THE CAMPBELL COLLABORATION

Systematic Review Protocol:
Later School Start Times for Supporting the Education, Health and Well-being of High School Students

Colleen M. Davison, Lisa Newton, Robert S. Brown, John Freeman, Lee-Anne Ufholz, J. David Smith

Submitted to the Coordinating Group of:

- ☐ Crime and Justice
- ☒ Education
  - ☐ Disability
- ☐ International Development
- ☐ Social Welfare
- ☐ Other:

Plans to co-register:

- ☐ No
- ☒ Yes   ☒ Cochrane   ☐ Other
- ☐ Maybe

Co-registration: This review is co-registered with both the Campbell and Cochrane Collaborations. A version of this protocol can be found in the Cochrane Library.

Date Submitted: 20 April 2011

Date 1st Revision Submitted: 20 July 2011

Date 2nd Revision Submitted: 20 September 2011

Approval Date: 1 October 2011

Publication Date: March 2012

BACKGROUND

The Problem, Condition or Issue

School start times for adolescents are increasingly being scrutinized within school communities. Recently, school officials have begun varying start times as a means to improve students' productivity, achievement, and health and well-being. This informal experimentation has proceeded in the absence of clear, detailed and well-substantiated guidelines about school start times that are based on empirical evidence. The purpose of this review is to synthesize existing evidence about school start times and contribute to knowledge that can ultimately support sound policy and decision-making. School time scheduling has implications globally and we also aim to explore whether or not this type of intervention works better in certain geographic areas, in certain kinds of schools or with certain types of students. The current and growing interest in secondary schools adopting later start times has been preceded by approximately 20 years of research concerning the issues of adolescent sleep needs, sleep deprivation, adolescents' peak periods of alertness and best time of day for learning.

It is our hope that this review will contribute towards the development of systematic review methodologies, particularly the approaches required for conducting reviews in public and population health and education, as well as for reviews that have a specific focus on the equity of these kinds of interventions.

We have not thus far been able to identify any existing systematic reviews on the effectiveness of later school start times. We have, however, identified some non-systematic reviews and summaries as well as systematic reviews on related topics. The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) at the Social Science Research Unit, Institute of Education, University of London recently completed a review on the effects of 'block scheduling' (one form of class scheduling) on students. As part of this review, a map of all interventions that were found relating to school time use was constructed. The map included, for example, interventions related to the length of the school day, the length of the school year, the length of classes and the use of Saturdays for schooling. Later school start time was identified as a type of intervention related to school time use. This topic was not considered for synthesis or in-depth review by the EPPI-Centre reviewers, however. An undergraduate thesis project was recently completed on the effectiveness of later school start times in the United States (Kuhfeld 2009). Non-systematic literature reviews related to later school start times have also been conducted (Yan 2006; Yan 2007) as part of a Blue Valley School District evaluation in Kentucky, USA and Newton 2010 with the Toronto District School Board, Toronto, Canada (also a co-author on this review). None of these previous studies or reviews has provided a comprehensive, systematic picture of the effectiveness of

later school start times for improving the education, health or well-being of high school students because:

1. the scope of the reviews was limited and evidence presented in these reviews may be only a partial representation of relevant studies;
2. systematic inclusion or exclusion criteria were not used; and
3. included studies were not systematically or critically appraised, so we know little about the quality of the evidence.

Although the topic has global significance, the majority of studies found to date were undertaken in North American settings, with school times being adjusted from early starts (between 7:00 and 8:00 am) to times after 8:15 am. There appears to be a preliminary indication from these previous non-systematic reviews and summaries that later school start times have positively impacted student learning by way of students being more alert, better focused on what they were doing, attending more first period classes and falling asleep less often in class. There has been some indication (especially in Wahlstrom 2002) that students in later starting schools eat breakfast more often, get more sleep and are less tense at home. School nurses reported fewer stress-related complaints and illnesses. Common disadvantages appear to be a longer school day for teachers; later times for sports and after-school activities resulting in students getting home later; and a decline in work hours for students with after-school jobs.

There is ample research available on the biology of sleep, specifically pertaining to adolescents. This body of research details adolescents' unique circadian rhythms, patterns of later melatonin release, bio-regulatory processes (delayed phase preference) associated with later evening sleepiness, adequate sleep needs (9.25 hours per night) and the growing prevalence of sleep lack or deprivation. The average total hours of adolescent sleep in the United States, for example, is reported as 6.5 to 7.5 hours per night (Carskadon 1990; Dahl 1996). If given a choice, as children progress into adolescence, they tend to go to sleep later and wake up later (Acebo 1991) but that this can change again in the late teens. Sleep research studies commonly find that teenagers have biologically different sleep and wake patterns than the preadolescent or adult populations (Wahlstrom 2002).

The initial research indicates that a number of conditions or behaviors may be remedied by a later school start time. In a previous study, students with a later school start time of 9:30 am averaged 7.5 hours of sleep, whereas those with an earlier start time of 7:20 am reported only 6.9 hours of sleep on week nights (Kowalski 1995). Even a moderate sleep increase (30 minutes to one hour) over a period of time has shown to be associated with increased scores on various performance tasks (e.g. reaction times and memory tests) (Sadeh 2003).

The collection of prominent sleep research studies, from the late 1980s and throughout the 1990s, have been critiqued (Wolfson 2003). Their summative report revealed that the association of sleep patterns and sleep quality, specifically

shortened sleep times (later bed times paired with earlier wake times), were associated with lower academic performance, grade point averages and motivation levels. They also reported that more sleep and later week day rise times were associated with better grades and a higher motivation to do well in school.

