Analysis protocol

Title: How do hospital birth volumes and travel time to hospital impact birth outcomes?

Roles:

Leader: Johan Håkon Bjørngaard

Analysis: Andreas Asheim, Sara Marie Nilsen, Signe Opdahl

Clinical experts: Kari Risnes, Elisabeth Magnussen

Methodological experts: Neil Davies, Fredrik Carlsen

Proposed journals: BMJ

Research questions: Is it safe to give birth at a hospital with low birth volumes? How do hospital birth volumes and travel time to hospital impact birth outcomes?

1. Background

Safe childbirth depends on access to health care services. The Norwegian health care system is set up to provide universal and high-quality free health care coverage for the whole population. However, large parts of the country are sparsely populated and there is a resource trade-off between size and location of hospitals. From 1974 to 2008, the number of birth institutions in Norway was reduced from 131 to 53.[1] Centralization of deliveries is driven by increased medical specialization, better transport infrastructure and changes in the geographical population patterns. The 2010 Norwegian practice guideline from the Health Directorate includes a clear recommendation that women who are expected to deliver very or extremely preterm (before 32 weeks) are transferred to hospitals with staffed neonatal intensive care units.[2] However, since most preterm births cannot be predicted, pregnant women living close to specialized hospitals, may have better access to appropriate care.

Because larger hospitals are more specialized and handle more high-risk cases, comparison of delivery outcomes between larger and smaller hospitals is likely to be confounded. For example, if women who are identified as being at high risk of complications during delivery are sent to larger and more specialized hospitals, more adverse outcomes may be expected at such hospitals. Also, women living in catchment areas of larger hospitals may have systematically different sociodemographic characteristics compared to women residing in catchment areas of smaller hospitals. Since registry data include limited information about patients' characteristics and risk factors, adjustment for available information would not be expected provide sufficient control for such confounding.

1.1. Key references
https://obgyn.onlinelibrary.wiley.com/doi/full/10.1002/ijgo.12832
https://bmcpregnancychildbirth.biomedcentral.com/articles/10.1186/s12884-021-03988-y

2. Objectives

The aim of this project is to estimate effects of hospital birth volume and travel distance to hospital on perinatal and infant mortality, neonatal immediate condition (indicated by Apgar score), neonatal unit admission, caesarean section, transport deliveries, and maternal complications in labour (indicated by blood transfusion). We will investigate the effects for different categories of gestational age.

3. Analysis plan

Instead of relying on adjustment for characteristics of the women to ensure comparability and limit bias, we will analyse observational data as a natural experiment by comparing women living close in proximity, but still in different hospital catchment areas. The analyses will be made independently of where the women gave birth, comparing women who live in catchment areas to different hospitals.

In the Norwegian healthcare system, hospitals with maternity wards have separate geographical areas of responsibility, usually a collection of municipalities. Therefore, a large proportion of deliveries in each municipality will likely happen at the same hospital. The proportion will not be 100% since women may also to some degree choose hospital, and the birth may happen when the woman is away from home. Women that are identified as having a high risk of complications during delivery are sent to larger hospitals. In the capital, Oslo, several hospitals serve one municipality.

Our analysis will be based on assessing which hospitals a woman is expected to give birth at, given the municipality of residence, which is likely a strong predictor of where the birth takes place. Analysing the association with the characteristics of the expected hospital rather than the hospital where the birth takes place avoids bias due to, e.g., selection of more severe cases to bigger hospitals. We will, in other words, do an instrumental variable analysis with the hospital characteristics linked to the municipality of residence as an instrument.

There may be substantial differences in patient characteristics between municipalities. For example, mothers with higher education are more likely to live in cities. Educated women also tend to postpone parenthood to age ranges associated with more complications during pregnancy. We may however assume that women living nearby each other may be comparable on measured as well as unmeasured characteristics, although they reside in different municipalities. Differences in which hospitals are used across municipality borders may therefore be a candidate instrument for the characteristics of the hospital where the birth takes place. We will match groups of births such that each group consists of births within the same year by women who reside in one of two municipalities belonging to different catchment areas. By only analysing variation in outcomes within such groups, we may effectively control for confounding that could be caused by geographical differences.

The analysis may be thought of as an analysis of a hypothetical intervention: If a woman's place of residence was in the same geographical area, but in another hospital catchment area, would the birth outcome have been different? We assume that women in reproductive age do not choose place of residence to situate themselves close to one or the other hospital. Therefore, within the described comparison groups we might expect few substantial associations between municipality of residence and characteristics of the women. In this sense, the municipality of residence may be analogous to the randomisation in a randomised experiment, within the groups. A central sensitivity analysis is to investigate whether there could be any associations between characteristics of the

hospital where delivery is expected to take place, and characteristics of the women who give birth, within the matched groups.

3.1. Eligibility criteria
All births in Norway between 1999 and 2016, as recorded in the Medical Birth Registry of Norway.

3.2. Definition and calculation of the exposure
The main exposure in the study will be hospital size, measured as the yearly number of births. We will also study travel time as an exposure.

Interactions will also be studied: effects of hospital volume given comparable travel time, and effects of travel time given similar hospital size.

Instrument

The exposures (volume and travel time) will be instrumented by the characteristics of the hospitals where women in a municipality usually give birth. To do this, we consider births at the two most used hospitals for each municipality. The fraction of births at each of these hospitals may be thought of as the probability of giving birth there, and the expected value of hospital characteristics for women residing in a given municipality may be computed from these. For example, if eight women from municipality *M* give birth at hospital *a,* and two women give birth at hospital *b*, the expected hospital size for women in *M* would be 0.8 times the total yearly number of births at hospital *a* plus 0.2 times the total yearly number of births at hospital *b*. This will likely be a strong predictor of characteristics of the hospital where a woman from this municipality gives birth. For this computation, we discard births at hospitals far away or births at hospitals that receive less than 10% of the mothers from the municipality. Such cases could result from the woman having moved without changing address or similar peripheral cases, or a selection of high-risk pregnancies. Since organisation of catchment areas has changed during the study period, we do the calculation separately for each year.

3.3. Outcomes
As outcomes we will use perinatal mortality (stillbirth or death within 1 week), infant mortality (death within 1 year), Apgar score < 7, neonatal unit admission, caesarean section (planned and emergency), birth before arrival at the hospital, and maternal blood loss.

3.4. Matching of participants
In the main analysis, participants will be grouped such that each woman is only compared with women from a neighbouring municipality giving birth in the same year. Women from two municipalities will only be grouped together if residing in each of the two municipalities is associated with substantially different probabilities of giving birth at the same hospital, indicating that they live in different catchment areas.

To assess whether two neighbouring municipalities are in different catchment areas, we will use a decision rule based on differences in proportion of births happening at different hospitals. If the highest ranked hospital is different for two municipalities, we assume that the municipalities belong to different catchment areas. If the highest ranked hospital is the same, but there is a difference of more than 20 percentage points between the municipalities in their proportions of births at this hospital, we also allow the municipalities to be grouped together.

Women from one municipality may potentially be matched with women from several neighbouring municipalities, since one municipality may have more than one neighbouring municipality in different catchment areas. In such cases, to ensure that standard hierarchical methods can be used, we will require that each woman is included in only one matched group. Participants from each municipality will then be split by neighbourhood (a subdivision of municipality) matching with the closest neighbouring municipality (see schematic illustration in Figure 1). To assess which is the closest municipality, we will calculate the number of neighbourhood units between the participant's neighbourhood and the border.

Figure 1: Comparable groups of women are assembled by matching across municipality borders. The matching is done such that births by women in one municipality are matched with those of the closest bordering municipality with different hospital preferences. In this schematic example, women in municipalities C and D give birth at the same hospital and will not be matched. Every birth by women in C are matched with those in A that live closer to C than B. Similarly, births by women in D are matched with those in B that live closer to D than A. Finally, births by women in A and B that live close to the border between A and B are matched in a third group.

