The Association Between Outdoor Free Play and Vitamin D Serum Levels: A Multiple Linear Regression Model Analysis

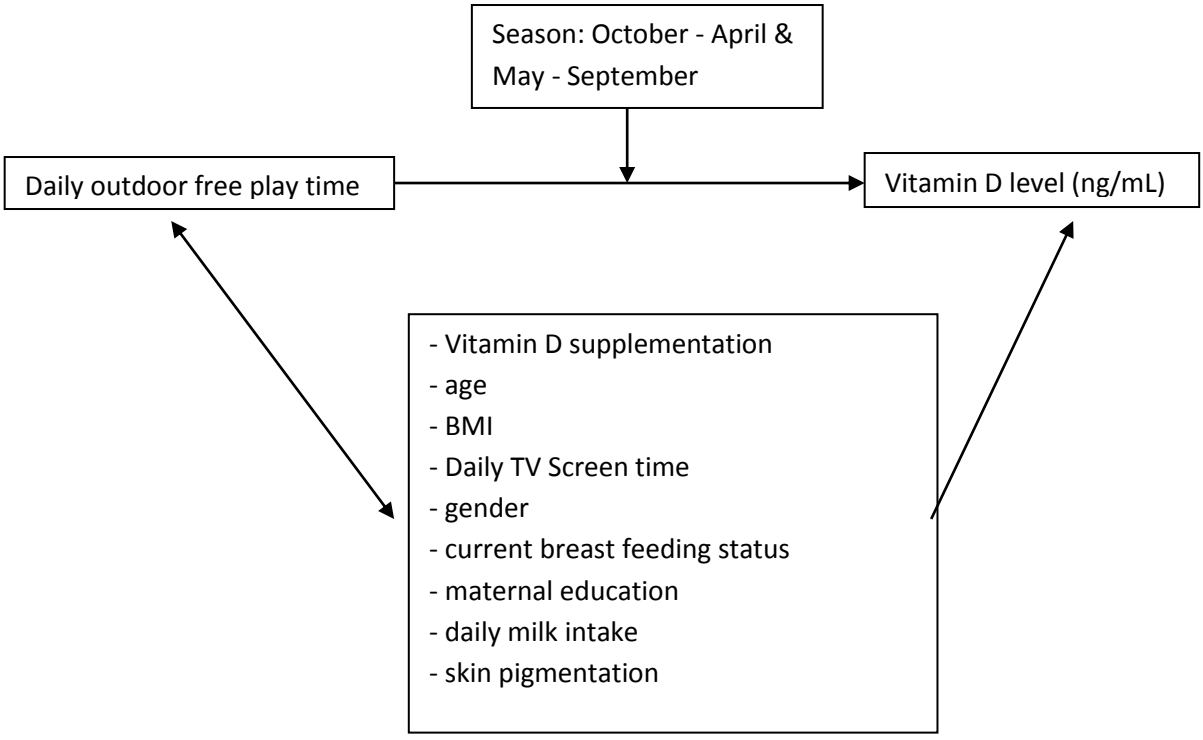
Nicolas Theodoric

Modelling Project as Requirement of CHL5021H

December 13, 2013

Introduction

Vitamin D is crucial for the normal development and maintenance of bones in children[1]. Deficiency in this vitamin can lead to manifestation of debilitating diseases such as rickets and arthritis[2]. The main source of vitamin D comes from exposure of UV light from the sun on the skin[1,3,4]. The daily recommended dose of vitamin D in Canada is 600 IU and 400 IU in the US[5]. Interestingly, exposure to sunlight for a over 8 minutes is enough to synthesize this daily recommended dose[6]. Therefore, it is crucial that children spend at least 10-15 minutes outdoor daily to achieve healthy daily intake of vitamin D. Of course, in order for time spent outdoors to be effective in obtaining vitamin D, direct sunlight exposure on the skin must occur. As such, this may be a problem for children living in Ontario where sun exposure vary greatly depending on the season as this may increase the risk of vitamin D deficiency for children that already do not spend much time outdoors in general. In light of the current knowledge of the factors that determine children's vitamin D levels, I would like to test the association between vitamin D levels and time spent for outdoor free play within children using a multiple linear regression model. Inherent to testing the association between vitamin D levels and outdoor free play, season at which vitamin D readings are taken should affect the amount of vitamin D levels elevated due to outdoor free play. As such, we first hypothesize that there is a positive linear association between vitamin D levels and time spent during outdoor free play. Second, we hypothesize that the effect of outdoor free play in increasing vitamin D levels depends on the season. Since we are using a multiple linear regression model to test the relationship between vitamin D and time spent outdoors, our study can test the relationship in a clinically meaningful way as the model takes into account other factors that may influence vitamin D levels such as vitamin D supplementation, skin pigmentation, as well as body mass index. To do this we obtained data of 1392 children from an Ontario pediatric registry who have had serum vitamin D levels measured. We fit our data using multiple linear regression with outdoor freeplay time as the main effect exposure and season at which data is taken as the effect modifier. We include predictors and other potential confounders as well into the full model:



Methods

Measures

Data for this study were obtained from an Ontario pediatric registry. Primary outcome vitamin D levels (ng/mL) and main exposure variable outdoor free play (min) were extracted from a sample of 1392 children ranging from 11 to 119 months old. The season at which vitamin D readings were taken is treated as an effect modifier variable with two categorical levels: October - April and May - September. The following variables were also obtained from the subjects and treated as confounders: daily intake of vitamin D supplement, age, BMI, daily TV screen time, gender, current breast feeding status, maternal education, daily milk intake, and skin color. These confounding variables may be predictors, but they are pooled into the confounding category as we will test their significance in affecting vitamin D as confounders. It's also important to note that skin pigmentation may be an effect modifier as well, but this variable will be treated as confounders within this study as our focus is on season as an effect modifier for the relationship between outdoor free play and vitamin D. Aforementioned confounder variables were chosen based on evidence within literature that suggests association with vitamin D levels.

Data Analyses

We used a multivariable linear regression model to identify the relationship between daily outdoor free play time and vitamin D levels. The mathematical model is as follows:

Primary outcome vitamin D level is denoted as Y, intercept as B0, ith independent fixed variable related to Y as Xi, and Bi as regression coefficients corresponding to the ith independent variable in the model. Interaction of season with outdoor free play on vitamin D levels was tested via crude linear regression using a p-value of smaller than 0.3 as a criterion for inclusion in the final multivariable model. Confounders were tested for significance using a full approach method. In this method, the regression coefficient for outdoor free play time in the full multivariable model were compared to the coefficient when the confounding variable is removed from the model. A change in effect (regression coefficient) of at least 10% was used as a criterion for significant confounder.

To test whether our multiple regression model follow the assumption of normality, we used quantile-quantile plots of our full model residuals with theoretical quantiles of a standard normal distribution and checked for any substantial departure from normality. In line with the assumption of constant variance across vitamin D levels at a given fixed outdoor free play time, we checked for homoscedasticity by plotting full model residuals with fitted vitamin D levels as well as with main exposure outdoor free play time values. We then looked for any discernible patterns characteristic of a heteroscedastic model . Furthermore, we test the assumption of the linearity of our model by inspecting for linear trends on plots of full model residuals versus outdoor freeplay and confounding variables. We also test the independence between independent variable by analyzing for multicolinearity. We do this by tabulating the variance inflation factors (VIF) of each independent variable. A VIF of greater than 10 indicates multicolinearity.

As a final diagnostic, we identify extreme and influential observations via dfbeta assessment of vitamin D level in each subject of our data.

Results

Children Subjects

Analyses included data from all 1392 children samples. From these subjects, we obtained data on vitamin D level as the primary outcome, outdoor free play time as main exposure, and ten other variables (predictors and confounders) that are treated as confounders (Table 1).

Gender distribution of the subjects is 49% female and 51% male. Subjects spent a median (IQR) of 60.00 (45.00) minutes for daily outdoor free play. Vitamin D levels were obtained during the May-September season for nearly half of the subjects (49%), while readings for the other half of the subjects were obtained during October-April(51%). 58%% of the subjects took daily vitamin D supplementation and 42% didn't. Median (IQR) age, BMI z-score, daily TV screen time, and daily milk intake are respectively 36.99 (30.07) months, 0.20 (1.31), 60.00 (77.14) minutes, and 500 (500) mL. 90% of subjects are not breastfeeding while 10% are. Most subjects' maternal parent had a college/university education (91)%, while a smaller portion had a highschool (8%) or public school (1%) education. Most subjects had 'light' skin pigmentation (86%) and a lesser portion had 'dark' pigmentation (14%).

