2019)*
	Natural settings **enable peers and teachers to interact differently**	Bjørgen (2016) Dowdell et al (2011) Liu (2020) Streelasky (2019).	*"The children are shouting 'X... can't you catch us? Please catch us, try to catch us ...'. The staffs join the situation and run after the children. The children are shouting 'Catch me ... can't catch me' ... There is excitement and the staff are running after the children, catching them and holding them before releasing them. The staffs have high energy, the children focus on the adults, avoiding being caught. The adults show empathy, holding and hugging the child when it is caught. The game is exciting and creates enthusiasm. A high level of physical activity is created, by climbing up, sliding down, running around and hiding in the tower to escape capture by the adults. They run at high speed and the children's body language shows that they are very much engaged in the game" (Bjørgen, 2016)* "*I try to do different things with them every day. Like I said, we play with them at least then minutes. So, I try to run, parachute, the blocks, climbing, sliding down the slides." (Liu, 2020).*
Natural and traditional settings provide similar affordances	**Movement types and intensity** similar across natural and traditional playgrounds	Wishart et al (2019).	Not available.
	Opportunity for and frequency of risky play is similar in both natural and traditional settings	Sandseter (2009)	*"Comparing the two play environments, they both seem to include an extensive amount of affordances for risky play. At both preschool playgrounds, there are opportunities for play in great heights such as climbing, jumping down, and balancing and as well as opportunities for play with high speed such as swinging, sliding/sledding, running, and bicycling."* Taken from authors conclusions – (Sandseter, 2009)

Natural environment is more diverse and engaging, and preferred by children for play compared to traditional settings		Bjørgen (2016) Streelasky (2019)	*"I like going outside and playing! I like playing with my friends, Sydney and Megan. We play hide and seek on the playground and hide in the forest in the logs and trees. I like outside because it's so fun and I really like to play. Sometimes I play with my sister too; I like all the colours outside and all the space." (Streelasky, 2019).*
Restorative and invigorating effect of nature		Puhakka et al (2019),	*"Now it's become very difficult to finish playing. They would rather continue, and those who need to take a nap, they've had a nice, long time outdoors and nice games so they fall asleep more easily, and it affects their energy in the afternoon. Some children have very long days here. They come in the morning and stay until five o'clock; they seem to be somehow energetic and lively in the yard. This is new for us. The contrast to the previous yard is so great that the effects can be seen here very quickly." (Puhakka et al, 2019).*

Supplementary file 6. Synthesis of qualitative and quantitative findings

Themes from qualitative studies	Quantitative results		
	Nature-based ECE	**ECE Natural playgrounds**	**Natural elements within ECE**
Natural settings enable children to **diversify their play**	Children engaged in a range of physical activity types; Better balance and locomotor skills (running, skipping, hopping), strength and jumping.	-	-
	Lower speed and agility	-	-
Natural settings enable children to engage in **high intensity physical activity**	Conflicting findings in MVPA doesn't suggest higher PA intensity	Decrease or no improvement in MVPA	Decreased MVPA as number of hilly landscapes, natural surfaces and vegetation increased.
Natural settings afford children with **higher levels of risk** compared to traditional settings	More minor injuries	-	-
Movement intensity similar across natural and traditional spaces	Conflicting findings in MVPA doesn't suggest similarity between spaces	-	No difference in MVPA
Movement types similar across natural and traditional spaces	No evidence of different object control skills and total motor competence which comprised of manual dexterity, ball skills and balance and total fitness scores	-	-
Restorative and invigorating effect of nature (playground)	Lower daytime sleepiness	-	Better sleep outcomes

Abbreviations: ECE = early childhood education; MVPA = moderate-to-vigorous physical activity; PA = physical activity

Green = confirmatory (i.e. quantitative results confirm qualitative themes); orange = conflicting (i.e. quantitative results conflict qualitative themes).