In the review of research that focused on laboratory testing for the impacts of sleep disturbances, sleep deprivation and adolescents' bio-regulatory processes, the following concerns were uncovered:

- aggressive behavior (Gibson 2006);
- memory loss (Poirel 1987);
- conflict with teachers, increased irritability and aggressive behavior (Maas 1995);
- tardiness (Gibson 2006; Wahlstrom 2002);
- falling asleep in class (Maas 1995);
- poor self esteem (Fredriksen 2004);
- greater risk of accidents (Dahl 1996);
- anxiety and depression (Chorney 2007);
- pre-frontal cortex dysfunction (inappropriate behavioral responses, negative impact on creative thinking, decreased goal-oriented behavior) (Drummond 1999; Harrison 1996);
- slow performance or lapses of performance, slower reaction times, difficulty concentrating (Sadeh 2003); and
- low achievement and achievement motivation (Epstein 1998; Meijer 2000).

It appears, therefore, that later school start times could logically have some positive education, health and well-being impacts for high school students and their families. It is the objective of this review to examine the evidence base for these potential relationships.

The Intervention

This review will be conducted specifically to examine the effectiveness and implementation of later start times in high schools. The intervention is an adjustment in the time school starts and, thus, the time school takes place during the day. We are interested in comparisons of schools that start at different times during the day, one later than the other, as well as comparisons between morning and afternoon or evening 'shifts' at a single school where double or multiple groups or split shifts of students exist.

How the Intervention Might Work

There are two main theories about why a later high school start time might improve the education, health or well-being of high school students. The first relates to the amount of sleep an adolescent gets depending on when they have to wake up to get to school and the various impacts of different amounts of sleep. The second theory relates to there being an optimal time of day for adolescent alertness and learning,

and that these factors are associated via many routes with health and educational outcomes.

Previous studies indicate that adolescents may not be behaviorally and physiologically ready to fall asleep until approximately 11:00 pm and that some show physiological signs of becoming more energetic and internally stimulated in the late evening (Carskadon 1981; Wolfson 2005). If given a choice, adolescents prefer going to bed later and waking later. There is general consensus throughout the majority of the biomedical and sleep literature that a mismatch exists between school schedules and demands, on the one hand, and the adolescent delayed sleep cycle on the other.

The second theory about why later school start time interventions might work to support the education, health and well-being of high school students is that implementing a later school start time may better align the school day with peak times for adolescent alertness and potential for learning, and that these factors in turn have direct implications for health and educational outcomes. Evidence for this theory relates to investigations of adolescent neuropsychologic performance, alertness and the time of day ideal for learning and short- and long-term memory tasks. Hansen 2005 found that students' performance was better later in the day than in the early morning. Adolescent short-term memory has also been shown to be better in the morning but long-term memory improved in late afternoon and early evening (Harrison 1996; Monk 1987).

Why it is Important to do the Review

There is a growing awareness amongst parents, teachers, schools and community groups of the possibilities of later school start times to support better health and education outcomes for adolescents. There is a growing body of literature on the effects of even mild sleep deprivation on young people and on the possible peak times for learning during the day in high schools. At the request of parents, teachers, school council members and others, school administrators and educational policy-makers in many jurisdictions are currently considering later school start times. A number of schools and school districts have changed times over the past 10 to 15 years, particularly in the North America. There have been a number of single studies of these kinds of interventions, but it is unclear whether there is sufficient evidence about the effectiveness of these kinds of approaches for students, their families or the school's wider community (Taras 2005). The results of this review can be used by school officials in making decisions about whether or not to adopt a later school start time within their own schools.

OBJECTIVES

The primary objective of this study is to assess the effectiveness of later high school start times for improving student educational, health and well-being outcomes. Specific objectives are the following.

1. In considering equity, to explore possible differential effects of later school start times among subgroups of students by (where possible to ascertain) age, grade, gender, ethnicity and socioeconomic status.
2. In considering equity, to explore possible differential effects of later school start times among subtypes of schools, such as comparing larger versus smaller schools, rural or urban schools, schools with students who walk to school versus those who take the bus or have mixed transportation, public or private schools, or schools in different geographic locations and climatic conditions.
3. To identify and explore any adverse or unforeseen effects of later school start times.
4. To explore the process of implementing later school start times to identify successful features as well as aspects of context and delivery that might lead to positive or adverse effects.
5. To identify and explore the effects of later school start times on the broader school community (including teachers and other school staff, the school neighborhood and the students' families).

METHODOLOGY

Inclusion Criteria

What types of studies designs are to be included and excluded?

We will include the following types of studies in this review:
- randomized controlled trials (RCTs) (including cluster-randomized controlled trials or randomized cross-over trials);
- non-randomised designs including:
- non-randomized controlled trials of an intervention (including potentially cluster and/or cross-over trials) (CCTs);
- controlled before and after studies (CBAs) of an intervention;
- interrupted time series studies (ITSs) with at least three pre and post measurements.

For the purpose of this review, the term 'intervention' will also include changes that are deemed 'natural' or 'natural experiments' (changes that occur and can be studied but are not under the design or immediate control of the research group). These could include, for example, comparisons of different 'shifts' of students when a high school enters a split-shift configuration, or studies of groups of students who leave a school with an earlier start time and enter a new school with a later one.

Note 1: we will only use qualitative data that are associated with an intervention study *included* in this review. These studies may help contextualize and provide explanations for the major findings. We will assess the qualitative studies for quality (rigor) by reviewing for evidence of:

- a clear and appropriate description of methods (for sampling, data collection and analysis);
- triangulation (of methods, data sources and/or of researchers);
- an audit trail.

We will not use qualitative studies that do not show evidence of these three points. For those that will be included, we will analyze them thematically (coding for major and minor themes presented) and include them in the review's narrative summary.

Note 2: we understand that there are likely other types of studies of late start interventions (for example, case-control studies, educational ethnographies, case studies or other types of observational studies). Based on our study type criteria for inclusion, these kinds of studies will not be formally included in our review. However, we will make note of those we find and include comments in the review about the different methodologies used. Where appropriate, we will offer advice on designing ideal intervention studies for evaluators of these kinds of schedule changes.

