FEMALE REPRODUCTIVE FACTORS OF PARITY AND LUNG CANCER SYSTEMATIC REVIEW AND DOSE RESPONSE META-ANALYSES: PROTOCOL

1 Background

1.1 Problem definition

Many studies have evaluated association between female reproductive factors of parity, age at menopause, age at menarche, age at first birth, menstrual cycle length and use of exogenous female hormones (oral contraceptive pills and hormone replacement therapy) and lung cancer, but results have been inconsistent and difficult to understand. It is possible that differences in study population in terms of other factors (e.g., smoking status and age) may have contributed to the inconsistent results.

1.2 Hypothesis

Female hormonal and reproductive factors such as menopause, menarche, parity, age at first birth, menstrual cycle length, and contraception and hormonal replacement are associated with lung cancer risk. There may be a dose-response relationship between some factors such as parity or age at menopause and lung cancer risk. Further, risk may be modified by factors such as smoking status, histological type of tumour, and age of women.

1.3 Objective

To review systematically the association between female reproductive factors of parity, female hormones, female hormonal factors, age at menopause, age

at menarche, age at first birth, parity, menstrual cycle length, and use of exogenous female hormones (oral contraceptive pills and hormone replacement therapy) and lung cancer risk, taking into account factors such as smoking status, age, and lung tumour histology, looking at data from observational studies using a case-control, cohort or nested case-control, case-cohort approach and where available randomized controlled trials.

1.4 PICO summary

Population: general female population living within the community

Intervention: Level of exposure to female hormones, as measured by the following proxies

a) age at menarche
b) age at menopause
c) menstrual cycle lengths
d) age at first birth
e) parity
f) exogenous female hormone exposure (contraceptive or replacement therapy)

Comparison: lowest parity group, lowest age at menarche group, lowest age at menopause group, lowest age at first birth group, category with shortest menstrual cycle length, non-female hormone users

Outcome: lung cancer incidence or mortality

2 Sources

2.1 Search Strategy:

Search terms used
Outcome: Lung cancer or lung neoplasm
Exposure: Reproductive factors or female hormones or female hormonal factors or menopause or menarche or parity or number of children or contraception or hormone replacement or menstrual cycle or menstrual cycle length or menstruation or estrogen or progesterone
Search strategy used

The following formula will be entered into the search field in all database used: (lung cancer OR lung neoplasm) AND (reproductive factors OR female hormones OR female hormonal factors OR menopause OR menarche OR parity OR number of children OR contraception OR hormone replacement OR menstrual cycle OR menstrual cycle length OR menstruation OR estrogen OR progesterone).
No restriction on publication date or languages.
The extracted articles from the different databases will be combined, with duplicates removed. A title scan will be conducted to remove articles that were unrelated to the objective of this meta-analysis.
Exclusion criteria for the title review is where the title indicates that the subject matter of the article is:

a) not about lung cancer,

b) about metastatic lesions to the lung

c) about treatment of lung cancer, or prognosis or survival of lung cancer

d) about studies conducted on cells, including human cell-lines, rather than humans.

e) about studies looking at ecological associations

This will be performed by a single investigator.
All articles selected from title review will then undergo a full text review to check that the study fulfils criteria. In the event that the full text is unavailable, a review of the abstract will be performed. Articles where neither the full text nor abstract are available will not be included if we are unable to contact the author for the paper.
2 reviewers will perform full text reviews. In the event where the decision for inclusion made by the 2 reviewers is different and an agreement cannot be reached, a 3rd reviewer will review the article and make the final decision.
Data extraction will be performed by 2 investigators.

2.2 Databases used

PubMed; Web of Science; Embase

2.3 Hand searching

Look at citation lists of the papers included in the meta-analysis, and of relevant review articles identified during the article review stage, for studies

that have primary data on the outcome and factors of interest. These papers will be added into the list of articles selected for meta-analysis if they meet inclusion criteria.

2.4 Papers published other than in English

To obtain full paper in original language, translate and obtain estimates

2.5 Abstracts

Attempt to contact authors for data (give authors 2 weeks to reply from date of email before study excluded)

2.6 Contact with authors

Yes, to obtain details for papers where only abstract is available.

3 Study selection

3.1 Types of study included

Case-control, cohort, or randomized controlled trials.

3.2 Exclusion criteria

Cross sectional studies, ecological studies, case series, studies without control group or comparison group; studies without informative effect estimates and unable to contact authors; studies looking at survival or prognosis of a lung cancer cohort.

3.3 Data extracted

Data extracted will include

i) study characteristics (country, type of study, number of cases and controls or number of cohort, year study ended, time frame of assessment, exclusion criteria, study duration)

ii) study population (race composition, age, proportion smokers, proportion response rate/loss-to-follow-up

iii) exposures:

Parity - number of times respondent has given birth, median and range in each category.

Age at first birth - age of woman at birth of first child; median and range in each category.

Age at menarche - age of woman at start of regular menstrual cycles; median and range in each category.

Age at menopause - age of woman at cessation of regular menstrual cycles; median and range in each category.

Menstrual cycle length - average length of cycle of menses; median and range for each category.