Multiple Regression Analysis

Univariable analysis of vitamin D level with outdoor free play reveals no significant relationship (β±S.E = 0.02±0.01, 95% C.I = -0.01, 0.05). Crude association of vitamin D levels and season at which readings were taken revealed, with significance, that vitamin D levels are -4.07 ng/mL (95% C.I. = -7.54, -0.97) lower during the months of October - April than May - September. Testing for interaction between season and free play on vitamin D levels revealed significant interaction between season and outdoor free play on vitamin D (β = -0.035, p = 0.027). After adjusting for effect modification and predictors/confounders, we find that there is still no significant relationship between vitamin D levels and outdoor free play time (β±S.E = 0.03±0.02, 95% C.I. = -0.01, 0.06) (Table 2). Furthermore, interaction of outdoor free play with season on vitamin D levels is no longer significant ((β±S.E = -0.03±0.03, 95% C.I. = -0.09, 0.03). Testing of confounders using the full approach method did not reveal any significant confounders (change in main exposure coefficient are all less than 10%). Although none of the confounders significantly confound the relationship between outdoor free play and vitamin D levels, they are still included in the model as they are regarded as important in the literature.

Despite insignificant confounding, a number of these proposed confounders significantly affect vitamin D level when all other variables are adjusted for. Daily vitamin D supplementation increases vitamin D reading levels by 10.41 (95% C.I. = 7.10, 13.7) ng/mL (Table 2). Every unit increase of BMI z-score decreases vitamin D levels by 1.98 (95% C.I. = 0.41, 3.55) ng/mL. Surprisingly, one unit increase in daily TV screen time increases vitamin D levels by a negligible yet significant 0.02 (95% C.I. = 0.0002, 0.04) ng/mL. Children with a maternal education of public schooling has vitamin D levels 19.53 (95% C.I. = 2.09, 36.97) ng/mL higher than children with maternal education at the college/university level. Interestingly as well, each unit increase of daily milk intake (mL) only increases vitamin D levels by 0.02 (95% C.I. = 0.01, 0.02) ng/mL. Vitamin D levels are 5.34 (95% C.I. = 0.66, 10.03) ng/mL lower in children with 'dark' skin pigmentation compared to 'light' skin pigmentation.

To verify the validity of our model, we tested the assumptions that follows a multiple linear regression model: residuals are normally distributed, homoscedastic variance of residuals across vitamin D level observations, linearity between exposures and outcomes, independence of exposure variables. Q-Q plot

analysis by plotting quantiles of residuals versus theoretical quantiles of a standard normal distribution showed fitting of residuals to a normal distribution. However, departure from normality (departure from Q-Q line) is observed for residuals accounting for extremely high values of vitamin D. Plotting of residuals with fitted vitamin D values, outdoor free play, and other independent variables in the full model showed no trend in residual spread indicating heteroscedasticity. This suggests that our model is fairly homoscedastic. Inspection of plotting residuals versus main exposure variable and other continuous independent variables in our full model confirmed the linearity of our model. We also fitted a lowess smooth line to aid in this inspection, which again showed no distinct departure from linearity. Next, we tested independence between independent variables through multicolinearity analysis. Tabulation of VIFs for each exposure variables showed no VIF values over 10 and therefore no evidence of multicolinearity (Table 2). Finally, dfbeta plots showed 0.64% influential observations. These influential observations consists of subjects with extreme vitamin D levels (e.g 352 ng/mL) and outdoor free play time (e.g 400 mins/day) both falling within greater than 95th percentile ranges.

Discussion

In this study we have sought to apply a multiple linear regression model to test the hypothesis that time spent on outdoor free play is linearly associated with vitamin D levels. We found that despite the current understanding that exposure to sunlight is achieved through time spent outdoors, which leads to vitamin D synthesis through skin, our analysis revealed no significant relationship between outdoor free play and vitamin D levels. Furthermore, after adjusting for other variables, season at which vitamin D levels were taken had no effect on the outdoor free play and vitamin D relationship. In stark contrast, crude analysis between season and vitamin D levels showed that vitamin D levels are lower during October-April months where sunlight is more limited. This is surprising, given that amount of UV irradiation is much higher during May-September months. One reason that may explain for this observation is a ceiling effect in vitamin D serum levels after exposure to sunlight for a certain amount of time. Holick (1995) showed that after 20 minutes of sun exposure, vitamin D serum levels reach an equilibrium and any more synthesized vitamin D is degraded[4]. Considering that median time spent outdoor in children is 60 minutes with a maximum of 400 minutes, we may be attenuating the effect of time spent during outdoor free play with increased vitamin D levels. Therefore, future experiments to test a relationship between time spent outdoor and vitamin D should consider limiting the range of outdoor free play time.