Types of participants

We hope to capture the student age group of 13 to 19 years, for whom this type of intervention is most pertinent. Participants will be high school or secondary school (or in some cases middle school or junior high school) students and their families, schools and communities, worldwide. We will not exclude students with special needs, alternative schools or boarding schools. We will note these differing circumstances in the narrative summary.

Types of interventions

This review will include studies of interventions that involve a comparison between two or more different high school start times. These can include 'late start' interventions, in which high school start time is purposely moved later and then studied in comparison to the previous or another earlier start time. It can also include studies that compare outcomes associated with start times at different, but matched, schools or different 'shifts' of students within the same school (the morning, afternoon or evening shift). We will also include studies of outcomes associated with day versus night school attendance. We will not exclude studies based on type of school (public, private, boarding etc.).

Types of outcome measures

We have divided the many possible outcomes of interest into two categories. We will include findings related to primary outcomes in the 'Summary of findings' table as well as a list of any possible adverse outcomes or harms associated with the intervention. We will include overall findings associated with secondary outcomes in a narrative summary.

Primary Outcomes

Primary outcomes include the following.

- Student <u>academic outcomes</u> measured, for example, by locally relevant standardized test scores and course grades, midterm and final exam results, and graduation records.
- Outcomes related to <u>amount or quality of sleep</u> for students.
- <u>Mental health indicators</u> including, for example, measures of stress, anxiety, feelings of isolation or exclusion, depression or suicidal ideation.
- Student <u>truancy or attendance</u>.
- Teacher or self reported student <u>alertness</u>.

We will also note any additional adverse outcomes or possible harms. These might include the following.

- Increased transportation costs for students, families or schools.
- Decrease in student supervision outside school time.
- Less time spent with family.
- Issues with childcare and before- and after-school care for younger siblings.
- Decreases in enrolment in extra-curricular and athletic activities.
- Polarizing or increased conflicts within the school community through this issue.
- Difficulties for school staff (making medical appointments, attending professional development workshops etc.).
- Increased costs and difficulties obtaining substitute teachers.

Secondary Outcomes

Secondary outcomes may include the following.

- <u>Outcomes related to health behaviors</u> (e.g. diet, exercise, tobacco use, alcohol or drug use, unsafe sexual practices).
- <u>Health and safety indicators</u> (e.g. vehicular accidents, falls or other accidents, abductions, waiting for buses in the dark, transportation issues).
- <u>Social outcomes</u> (e.g. changes in social supports, peer relations, participation in extra-curricular activities, student employment, discipline referrals at school).
- <u>Family outcomes</u> (e.g. changes in the supervision of children/adolescents and childcare, family communication and dynamic, routines, roles).
- <u>School outcomes</u> (e.g. changes in registration numbers, school climate or discipline referrals).
- <u>Community outcomes</u> (e.g. reactions from local businesses, feelings of safety on neighborhood streets).

We will record any mention of financial costs and resources required for implementation of the intervention and include this in the narrative summary.

Where possible, we will collect data on the schools and communities that are the settings of the included studies to provide context for the results. Qualitative companion studies may be helpful in this regard. Of particular interest are: social, political, climatic and geographic factors that might impact the delivery of, access to and experience of schooling, as well as any other school or community-based

interventions that have impacts on students, such as nutritional or other health and development interventions. We will also note, where possible, the season or time of year in which the included studies were undertaken.

Search Strategy

We will undertake electronic searches in scientific databases as well as hand searches of citation lists, recent relevant journal index lists, websites and publications from relevant conferences and scientific meetings. We will also contact key expert informants, including previous authors and researchers in this area.

Electronic searches

We gained input from three reference librarians to develop the search strategy. We will not exclude documents on the basis of language, country or publication date. The MEDLINE search will be conducted on the Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1948 to Present. This search will be modified for other databases and is included as Appendix 1.We will perform the search in the following electronic databases.

EBSCO Platform

- Academic Search Complete*
- Global Health (CAB Health)*
- CINAHL*

OVID Platform

- EMBASE*
- MEDLINE*
- PsycINFO*

GALE Platform

- Educator's Reference Complete*
- Educators e-collection*

SCHOLARS PORTAL Platform

- ERIC*
- Education Abstracts*
- Education Administration Abstracts*
- Sociological Abstracts*
- Social Science Citation Index*

WILSON WEB Platform

- Education Full Text*

PROQUEST Platform

- Dissertations and Theses

<u>DIALOG Platform</u>

- British Educational Index
- Australian Educational Index

<u>INDEPENDENT DATABASES</u>

- JSTOR*
- The Cochrane Collaboration Database of Systematic Reviews
- Cochrane Public Health Group Specialized Register
- Cochrane Central Register of Controlled Trials (CENTRAL)
- The Campbell Collaboration Database of Systematic Reviews (C2-SPECTR, C2-PROT)
- Evidence for Policy and Practice Information and Co-ordinating Centre Review Databases (including BiblioMap, DoPHER, TRoPHI, Database of Education Research, and CERUK)

A research assistant will complete 15 electronic database searches (*); CMD will complete the remainder. We will divide the search concepts roughly into three groups: population (lines 1 to 8 in the sample search); setting (lines 9 to 16) and intervention (lines 17 to 39). We will consider all outcomes and comparators.

Searching other resources

We will hand search the bibliographies of all included studies (as well as excluded but closely related studies). We will examine the 'similar studies' and 'cited in' options for all included studies where possible in Web of Science. We will contact authors of included studies and the authors of the other related non-systematic reviews by email. We will contact other prominent school-health researchers and members of our advisory group requesting any further relevant information, unpublished or in-progress studies.

