Challenges with the interpretation of Sami ethnicity and the height differences in epidemiologic research comparing Sami and non-Sami population

Susanna Ragnhild Andersdatter Siri, Vilde Lehne Michalsen

Susanna Ragnhild Andersdatter Siri, MPH, Ph.D., Department of Community Medicine, Centre for Sami Health Research, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø. E-mail: susanna.r.siri@uit.no
Vilde Lehne Michalsen, MD, Ph.D., Department of Community Medicine, Centre for Sami Health Research, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø. E-mail: vilde.l.michalsen@uit.no

Ethnicity clusters individuals with shared characteristics and is a common variable in epidemiology. Often it is not specified what is meant by a given ethnicity or what characteristics ethnicity entails in the different contexts. In public health research, ethnicity can be viewed as a non-modifiable risk factor of disease because modifying the ethnic composition in a population is not meaningful, nor ethical, public health policy. Importantly, ethnicity is typically associated with modifiable risk factors that are possible to intervene on for improvement of public health, and hence viewed as a risk marker, not a risk factor. In the SAMINOR 1 Survey (2003–2004) from Northern and Mid Norway, Sami people were on average 6 cm lower than non-Sami people that lived in the same rural area. Low height has been associated with a higher risk of coronary heart disease, stroke, and diabetes, lower risk of some cancer types, and lower socioeconomic status. This article aims to discuss the use and interpretation of ethnicity and height in epidemiologic research when comparing Sami to non-Sami populations. Body height is a highly heritable trait in individuals, whereas, at a population level, the average height may be a marker of long-term exposure to health and living conditions, and hence, a non-modifiable variable for interventions at present. We argue that height may be both an important risk factor for chronic lifestyle disease and an explanatory variable in studies of chronic lifestyle diseases in Sami vs non-Sami populations. For better interpretability of study results, we encourage researchers to express what components or characteristics are assumed to comprise Sami ethnicity and height and clarify which variables may provide interventions. Furthermore, we urge researchers to be aware of the potential for bias when studying health measures influenced by height, such as the popular body mass index.

Public health and challenges with non-modifiable epidemiological variables

Public health is among others the science of preventing disease and monitoring and improving the health of populations [1]. Often, it is relevant to disaggregate population data on ethnic belonging, as ethnic groups may differ in their distribution of sociocultural determinants of health, e.g., diet, socioeconomic status, access to health care, etc. However, in epidemiologic research including ethnicity as a variable, it is rarely expressed what ethnicity is assumed to comprise in terms of sociocultural factors. The use of ethnicity in medical and health research has been criticised as imprecise and vague [2]. An individual's ethnicity, and in particular self-identified ethnicity, is changeable throughout life. Albeit being changeable within an individual, ethnicity as a variable in epidemiology can be viewed as non-modifiable. The fact that a factor, or social group label, changes voluntarily or non-voluntarily in an individual, does not mean that it is modifiable from a public health perspective, because changing the ethnic composition in a population is not meaningful, nor ethical, public health policy. Principally, ethnicity is viewed as a risk marker, not a risk factor, in epidemiology [3]. Hence, ethnicity is rarely a cause of ethnic differences in health outcomes, which is why we need to identify the factors related to ethnicity and make public health interventions by modifying them.

The fluidity of ethnicity makes it challenging to interpret when included in statistical models. It is also challenging to convey meaningful information that is useful for public health interventions because we regard ethnicity as non-modifiable from a public health perspective. For instance, how can public health officials use the information that Sami have a 30% higher risk of cerebral stroke? [4]. Researchers usually aim to explain ethnic differences in disease, as the explanation of ethnic disparity often is a result of a different distribution of known risk factors, both distal and proximal. However, in some cases, explanations for ethnic differences in disease may complicate interpretability further. For instance, height explains some of the excess risk of cerebral stroke in Sami vs non-Sami [4]. Importantly, because height is yet another non-modifiable risk factor at a population level, the public health message is not clear [4]. Neither ethnicity nor height can be intervened on to improve the health of people.