In line with applying a multiple linear regression model, our model satisfy the basic required assumptions of the model: independence between proposed independent variables, linearity between vitamin D and independent variables, homoscedasticity of vitamin D outcomes across all values of independent variable, and normal distribution of model residuals. Influential observations were observed and there is a slight departure from normality towards subjects with extreme characteristics such as extreme serum levels of vitamin D and excessive time spent during outdoor free play. Exclusion of subjects with such characteristics are recommended for future experiments. Finally, it should be noted again that some covariables in this study are predictors (e.g vitamin D supplementation) and potential effect modifiers (skin pigmentation). These variables should be treated as their corresponding entities for future experiments if we are to build a more accurate regression model.

Table 1. Characteristics of children subjects.

Variables	Median (IQR)	Size (%)
Main Exposure		
Outdoor free play time (min)	60.00 (45.00)	
Effect Modifier		
Season of measurement:		
May - September		684 (49)
October - April		708 (51)
Predictors/additional confounders		
Daily vitamin D supplementation		
No		588 (42)
Yes		804 (58)
Age (months)	36.99 (30.07)	
BMI z-score	0.20 (1.31)	
Daily TV screen time (min)	60.00 (77.14)	
Gender		
Female		681 (49)
Male		711 (51)
Currently breast feeding		
No		1248 (90)
Yes		144 (10)
Maternal Education		
College/university		1273 (91)
High school		107 (8)
Public school		12 (1)
Daily milk intake (mL)	500 (500)	
Skin		
Light		1191 (86)
Dark		201 (14)

Table 2. Multiple Regression Analysis of Vitamin D.

Variables	β (S.E)	95% CI	VIF
Intercept	78.65 (3.21)	(72.34, 84.95)	
Main Exposure			
Outdoor free play time (min)	0.03 (0.02)	(-0.01, 0.06)	1.36
Effect Modifier			
Season: Oct-Apr	-4.07(2.46)	(-8.88, 0.75)	1.51
Outdoor free play time * season: Oct-Apr	-0.03 (0.03)	(-0.09, 0.03)	1.60
Predictors/additional confounders			
Daily vitamin D supplementation: Yes	10.41 (1.68)	(7.10, 13.7)	1.03
Age (months)	-0.08 (0.05)	(-0.18, 0.01)	1.10
BMI z-score	-1.98 (0.80)	(-3.55, -0.41)	1.02
Daily TV screen time (min)	0.02 (0.01)	(0.00, 0.04)	1.04
Gender: Male	-0.98 (1.63)	(-4.18, 2.22)	1.01
Currently breast feeding: Yes	-4.71 (2.93)	(-10.46, 1.04)	1.10
Parent education: highschool	-5.65 (3.10)	(-11.74, 0.43)	1.02
Parent education: public school	19.53 (8.89)	(2.09, 36.97)	1.03
Daily milk intake (mL)	0.02 (0.003)	(0.01, 0.02)	1.04
Skin: dark	-5.34 (2.39)	(-10.03, -0.66)	1.64
$R^2 = 0.069$			

References

1. Holick, M. F. Sunlight and vitamin D for bone health and prevention of autoimmune diseases, cancers, and cardiovascular disease. *Am. J. Clin. Nutr.* **80,** 1678S–88S (2004).

2. Wagner, C. L. & Greer, F. R. Prevention of rickets and vitamin D deficiency in infants, children, and adolescents. *Pediatrics* **122,** 1142–52 (2008).

3. Holick, M. F. Resurrection of vitamin D deficiency and rickets. *J. Clin. Invest.* **116,** 2062–2072 (2006).

4. Holick, F. Environmental factors that influence production of vitamin D13 the cutaneous. *Am. J. Clin. Nutr.* **61 (suppl.,** 638S–45S (1995).

5. Gartner, L. M., Lawrence, R. A., Naylor, A. A. & O'Hare, D. Prevention of rickets and vitamin D deficiency: new guidelines for vitamin D intake. *Pediatrics* **111,** 908–910 (2003).

6. Terushkin, V. *et al.* Estimated equivalency of vitamin D production from natural sun exposure versus oral vitamin D supplementation across seasons at two US latitudes. *J. Am. Acad. Dermatol.* **62,** 929.e1–9 (2010).