In addition, we will hand search the most recent issues of the following journals for publications that would not yet be included in the electronic databases:

- *Sleep* (2010 to present)
- *Behavioural Sleep Medicine* (2010 to present)
- *Journal of Adolescent Health* (2010 to present)
- *International Journal of Child and Adolescent Health* (2010 to present)
- *Journal of School Health* (2010 to present)
- *American Educational Research Journal* (2010 to present)
- *British Educational Research Journal* (2008 to present)
- *Canadian Journal of Education* (2009 to present)
- *Review of Educational Research* (2010 to present)
- *Educational Researcher* (2010 to present)
- *British Journal of Educational Studies* (2010 to present)
- *Educational Evaluation and Policy Analysis* (2010 to present)
- *The High School Journal* (2010 to present)
- *Journal of Sleep Research* (2009 to present)

We will search the following English language websites, conference listings, popular media sources and portals for relevant grey literature (such as evaluation reports and policy papers):

- System for Information on Grey Literature in Europe (SIGLE 1985-2005)
- American Educational Research Association
- American Psychological Association
- International Association for Adolescent Health
- The Canadian Sleep Society
- Academy for Educational Development
- World Health Organization
- The National Association of Secondary School Principals
- Ontario College of Teachers
- The Canadian Society for the Study of Education
- The Association of Supervision and Curriculum Development
- UK Dept. of Education
- European Sleep Research Society
- American Academy of Sleep Medicine
- Sleep Research Society
- The World Federation of Sleep Research & Sleep Medicine Societies (WFSRSMS)

In order to locate additional unpublished studies not identified through the procedures described above, we will carry out an Internet search using key words and prominent author searches. We will review five pages of hits for each keyword or author search in Google.

METHODS

Selection of Studies

Two review authors (CMD and LN) will independently scan the abstracts and titles of articles retrieved by the electronic and handsearches for eligibility, according to the inclusion or exclusion criteria outlined in the types of studies, participants and interventions listed previously. We will retrieve full copies of all those deemed eligible by at least one of the review authors for closer examination. In any instance in which it is difficult to make a selection decision on the basis of the title and abstract alone, we will retrieve the full article for screening. We will exclude all studies that initially appear to meet inclusion criteria but upon this closer inspection do not. These exclusions will be listed in a table of 'Characteristics of excluded studies' and we will provide reasons for exclusion.

Data Extraction and Management

Three review authors (CMD, LN and RB) will independently extract data and will thoroughly review each other's work. The data extraction form has been developed from the Cochrane checklist (Higgins 2008, p. 157). Any discrepancies among the authors for inclusion or exclusion of studies will be resolved through discussion among these authors and involvement of the remaining authors. If consensus cannot

be reached, a final decision will be made by CMD. We will then enter the data into the Review Manager 5 software (RevMan 2011) and record study details in the tables for 'Characteristics of included studies' and 'Characteristics of excluded studies'.

We will extract data on study methods (including study design, details related to follow-up measures and all timing data (see below)), description of the intervention or natural change (including where appropriate any process details, total numbers in each group, again all timing data), and details about participants. We will also extract additional items relating to implementation of the intervention, including levels of adherence to the intervention start times by schools and students, resources required for implementation, and whether a process evaluation was conducted. We will use the PROGRESS (place, race, occupation, gender, religion, education, socioeconomic status) checklist to provide detail about the included sample and to assess whether or not outcome data have been reported by socio-demographic characteristics known to be important from an equity perspective. We will also record whether or not studies included specific strategies to address diversity or disadvantage. We will extract school and community characteristics: geographic location (country, region, rural/urban), school type (grade levels, private versus public, single gender versus mixed gender, boarding school versus day school, alternative school), student transportation options, race distribution (of student majority and mix within school), social class (of student majority and mix within school), neighborhood characteristics (of the school), state infrastructure such as local roads, buildings and other social, political, geographic and climatic variables that may impact the delivery, access to, and experience of schooling. We will also extract any information relevant to the 'history of intervention' where possible (e.g. Has the school made other previous changes? What is the norm and what is the difference between that norm and the changed time? How long has the intervention been in place?). We will also note any other interventions impacting the study samples, especially school nutrition programs.

With respect to timing issues, it will be important to record all timing details systematically and comprehensively. This includes making note of the time of day that school begins and the schedule of how the school day unfolds in intervention and comparison schools. We will record all available information about start, finish, class and break times at intervention and comparison schools, as well as the times of year or season of intervention, data collection and follow-up. Of particular interest will be examining the chosen later start time to try to understand whether there is evidence indicating an ideal time of day for schooling 13 to 19 year olds.

We will include the primary outcomes (academic, sleep, mental health, truancy/attendance and alertness) in 'Summary of findings' tables. We will also list any adverse outcomes or harms found to be associated with the intervention. We will extract any mention of financial costs or resources required for the implementation of the intervention as well as the required details for statistical analyses and critical appraisal.

Assessment of Risk of Bias in Included Studies

We predict that there will be no or very few randomized or cluster-randomized controlled trials in this area and that we will mostly be working with non-randomized, controlled before and after studies and interrupted time series studies of interventions or natural experiments. As outlined by Guyatt 2011, for RCTs we are particularly aware of limitations associated with:

- lack of allocation concealment;
- lack of blinding;
- incomplete accounts of student and outcome events;
- selective outcome reporting;
- stopping early for benefit;
- the use of unvalidated outcome measures;
- incomplete washout period or carry-over effects in cross-over trials; and
- recruitment bias.

For observational studies we are particularly aware of limitations associated with:

- failure to develop and apply appropriate eligibility criteria for study and control populations;
- under or over-matching in case-control studies;
- selection of exposed or unexposed in cohort studies from different populations;
- flawed measurement of exposure or outcome;
- differences in measurement of exposure (e.g. recall bias);
- differences in the surveillance for outcome among exposed and unexposed populations in cohort studies;
- failure to adequately control for confounding;
- inaccurate measure of prognostic factors;
- lack of appropriate adjustment in statistical analysis; and
- incomplete follow-up.