Heterogeneity within Sami ethnicity

To give meaningful interpretations of Sami ethnicity in epidemiologic research, it is vital that researchers have an idea about what Sami ethnicity is a marker of, i.e., which components comprise Sami ethnicity as a variable. For instance, it may comprise exposure from a specific lifestyle, e.g., reindeer herding and traditional diet comprising reindeer meat and fatty fish, but only a minute fraction of Sami people are involved in traditional reindeer-herding, and there are variations in diet between and within Sami and non-Sami [5]. Geography and Sami ethnicity correlate, and geography correlates with distance to health services and higher education. Sami people may be exposed to ethnic discrimination and bullying [6], and Sami ethnicity is correlated with lower height [7]. However, all these components or characteristics of Sami ethnicity may overlap with non-Sami and are distributed with variance within the Sami people. Hence, there is a potential for heterogeneity of risk within the Sami population. For instance, Sami people living in urban cities may have different correlations to the abovementioned components and characteristics than Sami in rural areas, which again may differ from reindeer-herding Sami. Importantly, the characteristics of Sami ethnicity may vary by time or birth cohort as society changes. Sami children growing up in the first half of the 20th Century were sent to boarding school as part of the assimilation process [8, 9], and the Sami people have a different status today than during that period [10].

Sami ethnicity may be particularly challenging to use in research as compared to other non-indigenous ethnicities, due to the history of unethical racial skull measurements [11] and governmental assimilation [9] that stereotyped and stigmatised the population as inferior to the Norwegian population [11, 12]. Therefore, researchers that use Sami ethnicity should be careful not to cause further harm or stigma of the Sami people or depict Sami ethnicity to be inherently inferior or harmful in itself. We acknowledge the ethical challenges attached to the use of Sami ethnicity in epidemiologic research; however, in this article, we do not discuss these ethical perspectives any further.

In this article, we aim to discuss the implication and interpretation of ethnicity and height in epidemiologic research where the Sami population is compared to the non-Sami population that lives in the same region. We aim for two outcomes: 1) a better understanding of the challenges when using non-modifiable and multi-dimensional concepts such as Sami ethnicity in epidemiological studies, and 2) a discuss the public health challenges that arise from the use of non-modifiable variables (Sami ethnicity and height). The intention of this article is not to be instructive, but to provide some reflections of the interpretation of non-modifiable variables such as Sami ethnicity and height, when included in statistical models.

Ethnicity

Ethnicity has been defined as: "The social group a person belongs to and either identifies with or is identified with by others, as a result of a mix of cultural and other factors including language, diet, religion, ancestry and physical features traditionally associated with race. All people have an ethnicity—not only minorities"[13]. As such, ethnicity may include several dimensions of human life as listed in the definition, including socioeconomic status and culture, where the latter have itself many layers [14]. Some of these characteristics may have biological implications through their effects on disease and health, making ethnicity a relevant and common marker or a proxy variable (i.e., representing something else) in modern epidemiology. Ethnicity is not synonymous with race if "race" is narrowed down to distinct human genetic clusters, which is controversial and scientifically refuted [15, 16]. There is however genetic variation on a continuum between and within population groups worldwide [16]. In epidemiology, ethnicity has surpassed "race" as a label of population group belonging, although "race" as a concept is still in use, particularly in the United States, primarily as a sociocultural concept [15].

Norwegian Sami ethnicity

In 2019, a public health report on data from a random sample of the adult population in Troms and Finnmark County in northern Norway (43.5% participation rate), showed that approximately 40% had some connection to either Sami or Kven ethnicity [17]. Among these, roughly 30% were categorised as Sami, 20% as Kven, 16% as both Sami and Kven, and 5% as having Sami speaking grandparents. The remaining 25% could not be placed in an ethnic category. This illustrates the complicated and mixed composition of ethnicity in this population. Sami and non-Sami cannot be viewed as mutually exclusive population groups. Therefore, defining ethnicity is no straightforward task. The criteria to the Norwegian Sami Parliament electoral roll might also illustrate this: an individual must self-identify themselves as Sami, and in addition, must have Sami as the primary language spoken at home, or have parents, grandparents, or great grandparents that speak/spoke Sami, or be a descendant of one that is/was enrolled. In the SAMINOR Study, comprising the SAMINOR 1 Survey, the SAMINOR 2 Questionnaire Survey, and the SAMINOR 2 Clinical Survey [18-20], information on various aspects of ethnicity (language use, ethnic background, self-perceived ethnicity) was collected through self-reported questionnaires. In studies using data from the SAMINOR Study, participants are typically assigned an ethnicity based on replies to the self-reported questions on language use, ethnic background, and self-perceived ethnicity. The use of several items in the formal definition of Sami ethnic belonging of participants can be justified based on the structural assimilation of the Sami people lasting at least a century. However, which items or aspects of ethnicity, i.e., language use, ethnic background, or self-perceived ethnicity, are needed to be fulfilled to obtain a given ethnic label differ between studies and may be highly dependent on the given research question. The categorisation into ethnic groups is, however, outside the scope of this article. Research shows that the choice of criteria for defining Sami ethnicity impacts the final group in terms of size, geography, and partly characteristics such as household income and self-rated health [21]. Evidently, Sami ethnicity is a heterogeneous cluster comprising several subgroups with different characteristics.