Hormonal contraceptive use - use/no use; for users: duration of use (median and range in each category), age prior to outcome where use was first started (median and range in each category); type of contraception used (oral progesterone only, oral estrogen-progesterone combination, injection/implants).

Hormonal replacement therapy - use/no use; for users: duration of use (median and range in each category), age prior to outcome where use was first started (median and range in each category), type of HRT used (estrogen only, estrogen/progesterone combination)

iv) outcome (determination of outcome--self-report, registry, medical records, path review, all lung cancers, histological subtypes adeno vs non-adeno, covariates adjusted, 95%CI)

3.4 Method of data extraction

i) 2 authors will evaluate articles to determine eligibility; in event of uncertainty, a 3rd author will be consulted to arrive at a consensus on which articles will be included;

ii) data extraction performed by 2 authors; discrepancies resolved by repeating review and reaching a consensus;

iii) where there are more than 1 published article of the same study population, data from the latest article will be extracted

3.5 Assessment of confounding

Co-variates used in estimates will be extracted in data extraction; studies will be sorted based on extent of confounding control, and sensitivity analyses

and evaluation of heterogeneity will be performed based on this variable (i.e. control of confounding)

3.6 Assessment of study quality

2 researchers will independently evaluate study quality using the STROBE statements; differences in quality will be resolved by consensus with the whole group of researchers; `http://www.strobe-statement.org/index.php?id=available-checklists`
The following specific quality criteria will be assessed:

i) case-control vs cohort study

ii) for case-controls: community controls vs hospital controls

iii) for case-controls: blinded assessment of exposure vs unblinded assessment

iv) self-reported lung cancer vs registry/ medical records/path review

v) high participation (>80%) [for case-control studies] / low loss-to-follow up (<20%) [for cohort studies] rate vs low participation (<60%) /high loss to follow-up rate (>30%)

3.7 Assessment of heterogeneity

We will use

i) Cochrane Q and I^2 statistic to evaluate heterogeneity

ii) meta-regression to assess sources of heterogeneity: (type of study, ethnicity, proportion of smokers, number of participants in study, number of cases, adjustment for con-founders – smoking, ETS, age, family history of lung cancer, quality of study)

3.8 Statistical methods used

Stata programme to be used; random effects model; stratified analyses (see results d), evaluate publication bias (see results e);

3.9 Assessment and Analysis of Dose-response Associations

This will be performed for the following exposures: parity, age at first birth, age at menarche, age at menopause, menstrual cycle length and duration of hormonal contraception or replacement therapy use.
For every study, the mean or median value of the exposure of interest for each exposure category will be assigned to each corresponding effect estimates (OR/RR). If the median or mean per category is not reported in the article, the midpoint of the upper and lower boundaries of each exposure category will be used. If the lower or upper boundary for the lowest and highest category is not reported, a dose value which is at least as high as the lower bound plus the width of the second-highest category will be assigned to the uppermost, open-ended category. For the lower open-ended category, a value of its upper bound minus half the width of the next (second-to-lowest) category will be assigned.
To assess the shape of the dose-response relationship, all available data points from each study will be plotted; further, appropriate statistical methods will be employed to estimate dose-response associations.
For linear dose-response relationship, Generalized Least Squares for Trend estimation will be used to estimate the log-linear dose-response associations. This will be a two stage approach that estimates a slope/ beta-coefficient for each study at first stage, and then derives an overall estimate of dose-response effect by weighted average of the individual slopes.
For non-linear dose-response relationship, a restricted cubic spline model will be estimated using generalized least square regression taking into account the correlation within each set of reported effect estimates (OR/RR). Further, the study-specific estimates will be combined using the restricted maximum likelihood method in a multivariate random-effects meta-analysis. A probability value for nonlinearity will be calculated by testing the null hypothesis that the coefficient of the second spline is equal to 0.

4 Results

The data will be analysed separately for each exposure of interest:

a) parity

b) age at first birth

c) age at menarche

d) age at menopause

e) menstrual cycle length

f) hormonal contraception use

g) hormonal replacement therapy use

Result will include the following:

a) Figure 1: flow chart of study search

b) Table 1: summary of included studies

c) Figure 2: individual study estimates and overall estimates, for each exposure of interest

d) Figure 3: stratified estimates, for each exposure of interest

   i) type of study (case-control vs cohort),

   ii) smokers vs never-smokers

   iii) race (east asian/caucasian/other)

   iv) histological types (adenocar vs nonadenocarc)

   v) high quality vs low-quality studies (using quality parameters defined in Study Selection f)

e) Figure 4: evaluation of publication bias via Egger's test and funnel plot.

5 Discussion

Points to address in discussion:

a) Overall evaluation of data: is there an overall effect? if yes, is this consistent in all groups? if no, is an effect seen only in some groups (study type, smoking status, study quality, taking number of times respondent has given birth)

b) Strengths and weaknesses of study

c) Potential biases

d) Justification for exclusion of studies

e) Assessment of quality - does this affect estimates? sensitivity analyses

f) Alternate explanations for observed results - publication bias, study quality, type of study

g) Generalizability of results

h) Guidelines for future research

i) Disclosure of funding source