To assess risk of bias in the included studies, we will use the Cochrane Collaboration's Effective Practice and Organisation of Care (EPOC) 'Risk of bias' tools (EPOC RoB Tools). Nine standard criteria are used for all RCTs, CCTs and CBAs (Yes, No or Unclear).

1. Was the allocation sequence adequately generated?
2. Was the allocation adequately concealed?
3. Were baseline outcome measurements similar?
4. Were baseline characteristics similar?
5. Were incomplete outcome data adequately addressed? (each primary outcomes can be scored separately)
6. Was knowledge of the allocated interventions adequately prevented during the study?
7. Was the study adequately protected against contamination?
8. Was the study free from selective outcome reporting?
9. Was the study free from other risks of bias?

Seven standard criteria are used for all ITS studies (Yes, No, Unclear).
1. Was the intervention independent of other changes?
2. Was the shape of the intervention effect pre-specified?
3. Was the intervention unlikely to affect data collection?

4. Was knowledge of the allocated interventions adequately prevented during the study (each primary outcome can be assessed separately)?
5. Were incomplete outcome data adequately addressed?
6. Was the study free from selective outcome reporting?
7. Was the study free from other risks of bias?

Measures of Treatment Effect

For continuous data (in the RCTS, CCTs and CBAs), where baseline data are accessible, we will report pre-intervention and post-intervention means for both intervention and control groups and calculate the absolute change from baseline (this is the change in intervention group values minus change in control group values), along with standard deviations and 95% confidence intervals. When baseline data are not available, we will express results as the relative percentage change (difference between post-intervention values in the intervention and control groups expressed as a percentage of post-intervention values in the control group). For continuous data in ITS studies, we will calculate relative and absolute mean difference in before and after values. For any discrete variables we will present the risk ratio of the outcome in the intervention group compared to the control group. We will also calculate the risk difference (absolute risk reduction), which is the absolute difference in outcome rates between the intervention and control treatment groups. We will consult a statistician to confirm the approach to analysis and measurement of treatment effect once the included studies have been identified. This is particularly the case where there is dichotomous outcome data provided from cross-over trials. This is a similar approach to measuring treatment effect as was used in a recent systematic review of school-feeding interventions (Kristjansson 2007).

Unit of Analysis Issues

Studies which allocate by either whole schools or classroom clusters can have unit of analysis issues if between-cluster correlations are not adequately adjusted for. In cluster-randomized trials, it is not uncommon for the school or classroom to be the unit of randomization, and yet the data are analyzed at the individual student level (Whiting-O'Keefe 1984). There are other issues in cross-over trial designs, especially if cross-over trials are included in a meta-analysis and measurements from two or more intervention periods are combined as if it were a parallel-group trial design. In general, there is increased possibility for errors in analysis when there are studies with repeated observations on participants, events that re-occur, multi-intervention groups or multi-attempts at an intervention.

We will pay special attention to the unit of randomization, unit of analysis and underlying design issues for each included study, particularly identifying those studies with cross-over and clustered designs. We will assess whether or not analysis and reporting were appropriately done. JS will lead the assessment of unit of analysis issues. We will make decisions about whether or not to include studies with potential errors in analysis by consensus with the review authors, and consultation

with a statistician where necessary. We will use the following questions to guide our decision:

1. What is the unit of analysis issue?
2. Can the issue be corrected by the review authors (for example, the *Cochrane Handbook for Systematic Reviews of Interventions* gives some guidance to review authors if a study will be included in a meta-analysis and the original analysis did not adequately adjust for clustering (section 8.11.2.4 Higgins 2008)).
3. Does the unit of analysis issue impact all aspects of the results and analysis or only parts of it?
4. Does the unit of analysis issue represent an error that would warrant exclusion if all other aspects of the study indicated inclusion? Could some aspects of the study be used?

Dealing with Missing Data

We will contact study authors if methodological details or statistical data are missing. Authors will be given 30 days to respond. If further information cannot be retrieved, we will note that details are unclear on data extraction and critical appraisal forms and note this information in our published review. We will exclude articles with significant missing data from analysis.

Assessment of Heterogeneity

We will make all attempts to undertake a meta-analysis. However, we predict that because late start interventions can vary (especially around the time of start and the time difference between the late start and its comparison group) and because outcomes of interest will vary across studies, it is likely that meta-analysis will not be possible due to heterogeneity. We will only pool studies if they report the same primary outcome and are sufficiently homogeneous in design. As in the systematic review by Kristjansson 2007, to assess homogeneity we will first use sound scientific judgment to determine the degree of similarity in interventions, participants or outcomes. If a decision is not obvious, we will visually examine forest plots for outliers and between-study differences and perform a Chi^2 test for heterogeneity (where $P \leq 0.1$ will be taken to indicate significant heterogeneity).

Assessment of Reporting Biases

We recognize that reporting bias can occur via many routes (e.g. biases associated with publication and non-publication, rapid or delayed publication, multiple publication, the ease of access of publications, publications in certain languages, the citation or non-citation of research findings depending on their nature and direction and the selective reporting of outcomes) and can lead to the over-optimistic estimates of intervention effects. Our search strategy is sufficiently comprehensive so as to increase opportunities to identify all studies that meet the inclusion criteria. We will use a funnel plot (a scatter plot of the intervention effect estimates) if we have sufficient numbers of included studies. We will use appropriate techniques to assess asymmetry (guided by chapter 10 in Higgins 2008).

