The relationship among the age at menarche, anthropometric characteristics, and socio-economic factors in Bengali girls from Kolkata, India

Magdalena Żeglen¹ | Elisabetta Marini² | Stefano Cabras² | Łukasz Kryst¹ | Rituparna Das³ | Anindita Chakraborty³ | Parasmani Dasgupta³

1Department of Anthropology, Faculty of Physical Education, University of Physical Education in Kraków, Kraków, Poland
2Department of Life and Environmental Sciences, University of Cagliari, Cagliari, Italy
3Biological Anthropology Unit, Indian Statistical Institute, Kolkata, India

Abstract

Objectives: The aim of the study was to measure the causal effect of selected socio-economic factors and anthropometrical characteristics on the menarche occurrence.

Methods: The sample consisted of 2195 Bengali girls (aged 7-21) from middle-class families, from Kolkata city, India. The age at menarche was recorded from the retrospective data and status quo method. The causal effect of anthropometric and socio-economic variables on menarche occurrence was estimated by the nonparametrical analysis of survival probability (survival random forest).

Results: In the examined cohort menarche occurred, on average, at 11.8 years of age. The probability of menarche occurrence amplified with the increasing values of factors such as body mass index, height-for-age z-scores, number of family members, household rooms, and toilets, but decreased when expenditures increased. The relation maintained a similar pattern of causal effect with girls' age.

Conclusions: A complex pattern of relationship among sexual development, physique, and socio-economic characteristics was defined. The tendency toward early menarche, along with the observed causal relationships indicate that the analyzed sample is nearing the characteristics and standards of living noted in other middle and even high-income countries in the world.

1 | INTRODUCTION

The age of the biological maturation is vital for both the physical and psychological health of girls and young women. In fact, too early and too late onset of biological maturation can have negative effects. The premature occurrence of menarche is associated with a higher risk of obesity, breast and ovarian cancer, anxiety, depression, as well as metabolic syndrome (WHO, 2000). On the other hand, the too late onset of menarche is correlated with a higher risk of, once again, metabolic syndrome, but also oligomenorrhea and cardiometabolic abnormalities (Glueck, Morrison, Wang, & Woo, 2013). The age of menarche occurrence is influenced by many factors. Among others, they include woman's genetic background, as in the human genome, there are 10 single nucleotide polymorphisms, located in two loci, which have a significant impact on the onset of menarche (He et al., 2009). In addition, later studies suggest the presence of even 30 to 40 more loci containing genes with an analogous function (Elks et al., 2010).

Many of the genes that are crucial in regulating the onset of menarche are also significant for the hormonal regulation, energy homeostasis, as well as the determination of body height or weight (Elks et al., 2010). Moreover, it has been suggested that the height and weight of...
the body, as well as their mutual proportions also pose important constraints (Anderson & Must, 2005; Fredriks et al., 2005; Gohlke & Woelfle, 2009; Talma et al., 2013). Because of that, the so-called “critical weight theory,” where a certain body mass must be gained for the menarche to occur, was proposed (Frisch & Revelle, 1970). However, in later years, it was suggested that the presence of leptin, as well as the other hormones related to energy management, can also have a decisive role (Fernandez-Fernandez et al., 2006; De Bond & Smith, 2014). Due to the described phenomena, girls with higher weight and body adiposity generally experience menarche earlier than their leaner peers (Chowdhury et al., 2000; Khatoon et al., 2011).

On the other hand, the moment of menarche occurrence can influence some of the anthropometric features. In fact, it has been proven that girls who experience earlier menarche reach a shorter adult height (Gharravi, Marjani, & Golalipour, 2008). It is also generally suggested that the earlier biological maturation is correlated with a higher subsequent body mass index (BMI; Laitinen, Power, & Järvelin, 2001; Parsons, Power, Logan, & Summerbell, 1999; Pierce, Kuh, & Hardy, 2012; Pierce & Leon, 2005). Due to its close association with the growth process menarche status is considered more informative than anagrophic age about the body composition changes and biological development in adolescent girls (Buffa, Floris, & Marini, 2002; Buffa, Floris, Marini, & Ortu, 2003; Koury, Ribeiro, Massarani, Vieira, & Marini, 2019).

Other factors affecting the onset of menstruation include the socio-economic environment, such as, for example, the social class (Bagga & Kulkarni, 2000; Wronka & Pawlińska-Chmara, 2005). Moreover, family size, or living conditions, among others, can be distinguished as crucial factors influencing the age of biological maturation (Khalid et al., 2015). Due to its recognized sensitivity to environmental factors, age at menarche can be considered an indicator of changing socio-economic conditions. Therefore, it is a common subject of research concerning intergenerational changes. The negative secular trend regarding the age at menarche was observed in many countries. This phenomenon is one of the effects of the socio-economic progress, significant especially in developing countries, such as India, as acceleration of biological maturation is correlated with the improvement of the living conditions of people (Cole, 2003; International Statistical Institute, 2017; Kowal, Kryst, Woronkowicz, & Sobiecki, 2014; Das et al., 2016; Kryst, Kowal, Woronkowicz, Sobiecki, & Cichocka, 2012; Meng, Li, Duan, Sun, & Jia, 2017).

The aim of the study was to measure the causal effect of the selected socio-economic factors and anthropometrical characteristics on the menarche occurrence in a large sample of middle-class, Bengali girls.

2 | MATERIALS AND METHODS

2.1 | The sample

A cross-sectional sample of 2195 girls of Bengali population aged 7 to 21 years (Table 1), from 66 schools and colleges from Kolkata city, India, has been collected during one survey period (2005-2011). Girls' calendar age was calculated as a difference between the date of the survey and the birth date, which was expressed as a decimal fraction. Each child was then classified into one of the age categories on the basis of the said calendar age (eg, the group of 7 year-olds included children whose calendar age ranged from 6.50 to 7.49).

The girls came predominantly from middle-class families, classified as such on the basis of:

- monthly family expenditure—≥2200 INR;
- parental occupation—majority of fathers working in business field;
- parental education—majority had at least graduate degree;
- household assets and housing condition—included for example the area occupied by the household, number of rooms (≥in majority of the cases and bathrooms (at least 1) in the house as well as the source of drinking water (own tap/well in majority of the households).

<p>| TABLE 1 | Age distribution of the sample of 2195 girls of Bengali population |</p>
<table>
<thead>
<tr>
<th>Age group (y)</th>
<th>Number of individuals (N)</th>
<th>N in which menarche occurred</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>48</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>159</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>165</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>153</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>147</td>
<td>19</td>
</tr>
<tr>
<td>11</td>
<td>141</td>
<td>61</td>
</tr>
<tr>
<td>12</td>
<td>162</td>
<td>124</td>
</tr>
<tr>
<td>13</td>
<td>113</td>
<td>104</td>
</tr>
<tr>
<td>14</td>
<td>149</td>
<td>144</td>
</tr>
<tr>
<td>15</td>
<td>137</td>
<td>137</td>
</tr>
<tr>
<td>16</td>
<td>160</td>
<td>160</td>
</tr>
<tr>
<td>17</td>
<td>131</td>
<td>131</td>
</tr>
<tr>
<td>18</td>
<td>130</td>
<td>130</td>
</tr>
<tr>
<td>19</td>
<td>131</td>
<td>131</td>
</tr>
<tr>
<td>20</td>
<td>183</td>
<td>183</td>
</tr>
<tr>
<td>21</td>
<td>86</td>
<td>86</td>
</tr>
<tr>
<td>Total</td>
<td>2195</td>
<td>1420</td>
</tr>
</tbody>
</table>
The qualification to the research group was based on the good overall health of the girls and the consent of their parents or legal guardians. All procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

2.2 | The variables

The analyzed anthropometric characteristics included body height and weight, measured according to the protocol of the International Biological Program (Weiner & Lourie, 1969), using an anthropometer (GPM, Switzerland; accuracy 1 mm) and electric weight (Libra, India; accuracy 0.1 kg). Height was expressed as height-for-age z-scores (HAZ), calculated according to WHO standards (de Onis et al., 2007) and BMI (body weight [kg]/body height [m²] and as BMI z-scores [BAZ]). The analyzed socioeconomic characteristics of children’s families included: number of family members, sibship size and parity, family type, parents’ occupation and education, family expenditure, caste, number of rooms and toilets in the house, household duration, house possession, as well as drinking water source distance and adequacy.

3 | STATISTICAL ANALYSIS

3.1 | Age of menarche

The status quo and retrospective recall methods have been applied. The status quo method allows defining the “current status” regarding the birth date and menarche occurrence at the time of assessment (yes or not) by asking a girl (or her parent). The retrospective recall method allows the definition of menarche age by asking postmenarchal females (or their parents) to recall the date of first menses.

Age at menarche in the whole sample has been evaluated by means of a survival curve estimated from the Kaplan–Meier estimator (Klein & Moeschberger, 2003).

3.2 | Influential factors of the age of menarche

The causal effect of socio-economic and anthropometric variables on the age of menarche was analyzed using the survival random forest (SRF) model (Breiman, 2001; Mogensen, Ishwaran, & Gerds, 2012). SRF is a nonparametric survival regression method suitable for large datasets, as in this research, in which no linear predictor (ie, all variables along with all their interactions) has to be specified beforehand and where all possible interactions among all considered confounding factors can be accounted for. Hence, the analysis represents a significant improvement with respect to ordinary regression approaches, either for modeling or for estimating causal effects. In particular, as shown by Lu, Sadiq, Feaster, and Ishwaran (2018), SRF allows to measure causal effects also in cross-sectional samples. Such inference on the individual causal effect is obtained by means of the difference between the factual observed age at menarche (for the group of girls with menarche) and the counterfactual value (age at menarche estimated using SRF and variable values of the anthropometric or socio-economic variables). This difference represents an estimate of the individual causal effect due to the anthropometric or a socio-economic intervention variable. The means of such differences are used to estimate overall causal effects (for more details: Gelman & Meng, 2004). In this research, the analysis has been based on the survival curve of age at menarche in the whole sample. Causal effects have been estimated by assuming that the SRF model is able to predict the age of menarche, under a large set of possible causing variables. The possible error in menarche age due to the recall method has been accounted for by introducing the girls’ age (assumed to be related to retrospective data) as a confounding effect.

4 | RESULTS

4.1 | Age of menarche

The majority of girls in the sample (1420; 64.7%) had experienced menarche at the time of data collection: 715 girls recalled the age of menarche with an approximation no larger than 1 month (day, month, and year, or month and year), 22 with an approximation no larger than 6 months (1-6 months), 195 remembered the year only, and 488 did not mention any date (Table 1).

The estimated median age at menarche of girls calculated from status quo data by Probit analysis was 11.6 years. Considering retrospective data of girls who remembered the exact age with an error lower than 1 month (day, month, and year, or month and year), 22 with an approximation no larger than 6 months (1-6 months), 195 remembered the year only, and 488 did not mention any date (Table 1).

The estimated median age at menarche of girls calculated from status quo data by Probit analysis was 11.6 years. Considering retrospective data of girls who remembered the exact age with an error lower than 1 month (N = 715) the mean age of menarche was 11.8 ± 1.2.

Figure 1 represents the empirical marginal survival curve up to menarche event, that is, the probability of not having menarche at a given age.

4.2 | Influential factors on age of menarche

The model estimating the causal effect on age of menarche occurrence due to the specified anthropometrical
(BMI, BAZ, HAZ) and socio-economic variables (family expenditure, family members, sibship size, parity, family type, household rooms, household toilets, household duration, parents’ occupation and education, caste, house possession, drinking water source, distance, and adequacy) was quite explicative (variance explained of the above survival curve: 67.1%).

Figure 2 represents the possible causal effect of selected factors on age of menarche occurrence, that is, the proportion of girls in which menarche occurred according to factors variation, given fixed the effect of all other covariates (assumed as confounders). The age of 12 years, near the mean age of menarche, was selected to show these results. As can be seen, the probability of menarche occurrence at age 12 significantly increased (the survival, that is, the proportion of girls without menarche, decreases) with increasing values of BMI, BAZ, HAZ, number of family members, household rooms, but decreased when expenditures increases, while household duration did not show a significant effect. In particular, for example, the probability of menarche occurrence shifted from around 0.31 to 0.40, when BMI values changes from 15 to 27, indicating that higher values of BMI induce more precocious menarche. The probability also increased with the number of household toilets, sibship size, and parity, and decreased in case of household possession and adequate drinking source (Figure 3). Furthermore, a higher probability of menarche occurrence was perceived in girls whose parents attended a college, whose fathers were retired, who pertained to the “other backward class” caste, while the probability was lower among girls whose parents were graduated and whose mothers were working in business areas. Household duration (Figure 2), family type, and drinking water source distance seemed not to have a significant role in age of menarche occurrence in this sample of Indian girls.

**FIGURE 1** Curve of empirical survival probability up to menarche occurrence

**FIGURE 2** Survival probability up to menarche occurrence according to different causal factors. Vertical axis is the estimated survival probability up to menarche onset at 12 years (ie, the higher, the less probable is menarche onset)
Interestingly, the relations maintained a similar pattern of causal effect with girls’ age, even in some cases changing the scale of survival probability. For example, considering BMI, the peak of greater probability of menarche occurrence corresponded to a BMI = 26 in all groups of age, but increased its values with the increase of girls’ age, shifting from 14.5% at the age of 11 years to 63% at the age of 13 years (Figure 4). This means that an overweight or obese girl has a higher probability (up to 63% at the age of 13 years) of experimenting menarche than a normal weight one. Considering the whole sample, and the case of girls with 12 years as an example, the girls with a BMI lower than 26 experiment menarche with a probability comprised between 28% and 40%, while those with higher BMI values with an almost steady probability of 40%.

5 | DISCUSSION

The results of the study showed that the probability of menarche occurrence at a certain age in a sample of Bengali girls is significantly influenced by the anthropometric features (height and BMI), as well as by some socioeconomic factors (economic status, family, and household characteristics). The observed pattern of probability remains very similar in all age groups, even if the probability values change.

5.1 | The average age at menarche

In the examined cohort menarche occurred, on average, at 11.8 years of age. A similar result was obtained in another recent study of the West Bengali population, where the mean age at menarche was 11.9 years (De & Bose, 2016). The observed average age of the first menstruation is also similar to those detected in other countries. For example, among the Korean middle and high-school girls the mean age at menarche was 11.9, 11.6 in the girls from Bangladesh, and 11.4 years for the ones coming from Mexico (Han et al., 2016, Islam et al., 2017, Marván, Catillo-López, Alcalá-Herrera, & del Callejo, 2016). Similarly, in Portugal, it was 12 years of age, 12.1

In the same time, the mean age at menarche observed in the current study is lower than those detected in the previous researches carried out in the Indian population, thus suggesting an ongoing, negative secular trend regarding menarche age. In 2005, Pathak, Tripathi, and Subramanian (2014) examined 91,394 Indian females, aged 15 to 49. Among them, the average age at menarche was 13.8, for the ones born between 1955 and 1964 and 13.6 for the girls born from 1985 to 1989. A similar mean age (13.5 years) of the first menstruation was observed in the girls from a rural area in the Amravati district of Maharashtra, examined in 2013-14 (Wasnik et al., 2015).

The negative intergenerational changes in the age of menarche of Indian girls were also confirmed by John, Verma, and Chhatwal (2014) in their study on 400 middle-school Punjabi girls and their mothers. The average age at menarche was 14.3 for mothers when for daughters it was already 12.4 years of age.

The trend toward lower mean age at menarche observed in the present study is even strengthened considering that our result could represent an overestimate of the true mean age, due the possible error of retrospective data (Salehabadi, Sengupta, & Das, 2015).

5.2 Influential factors on age of menarche

The presently observed later menarche due to higher height is in line with the results of a previous study on the Indian population, in which taller girls experienced menarche later than their shorter counterparts (Bagga & Kulkarni, 2000; John et al., 2014; Khadgawat et al., 2016). There is also a vast amount of international literature reporting that taller girls reach their menarche at a relatively earlier age (Elks et al., 2010; Koprowski, Ross, Mack, Henderson, & Bernstein, 1999; Lin et al., 2016; Onland-Moret et al., 2005; Relethford & Lees, 1980; Shangold, Kelly, Berkeley, Freedman, & Groshen, 1989). It was also suggested that not only body height, but even more so the moment of the height spurt has a crucial influence on the timing of menarche (Chang, Tzeng, Cheng, & Chie, 2000).

The mutual proportion of body height and mass also had an influence on the age of menarche. It is in line with another study where body mass, alongside nutritional status, was even distinguished as one of the factors with the strongest influence on the onset of menarche (Batuabara, Soesanti, & van de Waal, 2010). In the present study, a higher probability of a relatively earlier onset of menarche was noted in girls with a higher BMI. This is consistent with the results of other studies carried out in the Indian population. The same relationship was also observed in girls from Kuwait and Germany, Italy and Chile (Bau et al., 2009, Al-Awadhi et al., 2013). However, the exact physiological mechanism linking BMI to menarche remains unclear. It is suggested that high BMI values in the prepubertal period could lead to increased estrogen production, which in turn predisposes to an earlier occurrence of menarche (Cheng et al., 2012).

The relation between BMI, as well as other causal factors, and menarche occurrence did not change with girls' age, but the probability of menarche become stronger in higher age groups, as expected.

Besides the anthropometric characteristics, the age at menarche can also be influenced by the socio-economic environment in which the girls are growing up (Khalid et al., 2015; Lee, Kim, Oh, Lee, & Park, 2016).

In the present study, a lower probability of menarche occurrence was observed with family expenditures increase, and with higher socio-economic, cultural, or working levels. Similarly, other studies on the Indian population noticed that females from more affluent families experience menarche later than their poorer counterparts (Pathak et al., 2014). There, the authors suggested that girls with a higher economic status were more likely to care about their figure and therefore take part in more intense physical activities, which could have reduced their body mass and shifted the age of their menarche to the upper thresholds (Pathak et al., 2014; Rees, 1993). A similar association was observed in the American population, where the relatively worse financial situation of the family was associated uniquely with earlier menarche (Deardorff, Abrams, Ekwaru, & Rehkopf, 2014). In turn, in Poland, the increase in expenses due to the rising food prices in the 1960s was correlated with a relatively later onset of menarche (Gomula & Koziel, 2018). However, opposite results have been obtained in other researches previously conducted in the Indian population, where the girls coming from wealthier families experienced menarche at an earlier age, compared to those growing up in less favorable economic conditions (Bagga & Kulkarni, 2000). Similarly, in Pakistani research, girls from the highest social class experienced menarche earlier, compared to their lower-class counterparts (Ahmed, Waheed, Ahmad, & Danish, 2016). The different direction of the association between expenditure and age of menarche observed in different periods or countries could be interpreted in relation to BMI variability and to the causal effect of BMI on menarche onset. In fact, assuming a negative relation between BMI and menarche age, it is to be noted that the association between BMI
and socio-economic level is different in low-income (LI) and high-income (HI) countries. In fact, in LI countries the higher BMI means values are present in people from the higher economic classes (where girls could experience menarche earlier), while in HI countries in people from the lower classes (Pampel et al., 2012). In the present study, India would perform in the same way as HI countries. This hypothesis is confirmed by the results on the negative secular trend of menarche age observed in this study and similar to other countries, where was associated with socio-economic development (Lee et al., 2016; Meng et al., 2017).

However, in the present sample, not all socio-economic factors affected the probability of menarche. In fact, household duration, or drinking water source distance did not have any significant causal effect. Similarly, among girls from Bangladesh (Islam et al., 2017) and Iran (Gharravi et al., 2008) almost no effect of socio-economic conditions on the age of menarche was observed. At the same time, in the study of the American population, parent’s occupation was one of the important factors, differentiating the onset of menarche in the studied group.

Furthermore, in contrast with the results on family expenditure, a relatively greater chance of menarche occurrence at a given age was related to the number of household’s rooms and toilets. Similarly, in Chinese studies, it has been proven that a better standard of living is related to the relatively earlier occurrence of the first menstruation (Lee et al., 2007; Wang, Dang, Xing, Li, & Yan, 2016). In the present study, the probability of menarche occurrence increased with the number of family members, sibship size, and parity of mothers, but family type did not show a significant effect. Similar results have been obtained by Khalid et al. (2015), regarding the number of family members, and by Wronka and Pawlińska-Chmara (2005), regarding the number of children in the family. However, a study on girls from Uganda and Nigeria showed that the number of siblings, as well as the order of births, did not have an effect on the moment of menarche (Ikaraoha et al., 2005; Odongkara Mpora et al., 2014).

In our study, girls from the “other backward class” had a higher probability of the onset of menarche at a certain age. In a previous study carried out on the Indian population, there was a statistically significant difference in the mean age at menarche between each caste—women from Forward castes statistically had their first menstruation later than their counterparts from Scheduled caste groups (Pathak et al., 2014). Also, another study of the Indian population has shown significant differences regarding age at menarche between each analyzed caste. However, girls from the Scheduled caste had menarche relatively later than their counterparts from the other caste groups (Bagga & Kulkarni, 2000).

### 5.3 Strengths and limitations of the study

The present study had some limitations, one of them being the use of the recall method for determining the age at menarche. According to some authors, it is possible for it not to be precise enough, due to the time interval between the menarche and the date of contact (John et al., 2014), or to represent an overestimate of the true age (Salehabadi et al., 2015). However, other studies have shown that such method of obtaining the information about the age at menarche can be sufficiently accurate (Koprowski, Coates, & Bernstein, 2001; Must et al., 2002; Pierce & Leon, 2005). Indeed, the parametric modelization realized by Salehabadi and Sengupta (2018) in a similar dataset has shown that, depending on the statistical model used, the recall effect can result irrelevant.

Furthermore, in our research, the inclusion of girls’ age in the regression model has permitted to account for the possible error due to recall data. The study also has significant strengths, such as the quite large study sample. Furthermore, body height and weight were measured by skilled operators and not recalled by the participants. It is important because the individuals with higher body mass in proportion to height tend to have underestimated BMIs based on self-reported measurements (Lawlor, Bedford, Taylor, & Ebrahim, 2002; Pierce & Leon, 2005). A major strength of the research is the statistical approach, which used a class of models much wider than those usually considered by regression approaches, thus improving the estimation of causal effects. Indeed, despite the cross sectional nature of the sample of Bengali girls, the Random Forest analysis allowed to measure the causal effects on the age at menarche, and not mere associations between variables as generally done.

### 6 Conclusions

In the analyzed sample of Bengali girls, a complex pattern of relationship between sexual development, physique, and socio-economic characteristics was defined. The probability of menarche occurrence was significantly influenced by the anthropometric features, as well as some of the socioeconomic variables. Although the observed pattern of probability remained very similar in all age groups, the probability changed with age.

The observed tendency toward earlier menarche, along with some of the observed causal relationships,
suggests that this sample of Indian population is performing similarly to girls from other middle and HI countries in the world.

ACKNOWLEDGMENTS
This study has been sponsored by the Neys van Hoogstraten Foundation, The Netherlands (ID158) and Indian Statistical Institute, Kolkata, India.

CONFLICT OF INTEREST
The authors declare no potential conflict of interest.

AUTHOR CONTRIBUTIONS
R.D., P.D., and A.C. carried out the field part of the research and collected the data. E.M. conceived of the presented idea with support from M.Z. S.C. performed the statistical analysis and prepared the figures with support from E.M. E.M., L.K., and M.Z. wrote the manuscript with support from S.C. E.M., L.K., and M.Z. collected the literature. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

ORCID
Magdalena Żegleń https://orcid.org/0000-0001-6028-8905
Elisabetta Marini https://orcid.org/0000-0001-8779-8745
Łukasz Kryst https://orcid.org/0000-0001-6938-192X

REFERENCES


How to cite this article: Żeglen M, Marini E, Cabras S, et al. The relationship among the age at menarche, anthropometric characteristics, and socio-economic factors in Bengali girls from Kolkata, India. Am J Hum Biol. 2019;e23380. https://doi.org/10.1002/ajhb.23380