Match municipalities A and C

Match municipalities B and D

Match municipalities A and B
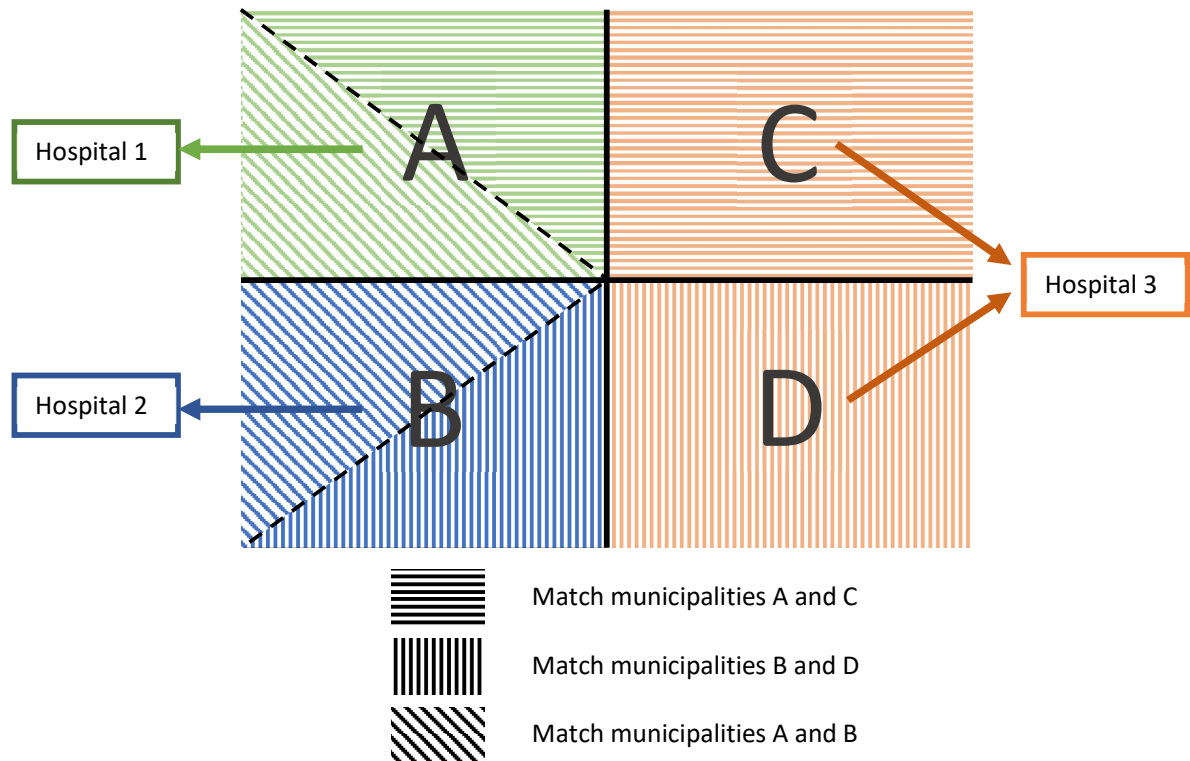
Matching across all such pairs of neighbouring municipalities may not yield a balanced design, since densely and sparsely populated municipalities with very different characteristics may share a border. This may be managed by restricting the comparison to municipalities with less than, e.g., 100 births per year. We may also limit potential bias from comparing densely and less densely populated areas by restricting the comparison to those living close to the border.

3.5. Adjustments

We will present adjusted and unadjusted estimates. For adjustment we will use:

- The woman's age at delivery (in categories of five years)
- The woman's level of education
- Parity (categories: 0, 1, 2, 3, 4, 5+)
- Municipality level indicators of socioeconomic position

3.6. Cluster correction of standard errors

Since the instrument (characteristics of expected hospital) is defined at the municipality level, we correct standard errors for clustered observations according to municipality, and on the woman's id-variable, since each woman may be represented with several births in the data.

3.7. Subgroup analyses

We will present outcomes for preterm (gestational age <37 weeks), early term (gestational age 37-38 weeks), full term (gestational age 39-41 weeks) and post term (gestational age 42-44 weeks) births.

The analyses will be performed on subgroups according to different years of the study period to assess whether effects would be heterogenous over time.

To assess possible non-linear effects of volume and travel-time, we will do the analyses for pairs of municipalities selected by the expected hospital characteristics. For example, a heterogenous treatment effect model, where we find the effect of hospital volume for those who are expected to give birth at a small hospital, we could consider only pairs of municipalities where one municipality is in the catchment area of a small hospital. We may also choose to do this as a non-linear instrumental variable analysis.[3]

3.8. Statistical methods

The analyses will be performed using within-groups estimators for Stata 15.1. We will use methods that allow for non-hierarchical clustering, *reghdfe*, *ppmlhdfe*, and *ivreghdfe* (*http://scorreia.com/software/*). By defining a grouping variable *id* that identifies births in the same matched groups of mothers, we will run these methods for computing cluster corrected within-group estimates by using the option *absorb(id) vce(cluster municipality_id mothers_id)*.

For estimates of the exposure-outcome effect, we will use *ivreghdfe*, giving estimates on a linear scale. For analyses of instrument-outcome associations, e.g., instrument relevance and sensitivity analyses, we will use *reghdfe* for continuous outcomes and *ppmlhdfe* for binary outcomes, providing results in risk ratios.

For visualisations we will use Rstudio with the ggplot2 library. For making map visualisations we will use the ggmap library.

4. Assumptions and additional analyses

The main assumptions for the analyses are that the instrument satisfies four instrumental variable assumptions. These should be satisfied when considering variation *within the matched groups*:

1. The municipality of residence should be a strong predictor of the volume of births at the hospital where a woman gives birth (relevance)
2. The municipality of residence should not be associated with any characteristics of the woman or the pregnancy which also associate with the outcome (independence)
3. The municipality of residence should only affect birth outcomes through affecting at which hospital the births take place (exclusion restriction)
4. Moving a mother from living near a high-volume hospital to a low volume hospital should never decrease the probability of giving birth in a low volume hospital (monotonicity)

Assumptions 1-3 are the identifying assumptions needed for the analysis to be valid. The fourth assumption is a point identifying assumption. This assumption will identify a local average treatment effect – i.e., the effect in individuals whose treatment is affected by the instrument (where they live).

Figure: Illustration of the identifying assumptions of the analysis with a directed acyclic graph. Within the matched groups, there should be no association between characteristics of the woman or pregnancy and which side of the border the woman resides. Neither should there be any causal link between where the woman resides and birth outcomes, except through the hospital of delivery.
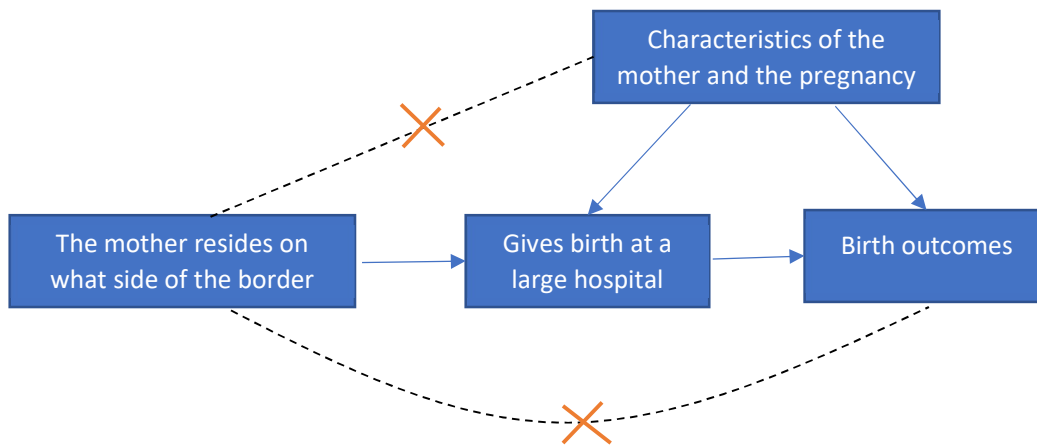
The results of an instrumental variable analysis must be interpreted as a local average treatment effect, meaning that the effect is an average for those births that were affected by the instrument. Some women will need more specialised health services and will have a much higher probability of giving birth at a more specialised hospital regardless of where they live. However, these women would have to be referred to such hospitals after identification of high risk in the prenatal care program.

4.1. Instrumental analysis assumptions

Instrument relevance

We will assess instrument relevance by computing the association between the instruments, expected hospital volume and travel time, and the exposures, reporting associations as estimated associations and F-scores.

This analysis will also be done for subsets of the population to assess which births are most affected by the place of residence.

- Preterm births (<week 37), early term births (week 37-38), term births (week 39-41), post-term births (>=week 42)
- Women younger than 25 years, 25-35 years, 35 year and older
- Larger and smaller municipalities

Instrument balance

To investigate instrument balance, we will assess the within matched group association between the instruments and characteristics of the woman or the pregnancy:

- Age of the woman (under 25 years, over 35 years)
- The woman's education level (university level or not)
- Age of partner
- Education of partner
- Pre-eclampsia
- Conception by artificial reproductive technology
- Multiple birth

Finding no substantial associations would strengthen the assumption of instrument balance.

Exclusion restriction

To be decided.

Monotonicity

We expect that the direction of our instrument would be in the same direction for all women. So far, we have no test for this.

4.2. Within-mother analysis

An alternate way of analysing the data is by comparing outcomes for women who moved between catchment areas of different hospitals between births. This analysis entails control over all stable characteristics of the woman. Some sources of bias remain, however, for example from conditioning on having moved. Time-trends may be an issue, for example if women tend to move to larger municipalities after their first delivery. This could be a source of bias that to a certain extent can be

managed by adjusting for age and parity. The analysis will give an estimate of the same effects as the main analysis, but with different assumptions, providing a path for triangulating the results.

In these analyses we will cluster standard errors by mother and municipality. We will adjust for age in five-year categories and parity. In addition, we will adjust for year of childbirth, which achieves adjustment of mother's age in years. We will use the same instruments as in the main analysis, defined on the municipality-year level.

4.3. Additional analyses

Exposures as mediator in outcomes of preterm birth

We will assess whether hospital birth volume and travel time are mediators in the association between gestational age and birth outcomes.

Timing of subsequent births

If complications during birth impair fertility, or affects a woman's wish to conceive again, this could lead to a delay in the timing of a subsequent birth. We will therefore use time to a subsequent birth as an outcome in an additional analysis. This will be analysed as a time-to event Cox regression, with a follow-up time of four years. We will use expected hospital size as exposure, i.e., this is not an instrumental variable analysis.

Gestational age differences

If different hospitals have different practices on when to perform a c-section or induction, this could result in selection issues when we condition on gestational age. Therefore, we will perform additional analyses of the association between the hospital volume and gestational age, to assess whether bigger hospitals may have different practices than smaller.

Births selected to bigger hospitals

Our analysis is not affected by selection of high-risk births to big hospitals, but analyses based on comparing hospitals may be. Mothers that give birth at big hospitals, although they reside in municipalities in the catchment areas of small hospitals, may to a large extent be selected for medical reasons. These births are interesting, since they might be used to assess the extent to which selection of high-risk births happens. Therefore, we will study the characteristics of such births to approach characteristics of births that are directed to bigger hospitals.

Geographical localisation of effect estimates

Only differentially exposed women will contribute to the analyses, and these will likely be concentrated near borders of hospital catchment areas. To visualise the geographic areas that do contribute, we will draw a heat-map showing the density of differentially exposed mothers.

Conventional analysis

The analysis is designed to minimise confounding due to selection of severe cases to specialised hospitals. We will do a direct analysis to investigate the potential for such confounding. This analysis will be the association between hospital volume, travel distance and birth outcomes, adjusted for

- Age of the woman
- The woman's education level (university level or not)
- Age of partner
- Education of partner
- Pre-eclampsia
- Conception by artificial reproductive technology
- Multiple birth

The estimated effect will be presented along with the effect estimates from the main analysis and sensitivity analyses.

Perturbations of parameters

We may perform the analyses with perturbed parameters. For example, the parameters that determine whether two municipalities have different hospital preferences (e.g., 20 percent difference in fraction of births at the top ranked hospital)

5. Data

Data on all deliveries in Norway from 1999 to 2016, as registered in the Medical Birth Registry of Norway (MBRN). The MBRN collects data on all pregnancies ending after 12 completed gestational weeks, including data on the woman's health before and during pregnancy, complications during delivery, vital status, and health of the new-born.

Data on residence and education level at the time of delivery were obtained from Statistics Norway.

Tables of travelling distances from all living circuits to all neighbouring municipality centres, and hospitals will be use, provided through the Norwegian Public Road Administration's routing services (https://labs.vegdata.no/ruteplandoc/). Data on municipality and living circuit topology is obtained via publicly available data (www.kartverket.no).

The study and data linkage were approved by the Regional Committee for Medical and Health Research (REK-Midt 2018/32).

1. Engjom HM, Morken NH, Norheim OF, et al. Availability and access in modern obstetric care: a retrospective population-based study. *BJOG: An International Journal of Obstetrics & Gynaecology* 2014;121(3):290-99.
2. Helsedirektoratet. Et trygt fødetilbud. Kvalitetskrav til fødeselsomsorgen. In: Helsedirektoratet, ed., 2010.
3. Burgess S, Davies NM, Thompson SG, et al. Instrumental Variable Analysis with a Nonlinear Exposure–Outcome Relationship. *Epidemiology* 2014;25(6)