Meta-analysis

If studies are found to be homogeneous (with respect to population, intervention, control group, outcome measure and design) and sufficient data are available for one or more of the primary outcomes, the following will take place:

1. we will input data into Review Manager 5;
2. we will plot effect sizes with 95% confidence intervals;
3. we will estimate overall effect size using meta-analysis.

If a meta-analysis is possible, we will conduct a separate analysis for different study designs (RCTs, CCTs, CBA studies and ITS studies). A similar approach was taken by Joyce (2009). If useful and if sufficient studies are available, we will also conduct meta-regression to look at the relationship between size of effect and characteristics of the different subgroups. Characteristics of interest will include gender, age and/or grade, and indicators of socioeconomic status and ethnicity.

Summary Tables

Whether or not a meta-analysis is possible, we will present findings in tabular form. We will include tables that give basic details for included studies (numbers of students, start times, numbers of different types of outcomes measured, types of schools, etc.). We will present a 'Summary of findings' in tabular form for each primary outcome. We will also present a 'Summary of findings' for the risk of bias assessment (using the EPOC Risk of Bias Assessment Tools) in a table. In addition, we will also use and report findings from the generic framework for assessing evidence quality from the GRADE Working Group (Guyatt 2011). We will assess study limitations (no serious limitations, serious limitations, very serious limitations) and quality of evidence (high, moderate, low or very low) for single studies, the group of included studies and, where data are sufficient, for each of the primary outcome across studies. Indeed, one benefit of the GRADE framework is that it can guide review authors in the assessment of evidence quality for individual outcomes rather than pooling the assessment over a number of outcomes within a single study. This is key in studies where limitations may influence findings for one outcome but not similarly for others.

Narrative Summary

We will also use a narrative summary. The narrative summary will include the following:

- A description of the types of interventions in the included studies with any details available about their implementation (some of this may be taken from rigorous qualitative companion studies which will be analyzed thematically).
- A description of the outcomes measured in the included studies and what might be perceived as obvious omissions.
- Any context details for the included studies that are available in the quantitative or rigorous qualitative literature or made available by contacting the study authors (e.g. school geographic location, socioeconomic or ethnicity details for the school or region, transportation options, potentially relevant details about broader social, political, climatic and geographic conditions etc).

- A review of findings divided by outcome and, where possible, reported by age, grade, gender, student body by ethnicity and socioeconomic status, and the existence of other school- or community-based interventions. We will approach the authors of included studies requesting these data if they are indicated but not available in the published articles or reports.
- A description of any adverse outcomes or potential harms found.
- Perceived strengths, weaknesses and contributions of each included study with specific emphasis on how it supports or detracts from health equity.
- Any mention of financial costs associated with later start interventions (this will not be a complete cost-benefit analysis).

As cautioned in Deeks 2008, we will aim to avoid introducing bias into the narrative synthesis by reporting the results of each study judiciously and make efforts to avoid inappropriate emphasis on the findings of any one particular study.

Report on Practical Significance

In addition to reporting statistical significance, we will also discuss 'clinical' or practical significance. With respect to sleep, for example, average nightly sleep gain of as little as 30 minutes has been found to positively impact education, health and well-being outcomes in adolescents where total nightly sleep times are below ideal levels (Sadeh 2003). A study may indicate that there is a statistically significant difference of 10 minutes in sleep per school night between intervention and control groups. This may have very little practical significance; however, as the time difference is very small. For other outcomes, such as indicators of alertness, mental health or academic success, even small statistically significant differences could have real value for students and schools. We will discuss these practical implications.

Report on Equity

We are using the Cochrane Collaboration Equity Field's Equity Checklist for Systematic Review Authors - 2009 to guide us as we aim to promote and support health equity in the conduct and completion of this review. We will dedicate a section of the discussion and review findings to the potential implications for health equity as well as whether there are research needs that have become evident in the conduct of the review that are specific to promoting health equity.

Subgroup Analysis and Investigation of Heterogeneity

We will assess heterogeneity among studies first using sound judgment (examining study design, participants, details of the intervention and measured outcomes). If the decision is uncertain we will examine forest plots for outliers and areas of overlap and perform a Chi² test for heterogeneity (where $P \leq 0.1$ will be taken to indicate significant heterogeneity). If heterogeneity exists, we will examine potential sources using meta-regression to look at the relationship between size of effect and characteristics of the different population subgroups and interventions. Characteristics of interest will include gender, age and/or grade, indicators of socioeconomic status and ethnicity, and time of school start, where available.

Sensitivity Analysis

We will conduct comparative analysis to test for sensitivity of the results of the review by:

- comparing results if we include studies that may have been excluded because only the abstract could be found (where some data and results are provided in the abstract);
- comparing results if we include studies that may have been excluded due to age range of participants (for example, a study may have included primary as well as secondary school students);
- comparing results that may have been excluded due to potentially confounding co-interventions (e.g. where a school adopted a later start as well as another intervention like a breakfast program);
- determining whether results differ when studies at high risk of bias are excluded;
- for dichotomous outcome variables, comparing results if relative risk ratios, odds ratios or risk differences are used; and
- for continuous outcome variables, where several scales have been used for measurement, comparing results if standard mean differences are calculated across all the scales or for each scale independently. We will also compare results if change scores or final values are used.

REFERENCES

Acebo C, Davis SS, Herman KB, Carskadon MA. (1991). Undergraduate sleep patterns: Evidence of adaptation over time. Sleep Research 20:111.

Brown RS, Presley A, Newton L, Davison CM. (2011). Eastern Commerce Collegiate Institute Late Start: Year one interim report. Retrieved from http://www.tdsb.on.ca/wwwdocuments/about_us/external_research_appli cation/docs/ECCILateStart-Yr1InterimRpt.pdf April 2011:1-58.

Carskadon M, Demont W. (1981). Cumulative effects of sleep restriction on daytime sleepiness. Psychophysiology 18:107-13.