Interpretation of the coefficient of ethnicity in statistical models

Sami ethnicity is often used as an exposure variable (independent variable) in statistical models. The estimate (often a regression coefficient) in such models may be interpreted as the "effect of Sami ethnicity". However, the "effect of Sami ethnicity" is not directly meaningful, as ethnicity is principally used as a proxy measure of something else that causes disease. That is, Sami ethnicity is itself not a causal factor but associated with factors that are, and hence, Sami ethnicity is regarded as an indirect cause of disease. Scholars in the causal inference field have challenged the idea that there is something inherent with ethnicity or race that causes disease [22, 23]. Only in very special circumstances, it is relevant to refer to the "effect of ethnicity", e.g., in experimental studies that randomise by ethnic belonging. For instance, in job application studies it is possible to randomise participants to either having a Sami or a non-Sami-sounding name, to estimate the effect of Sami ethnicity on an initial job application, i.e., number of positive responses.

In studies using statistical modelling, ethnic groups are compared based on the idea of a counterfactual. This offers a conceptual idea of what would have happened if all study participants with Sami ethnicity did not have Sami ethnicity. This is parallel to a randomised controlled trial, where different outcomes can be attributed to the different interventions. However, the counterfactual imagination of a different outcome might only be meaningful in circumstances where an intervention can be specified [23]. Regarding Sami ethnicity, this is a variable that is not meaningful to intervene on from a public health perspective because 1) it is unethical, and 2) it is rarely that ethnicity itself causes disparity, but rather related modifiable risk factors at the individual (e.g. level of physical activity, smoking status, dietary habits) and/or societal level (e.g. socioeconomic status, access to health care).

The counterfactual of a treatment A is to not have treatment A (it is not to have treatment B). Ethnicity does not have a counterfactual, technically speaking, as all individuals have an ethnicity. In the given context, those that do not have Sami ethnicity, have other ethnicity/-ies, which can be termed non-Sami, Norwegian, Kven, or even mixed, and differ in some aspects from the Sami. Similarly, the counterfactual to female sex is, technically, not having female sex. However, the "opposite" of the female sex is the male sex, which is technically different from "not having female sex". Hence, both ethnicity and sex do not have counterfactuals per se, but rather have different levels within that category, which are then compared against each other. Therefore, the interpretation of ethnicity, or sex, in a statistical model is the effect or difference in an outcome if the group with Sami ethnicity had non-Sami ethnicity (or e.g., if women were men). To clarify, it would *not* be the effect or difference in an outcome if the group with Sami ethnicity had removed their Sami ethnicity. The conceptual difference is subtle, but important. Moreover, in statistical modelling, the Sami category might be viewed as "the exposed", whereas the comparison group, often the non-Sami, is viewed as "the unexposed". Using counterfactual conceptual thinking may thus tacitly encourage ideas of ethnicity as something that is modifiable, possible to intervene upon, or removable. Furthermore, it may depict Sami ethnicity as being inferior, or that there is something harmful with being affiliated with the Sami community. Because the conceptual ideas and semantics used in statistical models might be less suitable for non-modifiable variables such as Sami ethnicity, careful interpretation is needed.

In their article "On the causal interpretation of race in regressions adjusting for confounding and mediating variables", VanderWeele and Robinson [23] suggest using either a strong or a weak interpretation of ethnicity in regression models, of which neither expects a counterfactual to ethnicity nor an "effect of ethnicity". The strong interpretation requires several assumptions to hold, whereas the weak interpretation requires fewer assumptions [23].

Weak interpretation of ethnicity

The weaker interpretation of the coefficient of ethnicity in a statistical model is described as the observed health disparity or ethnic difference in an outcome (dependent variable), without further explanation of the coefficient of ethnicity itself [23]. Instead, explanatory variables included in the model are given more attention. The weaker interpretation of the ethnicity coefficient (e.g., Sami ethnicity) may be interpreted as the remaining ethnic difference given that the explanatory variable(s) is equalised across the comparison groups (e.g, between Sami and non-Sami). As such, the concept of ethnicity (e.g., Sami ethnicity) persists as an umbrella term for various characteristics that are not described (e.g., culture, discrimination, etc.), whereas the explanatory variables entail possible interventions (e.g., diet, access to health care, etc.). Importantly, this interpretation presumes that the explanatory variable(s) included in statistical models can be manipulated, as ethnicity (e.g., Sami ethnicity) is something that is not meaningful nor ethical to modify. Rather than focusing on ethnicity, attention should be given to the explanatory variables included in the analyses or be identified by directed acyclic graph (DAG).

Strong interpretation of ethnicity

VanderWeele and Robinson's stronger interpretation of the coefficient of ethnicity is related to causality as opposed to the weak interpretation [23]. This interpretation requires several assumptions to be true, and ideally, researchers should draw these assumptions using DAG methodology [24]. Put simply, DAG methodology utilises expert knowledge on relationships between variables. That is, the relationships between ethnicity (e.g., Sami ethnicity), including the various characteristics that are attached to ethnicity (e.g., culture, lifestyle, dietary habits, and socioeconomic status), and the outcome, should be visualised, including direct and indirect effects, and possible backdoor pathways. From a set of mathematical rules, the DAG determines which variables should be adjusted for to obtain unbiased estimates [24]. Using DAGs, the direct effect of, e.g., Sami ethnicity, is the estimate of the relationship between Sami ethnicity and the outcome that is not mediated through other variables, e.g., lifestyle, dietary habits, and socioeconomic status. The indirect effect of Sami ethnicity in DAGs is the estimated effect of the relationship mediated through mediators, and the total effect is the sum of the direct and the indirect effects (figure 1). Note that a thorough explanation of DAG concepts is not the aim of this article, and the reader is referred to other sources [24].

The strong interpretation of the coefficient of Sami ethnicity in a causal model requires a clear, well-defined, and explicit definition of the various characteristics that comprise, or at least are associated with, Sami ethnicity in the given context. It also requires valid data on the various characteristics. Using examples from population surveys in the U.S., VanderWeele and Robinson suggest that the characteristics of e.g., black ethnicity could be the joint effects of culture, dietary habits, skin colour, history (e.g., assimilation), genetic background, and exposure to racism, wherefrom it is possible to decompose the total joint effect of ethnicity into direct and indirect effects. Using such causal diagrams, it is possible to estimate how much of the effect of ethnicity is due to one of the components constituting ethnicity, for instance, dietary habits [23].

In epidemiologic research on the Sami people, a clarification of the various characteristics of Sami ethnicity might also offer insight into specifically whom within the Sami community the "effect of Sami ethnicity" is relevant for, and what assumptions are required to hold, ideally specified in a DAG. If ethnicity is not well-defined in terms of its characteristics, then the interpretation and reporting of the public health messages may be challenging. Importantly, because Sami ethnicity is non-modifiable from a public health perspective, such causal modelling may provide useful information on modifiable risk factors (e.g., lifestyle, dietary habits, socioeconomic status).

Another key point is that interpretation of ethnic differences may become even more complicated if the explanatory variable - for instance height - is a marker or proxy variable that is not clearly defined.

Adult height is a stable and well-defined variable at an individual level, but at the population level, it might be a marker or proxy variable of the living conditions this population has had over time [25-27]. If neither height nor ethnicity is modifiable from a public health perspective, these findings fail to guide public health officials and health professionals.

Figure 1, panel A shows a DAG that illustrates how the suggested relationship between ethnicity and incidence of cerebral stroke was modelled in the paper by Siri et al., wherein conventional risk factors and height were the explanatory variables. [4]. Figure 1, panel B builds on the DAG in Figure 1A and has added complexity regarding the causal relationships. This more complex DAG includes variables we cannot adjust for because they have not been identified (e.g., unmeasured mediators) or are difficult to measure (e.g., history). Hence, the "effect" of these variables are all included in the umbrella term ethnicity.

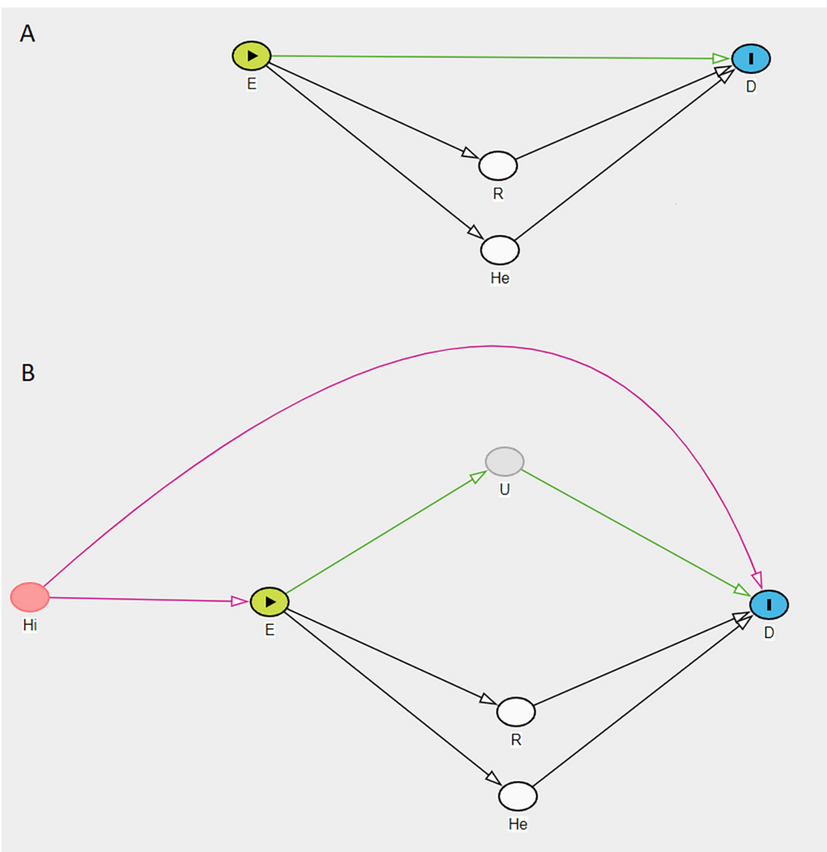

Figure 1. A simplified DAG of the suggested causal relationships in the paper by Siri et al. [4]. E = ethnicity, R = conventional cardiovascular risk factors, He = height, D = cerebral stroke, U = unmeasured mediators, Hi = history.
Panel A: Sami vs non-Sami ethnicity (E) had a hazard ratio (HR) of 1.31 for incidence of cerebral stroke (D). Adjustment for the mediator height (He) attenuated this HR to 1.18, and hence explained some of the excess hazard of cerebral stroke in Sami. Additional adjustment for conventional cardiovascular risk factors changed the HR to 1.19 and hence did not illuminate the excess hazard any further. Note that the HR of 1.31 refers to the total effect of ethnicity, whereas the adjusted HR of 1.19 refers to the direct effect of ethnicity. "Effect" refers to the difference in cerebral stroke incidence between Sami and non-Sami.
Panel B: Principally, ethnicity (E) is not viewed as a cause of disease. Rather, the adjusted HR of 1.19 is thought to be mediated through unmeasured mediators (U). The adjusted HR of 1.19 could also be confounded by unmeasured historical events (Hi), which may both affect ethnicity (E) and incidence of cerebral stroke (D). The adjusted HR of 1.19 may also be due to residual confounding, as the authors suggest.
The figure was made using the open software at www.dagitty.net by Textor et al.

In the next paragraphs, we make an effort to explain the implication and interpretation of height as an explanatory factor when interpreting epidemiological findings for Sami ethnicity.

Height

Average population height differs substantially across the world: the highest statures are found among European populations and the lowest among South/Central-American and Asian populations [28]. Height is a phenotype that is a highly heritable trait and a polygenetic trait, which means that it is determined by a combination of many genetic loci that each has a small effect [25]. Individuals inherit a genetic height potential, but whether the height potential is reached is partly dependent on environmental factors. Environmental factors during growth periods in childhood (i.e., below the age of 2 years) are crucial for variations in individual height [25-27]. Environmental factors that influence the final attained height include various nutritional factors, persistent infections, chronic diseases, use of medications, psychosocial stress, toxicants, and physical environment, wherein nutrition and disease are the most pervasive factors [29]. These factors are associated with social, cultural, political, and economic conditions [25, 26], and become markers or proxy measures of childhood social and economic conditions [27].

Sami ethnicity and height

In Norway, the Sami population has on average lower height than non-Sami [30], and this has also been observed in Finland and Sweden [31, 32]. In 1974–1975, the height difference between "Norse" and "Lappish" in Finnmark County was approximately 9 cm in men and 8 cm in women [7, 30]. In 2003–2004, the height difference between Sami and non-Sami in 24 rural municipalities in Northern and Mid-Norway was approximately 6 cm in both sexes [4, 33]. In 2012–2014, the height difference between Sami and non-Sami in 10 municipalities in Northern Norway was approximately 5 cm in both [34]. These figures indicate that the average height difference has diminished over time, but comparisons must be made with caution due to different geographical areas included in the sampling and variations in definitions of Sami ethnicity.

Several articles have discussed or included height in the context of differences in various health measures between Sami and non-Sami, e.g., cerebral stroke [4, 7], diabetes mellitus [34], and obesity [33]. In a 13-year follow-up study, using data from the SAMINOR 1 Survey (2003–2004) conducted in Northern and Mid-Norway, it was found that Sami had a higher incidence of stroke than non-Sami, and that height explained more of the excess risk than conventional cardiovascular risk factors did [4]. This supports previous findings by Njølstad et al.[7].

In a study comparing levels of various metabolic markers (e.g., triglycerides, glucose, blood pressure) at the same values of various obesity measures (e.g., BMI and waist circumference) in Sami and non-Sami, some ethnic differences were found after adjusting for lifestyle and prescribed drug use [33]. This marginal difference disappeared after adjusting for height. Moreover, the last author of this article recently showed that BMI comparisons in Sami and non-Sami may be invalid due to the negative correlation between BMI and height (published in this journal and issue). That is, even though the formula for BMI incorporates height (i.e., weight in kilograms divided by height in metres^2), there is a residual small negative correlation between BMI and height, which biases comparisons in groups that differ markedly in average height. These results also show that when correctly adjusting the BMI formula for height, there are no differences in obesity between Sami and non-Sami. This is potentially an important finding, as previous research has declared higher obesity prevalence in Sami vs non-Sami [5, 35].

The observation that height differences explain ethnic disparities between Sami and non-Sami, has been challenging to explain because it has not been clearly specified what Sami ethnicity and height represent in the different studies. What does it mean that height "explains" excess disease risk? Principally, height is both a risk factor for disease independent of ethnic belonging (i.e., in both Sami and non-Sami) and an explanatory factor for differences in disease rates between ethnic groups. From a life course perspective, height might be regarded as an intermediate variable between Sami ethnicity and the outcome that is by itself influenced by the historical, social, and political circumstances during childhood. When viewed as a risk factor for disease, height could be a marker of other exposures that cause disease, e.g., environmental and socioeconomic exposures [29], or genetic traits [36]. Height could also be a cause of disease directly through some biological pathways [37, 38]. Importantly, it is unknown how and why height correlates with Sami ethnicity, a primarily sociocultural construct, and the implication of this in epidemiology. Because adult height is a non-modifiable risk factor, it is of limited use in public health. Efforts should be made to understand why adult height is an explanatory variable of ethnic health differences and the public health implication of this.

Concluding remarks

Sami ethnicity may be referred to as the black box in epidemiologic research unless it is clearly defined what Sami ethnicity entails, or what explanatory factors or markers of these, are causing differences in health parameters between Sami and the comparison group. In this article, we have argued that for better interpretability of study results, researchers should aim to express which components or characteristics are assumed to comprise Sami ethnicity. Moreover, a large body of literature shows that low height is associated with Sami ethnicity, and that low height is both an important risk factor and an explanatory variable in studies of chronic lifestyle diseases in Sami vs non-Sami populations. However, similar to ethnicity, height is a non-modifiable variable that ideally should be dissected further to produce meaningful results for public health. Because height differs, on average, in Sami and non-Sami, we have argued that researchers need to be aware of the potential for bias when studying health measures influenced by height, such as the popular body mass index.

Referenser

1. Porta M. *A dictionary of epidemiology*. Sixth ed. Oxford: Oxford University Press, 2014.

2. Kaplan JB and Bennett T. Use of race and ethnicity in biomedical publication. *JAMA* 2003;289(20):2709-16. doi:https://doi.org/10.1001/jama.289.20.2709.

3. Bhopal RS. *Migration, Ethnicity, Race, and Health in Multicultural Societies*. Second ed. Oxford: Oxford University Press, 2014.

4. Siri SRA, Eliassen BM, Broderstad AR, et al. Coronary heart disease and stroke in the Sami and non-Sami populations in rural Northern and Mid Norway—the SAMINOR Study. *Open Heart* 2020;7(1):e001213. doi:http://dx.doi.org/10.1136/openhrt-2019-001213.

5. Petrenya N, Skeie G, Melhus M, et al. Food in rural northern Norway in relation to Sami ethnicity: the SAMINOR 2 Clinical Survey. *Public Health Nutr* 2018;21(14):2665-77. doi:https://doi.org/10.1017/S1368980018001374.

6. Hansen KL, Melhus M, Høgmo A, et al. Ethnic discrimination and bullying in the Sami and non-Sami populations in Norway: the SAMINOR study. Int J Circumpolar Health 2008;67(1): 97-113. doi: https://doi.org/10.3402/ijch.v67i1.18243

7. Njølstad I, Arnesen E and Lund-Larsen PG. Body height, cardiovascular risk factors, and risk of stroke in middle-aged men and women. A 14-year follow-up of the Finnmark study. *Circulation* 1996;94(11):2877-82. doi:https://doi.org/10.1161/01.CIR.94.11.2877.

8. Friborg O, Sørlie T, Schei B, et al. Do Childhood Boarding School Experiences Predict Health, Well-Being and Disability Pension in Adults? A SAMINOR Study. *J Cross Cult Psychol* 2020;51(10):848-75. doi:https://doi.org/10.1177/0022022120962571.

9. Minde H. Assimilation of the Sami – Implementation and Consequences[1]. *Acta Boreal* 2003;20(2):121-46. doi:http://dx.doi.org/10.1080/08003830310002877.

10. Lovdata [internet] Oslo,Norway: Stiftelsen lovdata [Lovdata Foundation]; 2008. Sameloven 1989 (2008-06-27 nr. 51) [The Sámi Act].(Updated; Accessed 2018 December 4th). Available at: Available at: https://lovdata.no/dokument/NL/lov/1987-06-12-56/KAPITTEL_1#KAPITTEL_1. [Norwegian]

11. Evjen B. Measuring heads: Physical anthropological research in North Norway. *Acta Borealia* 1997;14(2):3-30. doi:https://doi.org/10.1080/08003839708580465.

12. Baglo C. From universal homogeneity to essential heterogeneity: On the visual construction of "the Lappish race". *Acta Borealia* 2001;18(2):23-39. doi:https://doi.org/10.1080/08003830108580524.

13. Johnson MRD, Bhopal RS, Ingleby JD, et al. A glossary for the first World Congress on Migration, Ethnicity, Race and Health. *Public Health* 2019;172:85-88. doi:https://doi.org/10.1016/j.puhe.2019.05.001.

14. Kagawa Singer M, Dressler W and George S. Culture: The missing link in health research. *Soc Sci Med* 2016;170:237-46. doi:10.1016/j.socscimed.2016.07.015.

15. Fuentes A, Ackermann RR, Athreya S, et al. AAPA Statement on Race and Racism. *Am J Phys Anthropol* 2019;169(3):400-02. doi:https://doi.org/10.1002/ajpa.23882.

16. Maglo KN, Mersha TB and Martin LJ. Population Genomics and the Statistical Values of Race: An Interdisciplinary Perspective on the Biological Classification of Human Populations and Implications for Clinical Genetic Epidemiological Research. *Front Genet* 2016;7:22-22. doi:https://doi.org/10.3389/fgene.2016.00022.

17. Melhus M and Broderstad A. *Folkehelseundersøkelsen i Troms og Finnmark: Tilleggsrapport om samisk og kvensk/norskfinsk befolkning [The public health survey in Troms and Finnmark: Additional report on the Sami and Kven / Norwegian-Finnish population]*. Tromsø: UiT the Arctic University of Norwar; 2020. Available at: https://www.tffk.no/_f/p1/i2c5ac7f4-6b0d-485d-96d2-a68d9de030c2/rapport_troms_finnmark_sshf_redigert_april2020.pdf. [Norwegian]

18. Broderstad AR, Hansen S and Melhus M. The second clinical survey of the population-based study on health and living conditions in regions with Sami and Norwegian populations – the SAMINOR 2 Clinical Survey: Performing indigenous health research in a multiethnic landscape. *Scand J Public Health* 2019;48(6):583-93. doi:https://doi.org/10.1177/1403494819845574.

19. Brustad M, Hansen KL, Broderstad AR, et al. A population-based study on health and living conditions in areas with mixed Sami and Norwegian settlements – the SAMINOR 2 questionnaire study. *Int J Circumpolar Health* 2014;73(1):23147. doi:https://doi.org/10.3402/ijch.v73.23147.

20. Lund E, Melhus M, Hansen KL, et al. Population based study of health and living conditions in areas with both Sami and Norwegian populations-the saminor study. *Int J Circumpolar Health* 2007;66(2):113-28. doi:https://doi.org/10.3402/ijch.v66i2.18241.

21. Pettersen T and Brustad M. Which Sami? Sami inclusion criteria in population-based studies of Sami health and living conditions in Norway - an exploratory study exemplified with data from the SAMINOR study. *Int J Circumpolar Health* 2013;72:21813. doi:https://doi.org/10.3402/ijch.v72i0.21813.

22. Kaufman JS. Epidemiologic analysis of racial/ethnic disparities: Some fundamental issues and a cautionary example. *Soc Sci Med* 2008;66(8):1659-69. doi:https://doi.org/10.1016/j.socscimed.2007.11.046.

23. VanderWeele TJ and Robinson WR. On the causal interpretation of race in regressions adjusting for confounding and mediating variables. *Epidemiology* 2014;25(4):473-84. doi:https://doi.org/10.1097/EDE.0000000000000105.

24. Greenland S, Pearl J and Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10(1):37-48.

25. Jelenkovic A, Sund R, Hur YM, et al. Genetic and environmental influences on height from infancy to early adulthood: An individual-based pooled analysis of 45 twin cohorts. *Sci Rep* 2016;6:28496. doi:https://doi.org/10.1038/srep28496.

26. Batty GD, Shipley MJ, Gunnell D, et al. Height, wealth, and health: an overview with new data from three longitudinal studies. *Econ Hum Biol* 2009;7(2):137-52. doi:https://doi.org/10.1016/j.ehb.2009.06.004.

27. Stinson S. Growth Variation: Biological and Cultural Factors. Hoboken, NJ, USA: Hoboken, NJ, USA: John Wiley & Sons, Inc, 2012, p. 587-635.

28. NCD Risk Factor Collaboration (NCD-RisC). A century of trends in adult human height. *eLife* 2016;5:e13410. doi:https://doi.org/10.7554/eLife.13410.001.

29. Perkins JM, Subramanian SV, Davey Smith G, et al. Adult height, nutrition, and population health. *Nutr Rev* 2016;74(3):149-65. doi:https://doi.org/10.1093/nutrit/nuv105.

30. Tverdal A. Cohort study of ethnic group and cardiovascular and total mortality over 15 years. *J Clin Epidemiol* 1997;50(6):719-23. doi:https://doi.org/10.1016/S0895-4356(97)00021-8.

31. Laitinen J, Näyhä S, Sikkilä K, et al. Diet and cardiovascular risk factors among Lapp and Finnish reindeer herders. *Nutr Res* 1996;16(7):1083-93. doi:https://doi.org/10.1016/0271-5317(96)00113-3.

32. Edin-Liljegren A, Hassler S, Sjölander P, et al. Risk factors for cardiovascular diseases among Swedish Sami-a controlled cohort study. *Int J Circumpolar Health Suppl* 2004;63:292-97. doi:https://doi.org/10.3402/ijch.v63i0.17922.

33. Michalsen VL, Braaten T, Kvaløy K, et al. Relationships between metabolic markers and obesity measures in two populations that differ in stature—The SAMINOR Study. *Obes Sci Pract* 2020;6:324-39. doi:https://doi.org/10.1002/osp4.404.

34. Naseribafrouei A, Eliassen B-M, Melhus M, et al. Prevalence of pre-diabetes and type 2 diabetes mellitus among Sami and non-Sami men and women in Northern Norway – The SAMINOR 2 Clinical Survey. *Int J Circumpolar Health* 2018;77(1):1463786. doi:https://doi.org/10.1080/22423982.2018.1463786.

35. Nystad T, Melhus M, Brustad M, et al. Ethnic differences in the prevalence of general and central obesity among the Sami and Norwegian populations: the SAMINOR study. *Scand J Public Health* 2010;38(1):17-24. doi:https://doi.org/10.1177/1403494809354791.

36. Nelson CP, Hamby SE, Saleheen D, et al. Genetically determined height and coronary artery disease. *N Engl J Med* 2015;372(17):1608-18. doi:https://doi.org/10.1056/NEJMoa1404881.

37. Marouli E, Del Greco MF, Astley CM, et al. Mendelian randomisation analyses find pulmonary factors mediate the effect of height on coronary artery disease. *Commun Biol* 2019;2:119. doi:https://doi.org/10.1038/s42003-019-0361-2.

38. Lai FY, Nath M, Hamby SE, et al. Adult height and risk of 50 diseases: a combined epidemiological and genetic analysis. *BMC Med* 2018;16(1):187. doi:https://doi.org/10.1186/s12916-018-1175-7.