Carskadon M. (1990). Patterns of sleep and sleepiness in adolescents. Paediatrician;17:107-13.

Chorney D, Detweiler M, Morris T, Kuhn B. (2007). The interplay of sleep disturbance, anxiety and depression in children. Journal of Pediatric Psychology 33(4):339-48.

Dahl R. (1996). The impact of inadequate sleep on children's daytime cognitive function. Seminars in Pediatric Neurology 3(1):44-50.

Deeks JJ, Higgins JPT, Altman DG. (2009). Chapter 9. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S, editors(s). Cochrane Handbook for Systematic Reviews of Interventions. Version 5.0.0 edition. Chichester: John Wiley & Sons, February 2008.

Drummond S, Brown G. (1999). The effects of total sleep deprivation on cerebral responses to cognitive performance. Neuropsychopharmacology 1999;25(S):S68-73.

Epstein R, Chillag N, Lavie P. (1998). Starting times of school: effects on daytime functioning of fifth-grade children in Israel. Sleep 1998;21(3):250-6.

Fredriksen K, Rhodes J, Reddy R, Way N. (2004).Sleepless in Chicago: tracking the effects of adolescent sleep loss during the middle school years. Child Development;75(1):84-95.

Gibson E, Powles A, Thabane L, O'Brien S, Molnar DS, Trajanovic N, et al. (2006). Sleepiness is serious in adolescence: two surveys of 3235 Canadian students. BMC Public Health 6:116.

Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. (2011). GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). Journal of Clinical Epidemiology 64(4):407-15.

Hansen M, Janssen MS, Schiff A, Zee P, Dubocovich ML. (2005). The impact of school daily schedule on adolescent sleep. Pediatrics;115(6):1555-61.

Harrison Y, Horne J. (1996).Sleep loss impairs short and novel language tasks having a prefrontal focus. Journal of Sleep Research 7:95-100.

Higgins JPT, Green S. (2008). Cochrane Handbook for Systematic Reviews of Interventions. Chichester, UK: Wiley-Blackwell.

Joyce K, Critchley JA, Bambra C. (2009). Flexible working conditions and their effects on employee health and well-being. Cochrane Database of Systematic Reviews Issue 4. Art. No.: CD008009 DOI: 10.1002/14651858.CD008009.

Kowalski NA, Allen RP. (1995). School sleep lag is less but persists with a very late starting high school. Sleep Research 24(124):124.

Kristjansson B, Petticrew M, MacDonald B, Krasevec J, Janzen L, Greenhalgh T, et al. (2007). School feeding for improving the physical and psychosocial health of disadvantaged students. Cochrane Database of Systematic Reviews, Issue 1. Art. No.: CD004676 DOI: 10.1002/14651858.CD004676.pub2.

Kuhfeld M, Cooper H. (2009). Do later high school start times benefit or create problems for families and schools? Unpublished paper: Department of Psychology and Neuroscience, Duke University.

Maas JB. (1995). Power Sleep. New York, NY: Villard-Random House.

Meijer A, Habekothe H, Van Den Wittenboer G. (2000). Time in bed, quality of sleep and school functioning in children. Journal of Sleep Research 9(2):145-53.

Monk TH. Circadian rhythms in human performance and subjective activation. In: Halaris A, editors(s). Chronobiology and Psychiatric Disorders. New York, NY: Elsevier Science, 1987:241-62.

Newton L. (2010). Literature review: a collection of research studies on implementing later high school start times and the impacts of this school schedule change. Unpublished Report: Toronto District School Board, Toronto, Canada.

Poirel C, Larouche B. (1987). Human circadian patterns and memory process: chronopsychology of rates of forgetting, a rhythmometric study. Psychology Reports 61:3-12.

Review Manager (RevMan) [Computer program]. Version 5.1. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011.

Sadeh A, Gruber R, Raviv A. (2003). The effects of sleep restriction and extension on school-age children. Child Development 74(2):444-55.

Taras H, Potts-Datema W. (2005). Sleep and student performance at school. Journal of School Health 75(7):248.

Wahlstrom K. (2002). Changing times: findings from the first longitudinal study of later high school start times. NASSP Bulletin 86(633):3-21.

Whiting-O'Keefe QE, Henke C, Simborg DW. (1984). Choosing the correct unit of analysis in medical care experiments. Medical Care 22:1101-14.

Wolfson AR, Carskadon MA. (2003). Understanding adolescents' sleep patterns and school performance: a critical appraisal. Sleep Medical Review 7(6):491-506.

Wolfson A, Carskadon M. (2005). First bell survey: a survey of factors influencing high school start times. NASSP Bulletin 89(642):47-66.

Yan B, Slagle M. (2006). School start time and academic achievement: a literature review. Report from the Blue Valley School District, Kentucky, USA (www.bluevalleyk12.org).

Yan B, Slagle M. (2007). School start time follow-up report. Blue Valley School District, Kentucky, USA (www.bluevalleyk12.org).

MEDLINE sample search strategy

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1948 to Present>

1. Adolescent/
2. Students/
3. teen$.ab,ti.
4. adolescent$.ab,ti.
5. high school students.ab,ti.
6. "young person$".ab,ti.
7. "young people".ab,ti.
8. youth$.ab,ti.
9. Schools/
10. high school education.ab,ti.
11. secondary education.ab,ti.
12. secondary school.ab,ti.
13. "junior high school$".ab,ti.
14. middle school$.ab,ti.
15. "senior high school".ab,ti.
16. "high school$".ab,ti.
17. (start* adj3 tim*).ab,ti.
18. (late* adj3 start*).ab,ti.
19. (earl* adj3 start*).ab,ti.
20. (delay* adj3 start*).ab,ti.
21. (delay* adj3 time*).ab,ti.
22. (stagger$ adj3 start$).ab,ti.
23. (stagger$ adj3 tim$).ab,ti.
24. (split$ adj3 shift$).ab,ti.
25. (double adj3 session$).ab,ti.
26. (start$ adj2 school$).ab,ti. (415)
27. (tim$ adj2 school$).ab,ti. (704)
28. (time$ adj3 day).ab,ti. (22683)
29. (begin$ adj3 day).ab,ti. (3741)
30. (start$ adj3 day).ab,ti. (6323)
31. ((double$ or morning or afternoon or evening or different) adj3 shift$).ab,ti.
32. (tim$ adj3 learn$).ab,ti.
33. ((earl$ or late$ or school$) adj3 schedul$).ab,ti.
34. Wakefulness/
35. sleep.ab,ti.
36. chronobiology.ab,ti.
37. "human biological rhythm$".ab,ti.
38. circadian.ab,ti.
39. "phase shift".ab,ti.
40. or/1-8
41. or/9-16
42. or/17-39
43. 40 and 41 and 42

Sources of Support

Internal sources

Toronto District School Board, Canada
In-kind research assistance time provided for second author.

External sources

Canadian Institutes of Health Research, Canada
Salary support for first author.
CHSRF/CIHR Chair in Nursing Research, Canada
Salary support for the graduate student performing the electronic database searches.

DECLARATIONS OF INTEREST

The authors of this review have also recently been involved in an evaluation of a high school late start program in Toronto, Canada (Brown 2011). This may be one of the studies included in this review.

REVIEW TEAM

Lead reviewer:

Name:	**Colleen M. Davison**
Title:	CIHR Postdoctoral Fellow
Affiliation:	University of Ottawa, Institute of Population Health
Address:	451 Smyth Ave.
City, State, Province or County:	Ottawa Ontario
Postal Code:	K1N6N5
Country:	Canada
Email:	colleen.davison@queensu.ca

Co-author(s):

Name:	**Lisa Newton**
Title:	Research Analyst Assistant
Affiliation:	Toronto District School Board, Research & Information Services
Address:	1 Civic Center, Lower Level
City, State, Province or County:	Toronto
Postal Code:	M9C 2Bc

Country:	Canada

Co-author(s):

Name:	**Rob Brown**
Title:	Director
Affiliation:	Toronto District School Board, Research & Information Services
Address:	1 Civic Center, Lower Level
City, State, Province or County:	Toronto
Postal Code:	M9C 2Bc
Country:	Canada

Co-author(s):

Name:	**John Freeman**
Title:	Professor, Faculty of Education & Director
Affiliation:	Queen's University Social Program & Evaluation group
Address:	
City, State, Province or County:	Ontario
Postal Code:	
Country:	Canada

Co-author(s):

Name:	**Lee-Ann Ufholz**
Title:	Director
Affiliation:	University of Ottawa, Health sciences Library
Address:	451 Smyth Road
City, State, Province or County:	Ottawa
Postal Code:	ON K1H 8 M5
Country:	Canada

Co-author(s):

Name:	**J. David Smith**
Title:	Professor, Faculty of Education

Affiliation:	University of Ottawa
Address:	145 Jean-Jacques-Lussier Private
City, State, Province or County:	Ottawa
Postal Code:	K1N 6N5
Country:	Canada

REQUEST SUPPORT

Not specified.

ROLES AND RESPONSIBLIITIES

Colleen M. Davison led the writing of the protocol. She will perform some of the searches, assess documents for inclusion and exclusion, extract data, monitor the equity checklist and lead the writing of the review.

Lisa Newton wrote the review background, and reviewed and edited the protocol. She will assist with the handsearches, assess documents for inclusion and exclusion, extract data and be involved in the writing of the review.

Robert S. Brown reviewed and edited the protocol. He will extract data, lead the synthesis of the quantitative data (where required) and help write the review.

Lee-Anne Ufholz led the development of the electronic database search strategy and will continue to advise in this area, she will supervise a research assistant doing 15 of the electronic database searches and review and edit the final report.

John Freeman reviewed and edited the protocol. He will lead the assessment of risk of bias and be involved in the writing of the review.

J. David Smith reviewed and edited the protocol. He will lead the assessment of study quality and unit of analysis issues and will be involved in the writing of the review.

PRELIMINARY TIMEFRAME

31 March 2012

Plans for Updating the Review

Not specified.

AUTHORS' RESPONSIBILITIES

By completing this form, you accept responsibility for preparing, maintaining and updating the review in accordance with Campbell Collaboration policy. The Campbell Collaboration will provide as much support as possible to assist with the preparation of the review.

A draft review must be submitted to the relevant Coordinating Group within two years of protocol publication. If drafts are not submitted before the agreed deadlines, or if we are unable to contact you for an extended period, the relevant Coordinating Group has the right to de-register the title or transfer the title to alternative authors. The Coordinating Group also has the right to de-register or transfer the title if it does not meet the standards of the Coordinating Group and/or the Campbell Collaboration.

You accept responsibility for maintaining the review in light of new evidence, comments and criticisms, and other developments, and updating the review at least once every three years, or, if requested, transferring responsibility for maintaining the review to others as agreed with the Coordinating Group.

PUBLICATION IN THE CAMPBELL LIBRARY

The support of the Campbell Collaboration and the relevant Coordinating Group in preparing your review is conditional upon your agreement to publish the protocol, finished review and subsequent updates in the Campbell Library. Concurrent publication in other journals is encouraged. However, a Campbell systematic review should be published either before, or at the same time as, its publication in other journals. Authors should not publish Campbell reviews in journals before they are ready for publication in the Campbell Library. Authors should remember to include the statement: "This is a version of a Campbell review, which is available in The Campbell Library" when publishing in journals or other venues.

I understand the commitment required to undertake a Campbell review, and agree to publish in the Campbell Library. Signed on behalf of the authors:

Form completed by:

Date: