Pathways to Early Coital Debut for Adolescent Girls: A Recursive Partitioning Analysis

Matthew R. Pearson, Tatyan Kholodkov, and James M. Henson
Department of Psychology, Old Dominion University

Emily A. Impett
Department of Psychology, University of Toronto Mississauga

This study examined pathways to early coital debut among early to middle adolescent girls in the United States. In a two-year longitudinal study of 104 adolescent girls, recursive partitioning (RP) analyses were conducted to examine the specific factors that were related to engaging in first intercourse by the 10th grade among adolescent girls who had not yet engaged in sexual intercourse by the 8th grade. RP analyses identified subsamples of girls who had low, medium, and high likelihoods of engaging in early coital debut based on six variables (i.e., school aspirations, early physical intimacy experiences, depression, body objectification, body image, and relationship inauthenticity). For example, girls in the lowest likelihood group (3% had engaged in sex by the 10th grade) reported no prior experiences with being touched under their clothes, low body objectification, high aspirations to complete graduate education, high aspirations to complete graduate education, and low depressive symptoms; girls in the highest likelihood group (75% had engaged in sex by the 10th grade) also reported no prior experiences with being touched under their clothes, but had high levels of body objectification. The implications of these analyses for the development of female adolescent sexuality, as well as for advances in quantitative methods, are discussed.

Despite continued efforts to educate the public, troubling trends are emerging in the rates of sexually transmitted infections (STIs) and unwanted pregnancies among young people in the United States. In 2006, it was estimated that about one million adolescents and young adults in the United States were living with chlamydia, gonorrhea, or syphilis (Centers for Disease Control [CDC], 2009). Adolescent girls and young adult women are particularly vulnerable to contracting certain types of STIs. An estimated 24.5% of adolescent girls 15 to 19 years old and 44.8% of those in the 20 to 24 age group are believed to be infected with the human papillomavirus (CDC, 2009). Further, chlamydia rates are almost three times as high among women than men, and women are more likely than men to contract gonorrhea (CDC, 2007). Furthermore, girls and women bear the primary burden of unwanted pregnancies stemming from unprotected sex. Early coital debut has been shown to be a risk factor for having multiple sexual partners (World Health Organization, 2006) and HIV infection (Drain, Smith, Hughes, Halperin, & Holmes, 2004). In addition, early onset of sexual intercourse in girls is associated with more problematic behavior, such as minor deviance (Bingham & Crockett, 1996), failure to attain a high school diploma and post-secondary education (Frisco, 2008), and may also be associated with complications in sexual health in adulthood (Magnusson & Trost, 2006). Thus, it is important to identify pathways of adolescent girls’ transition to first intercourse.

Predictors of Early Coital Debut

Researchers have demonstrated a number of a priori hypothesized factors that relate to the timing of adolescent girls’ first sexual intercourse experience. In this study, we complement this previous research by exploring a broad set of potential predictors of coital debut, including demographic variables, school-related variables, religiosity, mental health, experiences and...
comfort with physical intimacy, and femininity ideology constructs. In previous research, these variables have been examined piecemeal, such that only a few hypothesized variables were examined in any particular study. In contrast, this study builds on that prior work in order to examine all of these variables simultaneously by taking advantage of an exploratory statistical technique called recursive partitioning (RP; Zhang & Singer, 1999). RP offers a unique perspective to research on sexuality because it enables researchers to examine a large number of potential predictors of coital debut while simultaneously identifying which factors are the strongest predictors.

**Demographics**

In terms of demographic variables, existing research that has examined racial and ethnic differences in coital debut has revealed mixed findings. With regard to differences according to race and ethnicity, several studies indicate that African American girls are at the highest risk of early coital debut (Carvazos-Rehg et al., 2010; Feldman, Turner, & Araujo, 1999; Regan, Durvasula, Howell, Ureño, & Rea, 2004). However, other research indicates no significant differences between White, Black, and Asian girls living in an urban city (Upchurch, Aneshensel, Sucoff, & Levy-Storms, 1999), as well as no differences among White, Black, and Latina low-income adolescent girls (Jordahl & Lohman, 2009). Studies also indicate that Latina girls have intercourse experiences later than African American girls (Feldman et al., 1999; Regan et al., 2004) and White girls (Upchurch et al., 1999). Therefore, extant research on the influence of race and ethnicity on coital debut has not produced a clear set of findings.

In contrast to the effects of race and ethnicity, socioeconomic status has consistently predicted coital debut in adolescent samples. In a large Norwegian study, maternal and paternal education, as well as family income, significantly predicted early sexual debut among 15- to 16-year-old adolescent girls, such that girls with lower parental education were more likely to experience coital debut by age 15 or 16, as compared to girls with higher parental education (Valle, Roysamb, Sundby, & Klepp, 2009). This trend is consistent with other studies conducted in the United States (Longmore, Manning, Giordano, & Rudolph, 2004), Sweden (Magnusson & Trost, 2006), and New Zealand (Paul, Fitzjohn, Herbison, & Dickson, 2000).

**Educational Predictors**

Researchers have shown that both educational achievement (Voisin & Neilands, 2010) and future educational goals (Marchi & Guendelman, 1994; Valle et al., 2009) are related to a lower age of coital debut. For example, greater educational aspirations were associated with a delay in coital debut among 15- to 16-year-olds (Valle et al., 2009). Moreover, a longitudinal study of children in New Zealand conducted by Paul et al. (2000) demonstrated that plans to not finish school or being in “trouble” in school at age 13 predicted a higher likelihood of sexual initiation by age 16. Similarly, a longitudinal study of girls ages 13 to 16 by Bonell and colleagues (2005) revealed that negative attitudes toward academics were associated with an earlier age of coital debut. Among Mexican American adolescents, a delay in coital debut was also associated with familial expectations of attaining higher education (Gilliam, Berlin, Kozloski, Hernandez, & Grundy, 2007). Thus, school involvement and educational goals appear to be protective factors that can delay sexual debut, which, in turn, may also be associated with future academic success (Frisco, 2008).

**Religiosity**

Multiple studies have shown that religiosity, typically defined as the importance of religion in an individual’s life (Schwartz & Huismans, 1995), is related to an older age of transition to first intercourse (Day, 1992; Edwards, Fehring, Jarrett, & Haglund, 2008; Rostosky, Regnerus, & Wright, 2003). In a review, Rostosky, Wilcox, Wright, and Randall (2004) found consistent evidence that religiosity was related to an older age of coital debut among adolescent girls; however, the relationship between religiosity and coital debut was inconsistent for adolescent boys, such that some studies found that religiosity predicted delayed coital debut among adolescent boys whereas other studies found no such relationship.

**Mental Health Predictors**

Researchers have found that girls who experience depressive symptoms prior to their first experience with sexual intercourse have an earlier coital debut than their less depressed counterparts (Longmore et al., 2004; Paul et al., 2000; Spriggs & Halpern, 2008). Studies with regard to the link between a girl’s feelings of self-esteem and the age at which she first engages in sexual intercourse have revealed a mixed pattern of findings. Among New Zealander girls, higher self-esteem was associated with engaging in sex before the age of 16 (Paul et al., 2000). However, several studies have found that self-esteem was not associated with early coital debut, as demonstrated in a sample of Canadian girls in the 9th and 11th grades (Boyce, Gulle, & Fergus, 2008), of U.S. adolescent girls (Bingham & Crockett, 1996), and of university students recalling their first intercourse experience (Slicker, Patton, & Fuller, 2004). Some research also seems to indicate that perhaps low self-esteem is compensated for by engaging in earlier sexual activity. For example, among U.S. adolescents,
low self-esteem was found to be associated with engaging in oral sex by the 9th grade, and high self-esteem was associated with delaying sexual activity (Price & Hyde, 2009). Additionally, a study of 13-year-old Canadian girls (Brendgen, Wanner, & Vitaro, 2007) found a negative relationship between self-esteem and sexual intercourse, such that low self-esteem increased the chances that a girl would engage in her first intercourse experience by age 13.

Another important indicator of mental health, especially during a time when most adolescent girls are undergoing puberty and its ensuing bodily changes, is a girl’s satisfaction with her developing body (Tolman, Impett, Tracy, & Michael, 2004). Research has shown that body dissatisfaction is associated with early coital debut. For Norwegian 10th grade girls 15 to 16 years old, negative feelings about one’s body were associated with earlier ages of engaging in first sexual intercourse (Valle et al., 2009). Further, symptoms of bulimia were associated with more advanced sexual behavior (e.g., progressing from kissing to light or heavy petting) among Finnish 14- to 16-year-old girls (Kaltila-Heino, Rimpela, Rissanen, & Rantanen, 2001). In contrast, Schooller, Ward, Merriwether, and Caruthers (2005) reported that college student women who had a poorer body image reported engaging in less overall sexual experience than women who felt more satisfied with their bodies (Schooler et al., 2005). With a few exceptions, this literature suggests that experiencing poorer mental health, such as more depressive symptoms, reduced self-esteem, and body dissatisfaction, may increase the likelihood that adolescent girls will engage in their first experience of sexual intercourse at younger ages than girls who experience greater mental health.

**Physical Intimacy Predictors**

Researchers have found that adolescent girls’ experiences with dating (e.g., having any dating experience) and romance or physical intimacy (e.g., kissing and flirting) have also been shown to predict earlier coital debut (Davila et al., 2009; Longmore et al., 2004). It is not surprising that girls with dating or romantic experience are more likely to engage in their first experience of sexual intercourse at earlier ages than their less experienced counterparts. Various studies have found that sexual behaviors occur in a similar progression, such as kissing, French kissing, breast fondling, genital contact, and intercourse (Bingham & Crockett, 1996; Feldman et al., 1999) and that late initiation of sexual intercourse is associated with an absence of “necking” and “petting” behaviors earlier in adolescence (Bingham & Crockett, 1996).

Sexual self-acceptance is another factor relevant to physical intimacy that may relate to coital debut (Winter, 1988). Sexual self-acceptance has been defined as an individual’s overall concept or acceptance of the self as a sexual person, including both positive aspects (e.g., passion, arousability, and agency), as well as negative aspects (e.g., anxiety, shame, and embarrassment; Andersen & Cyranowski, 1994; O’Sullivan, Meyer-Bahlburg, & McKeague, 2006). Although no research has examined how a girl’s acceptance of her own sexuality relates to the age at which she transitions to first intercourse, other related research has found that girls with higher levels of sexual self-acceptance tend to find their first sexual experiences more satisfying than girls with lower levels of sexual self-acceptance (Impett & Tolman, 2006). It is unclear, however, if girls with high levels of sexual self-acceptance engage in sex earlier (because they see sexuality as a positive part of their identity) or later (because they wait until they feel fully ready and prepared).

**Femininity Ideology Predictors**

Finally, we included another set of variables that may be especially important to girls’ developing sexuality and age of coital debut. Research stemming from Deborah Tolman’s femininity ideology framework suggests that two variables are specific to girls’ sexuality and sexual experiences and may be particularly important predictors of adolescent girls’ coital debut (e.g., Tolman et al., 2004; Tolman & Porche, 2000). This perspective holds that girls are expected to behave in “feminine” ways in the context of their relationships (e.g., by avoiding conflict, suppressing anger, and being “nice”) and through their relationships with their own bodies (i.e., by managing their own bodies to conform to the prevailing images of beauty and attractiveness). These pressures translate into two central femininity ideology constructs. First, inauthenticity in relationships represents the tendency to silence one’s own authentic thoughts and feelings in relationships in order to “keep the peace,” (Impett et al., 2008) and is conceptually akin to Jack’s (1991) construct of self-silencing in adult women. Although no research has yet examined the links between inauthenticity in relationships and coital debut, related research provides some potential insights. For example, a recent daily diary study found that college-aged women who were high in relationship inauthenticity were less likely to use condoms with a dating partner, especially on days when they felt highly threatened in their relationships (Impett et al., 2010); this research is consistent with the hypothesis that inauthentic girls may find it difficult to say “no” to early sexual episodes, especially out of desires to please or appease their male partners.

A second aspect of femininity ideology is body objectification, defined as the tendency to view one’s body from the perspective of an outsider and, subsequently, being more concerned with how one’s body looks instead of how it feels. The theoretical construct of body
objectification stems from Objectification Theory, which posits that sexual objectification occurs when individuals are treated as bodies for the use and pleasure of others (Bartky, 1990; Fredrickson & Roberts, 1997) and when their value is appraised solely from their sexual appeal (American Psychological Association, 2007). Again, no research has explicitly examined the link between body objectification and coital debut; however, we can draw on other related research to derive potential predictions. One study of late adolescent girls revealed that girls who objectified their bodies reported using condoms less consistently by the 12th grade than more embodied girls (Impett, Schooler, & Tolman, 2006). In a qualitative study of late adolescent girls, girls who objectified their own bodies felt less comfortable talking about sex and expressed greater regrets about having engaged in sex (Hirschman, Impett, & Schooler, 2006). Neither of these studies assessed adolescent girls' coital debut specifically. However, it is possible that highly objectified girls who are distanced from their own bodily feelings and desires have regrets about engaging in sex, feel less comfortable talking about sexuality, and perhaps feel less assertive in protecting themselves during sexual interactions may be more likely to engage in sex at younger ages than less objectified girls.

This Two-Year Longitudinal Study of Adolescent Girls

Although most researchers would agree that there are multiple pathways to early sexual intercourse among adolescent girls, most studies are limited to examining only a few predictors at a time and fail to account for the potentially complex interactions of multiple variables that may relate to coital debut. We believe that there are three main reasons that more complex relationships have not yet been examined in the existing literature on adolescent girls’ sexuality: (a) A small set of predictor variables are chosen due to time or assessment constraints, (b) only a small set of predictor variables are theoretically relevant to individual researchers’ particular interests, and (c) standard quantitative methods are ill-suited for exploring data with a large set of predictor variables and modest sample sizes. The first two reasons are perhaps inevitable or acceptable on practical or theoretical grounds; however, the last reason points to the need for sexuality researchers to examine these relationships with quantitative methods that are well-suited to identifying complex relationships among a large number of predictors.

In multiple fields, tree-based models have been used to circumvent the problem of having a large number of predictors with limited sample sizes (e.g., see Strobl, Malley, & Tutz, 2009; Zhang & Singer, 1999). This “small n–large p” dilemma is especially common in genetics research where thousands are genes are considered as potential markers for a disease. Even when faced with only a dozen predictors, the combination of all main and interaction effects of interest often leads to sparse cell counts that preclude accurate parameter estimates using traditional data analytic techniques (Strobl et al., 2009).

As an alternative, RP is a technique that examines all available predictors and creates a hierarchy of variables that are the most related to the outcome measure of interest. RP is best considered a technique that allows researchers to conduct exploratory analyses that can aid future theory development. In this study, we used RP to identify the strongest relative predictors of early coital debut among adolescent girls. Specifically, we examined girls who indicated that they had not yet engaged in sexual intercourse when assessed in the 8th grade, and then we re-assessed these girls again in the 10th grade. Using RP, we examined 19 unique predictors of coital debut, which we specifically selected based on prior research on adolescent girls’ sexuality and the research we reviewed earlier (see Table 1). All predictor variables were assessed in the 8th grade, and the outcome variable of coital debut was assessed in the 10th grade. To examine girls who transitioned to sexual intercourse between 8th grade and 10th grade, we limited our sample to girls who had not engaged in intercourse by the 8th grade in order to control for previous sexual experience.

Method

Participants and Procedure

This research uses data from an existing longitudinal study of adolescent girls. A more detailed description of the participants and procedure can be found elsewhere (see Impett et al., 2008). The complete dataset includes 183 adolescent girls from one New England public school district who participated in a three-wave longitudinal study, completing surveys in the 8th (1998), 10th (2000), and 12th grades (2002). Of the 147 girls who participated in the 8th grade, 80% (N = 117) were successfully recruited in the 10th grade. Eighty-nine percent of these girls (N = 104) reported that they had not engaged in sexual intercourse by the 8th grade; these girls represent the subset of the larger dataset of girls whom we included in our data analyses in this study.

Most of the girls reported their race as White (56%) or Latina (34%), with fewer girls reporting Black (6%) and “other” (4%). Most of the girls were 13 years old (77.9%) or 14 years old (19.2%) in the 8th grade (M = 13.19, Mdn = 13, SD = 0.45), and either 15 years old (44.7%) or 16 years old (44.7%) in the 10th grade (M = 15.61, Mdn = 16, SD = 0.67).

All of the girls completed a survey that included questions about friendship, dating, sexuality, and demographic
characteristics. Only those measures relevant to these analyses are described later. Because of the large numbers of Spanish-speaking students, bilingual and Latina girls were offered the option of completing surveys in Spanish (translated and back-translated) with a Spanish-speaking researcher present. All predictor variables were assessed in the 8th grade, whereas the outcome variable was assessed in the 10th grade.

Predictors

We have a variety of predictors that differ in terms of scale of measurement (e.g., dichotomous, ordinal, and continuous). Table 1 lists all of the predictors and their scale of measurement, and includes the coding schemes for the dichotomous variables, the anchors for the ordinal variables, and the reliability coefficients for continuous measures.

**Race and ethnicity.** Girls chose any number of six supplied racial and ethnic categories (Black, African American, or Caribbean; White; Hispanic or Latina; Brazilian or Portuguese; Asian or Pacific Islander; or American Indian or Alaskan Native), and some supplied their own category. Although girls self-identified as belonging to a variety of racial and ethnic categories, there were not enough girls in these groups to adequately test for group differences with regards to our questions of interest in this study. For the purposes of examining race and ethnicity in the model, we coded racial and ethnic group membership with two dummy codes: (a) White (0) versus non-White (1), and (b) Latina (0) versus non-Latina (1). We should note that in this sample, the majority of the non-White girls self-identified as Latina (predominantly Dominican).

**Socioeconomic status.** Participants were asked to report the highest level of education achieved by their mother/mother figure and father/father figure. Maternal education has been shown to be an adequate general index of socioeconomic status (Entwisle & Astone, 1994), and past research has examined both maternal and paternal education as indicators of socioeconomic status (e.g., Valle et al., 2009). To the best of their knowledge, girls reported the highest level of formal education achieved by both their mother and their father (1 = did not finish high school, 2 = finished high school/obtained General Equivalency Diploma, 3 = completed some college, 4 = finished college, 5 = attended school beyond college).

**Educational achievement.** We assessed educational achievement with the question, "Compared to other..."
students in your class, what kind of student would you say you are, in terms of grades?,” on a five-point scale ranging from 1 (near the bottom) to 5 (one of the best).

**School aspirations.** School aspirations were assessed with the question, “How far do you want to go in school?,” on a four-point scale ranging from 1 (quit as soon as I can), 2 (finish high school), 3 (college) to 4 (graduate or professional school).

**Religiosity.** We measured religiosity with a single item: “How important is religion in your life?” The scale ranged from 1 (not at all) to 4 (very). This one-item measure has been used in previous research with adolescents (Tolman et al., 2004) and is standard in the literature on religiosity more generally (Schwartz & Huismans, 1995).

**Self-esteem.** The 10-item Rosenberg Self-Esteem Scale (Rosenberg, 1965) was used to assess global self-esteem. Girls responded to such statements as, “I take a positive attitude toward myself,” on four-point scales ranging from 1 (disagree a lot) to 4 (agree a lot). Several items were reverse-coded, and all of the items were averaged to create a summary measure of global self-esteem (ranging from 1–4), with higher scores indicating more positive self-regard.

**Depression.** Depressive symptoms were assessed using the short form of the Children’s Depression Inventory (Kovacs, 1992). Girls responded to 10 statements such as, “How often did you feel sad in the past 2 weeks?,” with one of three possible answers (0 = once in awhile, 1 = many times, or 2 = all the time). Items were summed and used as a continuous variable, reflecting a range of depressive symptoms (from 0–20), rather than as a criterion for determination of clinical depression.

**Body satisfaction.** The body image subscale of the Self-Image Questionnaire for Young Adolescents (Petersen, 1984) was administered to measure girls’ body satisfaction. The body image scale contains seven items, and girls responded to questions such as, “I am proud of my body,” on a four-point scale ranging from 1 (disagree a lot) to 4 (agree a lot). Items were averaged to create a summary measure of body satisfaction (ranging from 1–4), with higher scores indicating greater satisfaction with one’s body.

**Body objectification.** We used a measure of body objectification created specifically for adolescent girls (Tolman et al., 2004; Tolman & Porche, 2000), rather than using an existing measure of objectification designed for young adult samples (McKinley & Hyde, 1996; Noll & Fredrickson, 1998). This measure was developed using extensive focus groups with an ethnically diverse sample of teen girls to ensure that items were representative of the developmental concerns of girls. This measure has been used in previous research with early adolescent (Tolman et al., 2004; Tolman & Porche, 2000), middle adolescent (Impett et al., 2011), and late adolescent girls (Impett & Tolman, 2006). Girls responded to 10 statements such as, “I think a girl has to be thin to feel beautiful” and “I often feel uncomfortable in my body,” on six-point scales ranging from 1 (strongly disagree) to 6 (strongly agree). Several items were reverse-coded, and mean scores were computed, with higher scores reflecting greater body objectification.

**Relationship inauthenticity.** The inauthentic-self-in-relationships subscale of the Adolescent Femininity Ideology Scale (AFIS; Tolman & Porche, 2000) was used to measure relationship inauthenticity. The AFIS was developed with and specifically for girls of varied ages in adolescence. Girls responded to 10 statements such as, “I wouldn’t change the way I do things to please someone” and “I tell my friends what I honestly think even when it’s an unpopular idea,” on six-point scales ranging from 1 (strongly disagree) to 6 (strongly agree). Several items were reverse-coded, and mean scores were computed, with higher scores reflecting greater inauthenticity in relationships.

**Physical intimacy experiences.** Participants were asked whether they had any experiences with dating, and they could respond with “yes” (e.g., currently dating), “no” (e.g., not currently dating), or “never” (e.g., never dated before). From these responses, we created two dummy-coded variables: dateyes (1 = yes, 0 = no or never) and dateno (0 = yes or never, 1 = no). Participants were also asked if they had engaged in four different types of physical intimacy experiences: holding hands, kissing, touching someone underneath their clothes, and being touched by someone underneath their clothes (1 = yes, 0 = no). In total, we measured six specific indicators of dating and physical intimacy experiences.

**Sexual self-acceptance.** Sexual self-acceptance was measured with eight items adapted from Winter’s (1988) Sexual Self-Concept scale to be more appropriate for use with early adolescent girls (for use of this adapted measure, see Impett, Schooker, Tolman, 2006). Items such as, “It’s normal for me to have sexual feelings,” were measured on four-point scales ranging from 1 (disagree a lot) to 4 (agree a lot).

**Outcome Variable**

**Coital debut.** In the 10th grade, girls were asked whether they had ever engaged in sexual intercourse. Thirty-three percent of the girls (n = 34) reported engaging in sexual intercourse by the 10th grade. Considering that we dropped girls who reported having engaged in sexual intercourse by the 8th grade from our final sample,
the percentage of girls in our dataset who reported being sexually active by the 10th grade (36%) is comparable to the percentage found in the National Longitudinal Study of Adolescent Health (40.9%), which is a large national survey of adolescents (Harris et al., 2009).

Data Analysis Plan

To be consistent with the traditional approach taken in the field of sexuality research, we used correlations and logistic regressions to examine the bivariate relationships between each predictor and the outcome of coital debut. These analyses allow us to determine which predictors are most strongly related to the outcome of coital debut in an absolute sense. In other words, these bivariate relationships show the predictive validity of each of these theoretically chosen variables in predicting coital debut in the sample of adolescent girls.

Then, in our central set of analyses, we used RP (Zhang & Singer, 1999) to further explore the multivariate relationships between all of the predictor variables and coital debut. Despite wide application of RP in a number of fields, we have yet to have seen its application to the study of coital debut specifically or in the field of sexuality research more generally. Traditional parametric techniques that are typically used with binary outcome variables such as coital debut (e.g., logistic regression) make assumptions regarding the distribution of the residuals and the form of the relationship between the predictor variables and the outcome variable (Agresti, 2007). Specifically, in logistic regression, the predictors are assumed to relate additively and linearly to the logit of the outcome variable. RP is a nonparametric technique that does not make any assumptions about the distribution of the dependent variable. Another benefit of RP is that it creates a decision tree that is both easy to understand and useful for practitioners who may not have access to sophisticated statistical software. For this reason, tree-based models are used quite frequently in medical fields (e.g., Calvocoressi, Stolar, Kasl, Claus, & Jones, 2005; Koziol et al., 2003).

RP is a data analytic technique that can be used in studies that contain a large number of predictor variables because, rather than examining all predictors simultaneously, it examines them in a stepwise fashion. First, the data are split along the coordinate axes of all predictor variables, and whatever split creates the two most homogenous groups on the outcome variable, or maximally distinguishes between, in this case, engaging in sex by the 10th grade versus not engaging in sex, is selected. Reflecting the strongest predictor, this first split occurs at the “root” of the tree, creating the first two branches that lead to two subsamples. Then, in each subsample, the process is repeated such that the predictor that most accurately distinguishes between success and failure is selected to split the data further into increasingly homogenous groups. It is important to note that a variable that has been used to make a split can be used again to make a split elsewhere in the tree. To create a parsimonious tree, we employed a statistical criterion that stops splitting the data when the splits no longer result in a statistically significant reduction in error (Strobl et al., 2009).

To further justify our use of RP in this study, we review the advantages and disadvantages of both stepwise logistic regression and RP. Both stepwise logistic regression and RP are limited by their exploratory approach, which makes them vulnerable to capitalizing on chance variation in the sample. Thus, when applied to a new/validation sample, one can expect the strength of prediction to decrease (i.e., shrinkage). Past research using real data has found that this shrinkage is similar for stepwise logistic regression and RP (James, White, & Kraemer, 2005). However, just as confirmatory research is required for theory-testing, exploratory research is required for theory development; thus, exploratory data analytic techniques are essential to parse a myriad of predictors and guide future theory.

One benefit of logistic regression is that it calculates an estimated probability of the outcome for each individual. In contrast, RP yields a decision tree that is both simple to use and understand, and does not require difficult calculations. Furthermore, RP does not share the linearity assumption imposed by a logistic regression, which makes it more robust to detect interactions among variables. An interaction is evidenced in RP when, after the first split is made, one predictor further splits only one of the remaining subsamples and not the other. Given that RP handles data differently than regression approaches, we cannot say that interactions in the RP framework are equivalent to a linear-by-linear interaction in regression, but RP does intuitively model interactive effects.

Results

In this study, 19 variables measured in the 8th grade were considered as potential predictors of coital debut by the 10th grade. Initially, we examined the bivariate associations between each predictor and the criterion variable of coital debut. Table 1 lists all predictor variables, as well as the correlations between each predictor and coital debut. In addition, odds ratios from logistic regressions are also included to show the independent main effects of each predictor on the criterion. These analyses revealed that higher maternal education and higher school aspirations were associated with a decreased likelihood of coital debut by the 10th grade, whereas dating and three specific physical intimacy experiences (i.e., kissing, touching someone underneath clothes, and being touched by someone underneath...
clothes) were associated with an increased likelihood of coital debut by the 10th grade. Table 2 depicts the correlations between all predictor variables.

We proceeded by constructing the tree model using the \textit{rpart} package (Therneau & Atkinson, 2010) in the statistical software program R (R Development Core Team, 2010). Using the statistical criterion to stop the growth of the tree, Figure 1 depicts the final tree model. The prediction error rate for this tree was .679, indicating that the tree explained 32.1% of the variance in coital debut in this sample. The final tree contains six predictor variables that identify seven homogenous subsamples that vary based on the likelihood of engaging in sexual intercourse by the 10th grade. Of the subsamples generated by our decision tree, two are classified as having a low likelihood of early coital debut (<10% of subsample reported coital debut), three as having a medium likelihood (>10% but <50% of subsample reported coital debut), and two as having a high likelihood of engaging in intercourse by the 10th grade (>50% of subsample reported coital debut).

There were two homogenous subsamples of girls generated by the decision tree that had a low likelihood of engaging in sex by the 10th grade. The group with the lowest likelihood (Low Group 1, \( n = 31 \); 3.2% engaged in intercourse by the 10th grade) included girls who reported never having been touched underneath their clothes, did not have high body objectification (scored <1.16 \( SDs \) above the mean), aspired to complete graduate education or beyond, and reported low or average levels of depressive symptoms in the 8th grade (scored <0.17 \( SDs \) above the mean).

The other group with a low likelihood of engaging in sex by the 10th grade (Low Group 2, \( n = 14 \); 7.1% engaged in intercourse) included girls who also reported never having been touched underneath their clothes and did not have high body objectification; however, this subsample did not aspire to pursue graduate level education, had relatively low feelings of body satisfaction (scored <0.39 \( SDs \) above the mean), and endorsed relatively high relationship inauthenticity in the 8th grade (scored >0.11 \( SDs \) above the mean).

There were three subsamples of girls generated by the decision tree that had a medium likelihood of engaging in sex by the 10th grade. These three subsamples of medium-risk girls differed from the low-risk subsamples by one of three key dimensions: depressive symptoms, relationship inauthenticity, or body satisfaction. For example, similar to the group of girls with the lowest likelihood of engaging in sexual intercourse by the 10th grade (i.e., Low Group 1), the first group of girls with a medium likelihood of having an early coital debut (Medium Group 1, \( n = 11 \); 36.4% engaged in intercourse) included girls who reported never having been touched underneath their clothes, did not have high body objectification, aspired to complete graduate education or beyond, yet reported high levels of depressive symptoms in the 8th grade. The second group, which had a medium likelihood of engaging in sex by the 10th grade (Medium Group 2, \( n = 9 \); 44.4% engaged in intercourse), included girls who were quite similar to girls in Low Group 2, which had a low likelihood of having an early coital debut, in that it included girls

\begin{table}[h]
\centering
\caption{Correlations among All Predictor Variables}
\begin{tabular}{lcccccccccccccccc}
\hline
Variable & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 \\
\hline
7. Self-esteem & -.13 & .03 & .00 & .02 & -.14 & -.20 & & & & & & & & & & & & & \\
8. Depression & .15 & -.14 & .01 & .07 & .10 & .20 & -.68 & & & & & & & & & & & \\
10. Religiousity & -.07 & .05 & -.03 & -.05 & -.02 & -.06 & .20 & -.08 & -.15 & & & & & & & & & & \\
11. Educational aspirations & .09 & .08 & .11 & .12 & .26 & .15 & -.33 & .19 & .08 & -.09 & & & & & & & & \\
12. School aspirations & -.13 & .00 & -.06 & -.01 & -.21 & -.31 & .22 & -.15 & -.09 & -.02 & -.30 & & & & & & & \\
13. Body satisfaction & -.09 & .19 & -.05 & .10 & -.13 & -.15 & .60 & -.48 & -.68 & -.01 & -.09 & .07 & & & & & & \\
14. Relationship inauthenticity & .13 & -.03 & .10 & .06 & .02 & .12 & -.47 & .40 & .39 & .10 & .10 & -.24 & -.45 & & & & & \\
15. Sexual self-acceptance & .17 & .04 & .34 & .35 & .25 & .28 & .28 & -.18 & -.13 & -.01 & -.06 & .05 & .06 & .02 & & & & \\
16. Maternal education & -.10 & .10 & -.12 & -.29 & -.27 & -.31 & .00 & -.07 & -.16 & .11 & -.05 & .03 & .09 & .06 & -.41 & & & \\
17. Paternal education & -.12 & .03 & .08 & -.08 & -.16 & -.15 & .13 & -.08 & -.27 & -.12 & -.13 & .06 & .27 & -.05 & -.07 & .38 & & \\
18. Race1 & -.02 & -.10 & -.03 & -.19 & .03 & .11 & .01 & .10 & -.10 & -.05 & -.03 & -.22 & -.01 & -.03 & -.01 & .20 & .04 & \\
19. Race2 & -.01 & .01 & .07 & .11 & .00 & -.08 & -.11 & -.04 & .16 & .02 & .04 & .14 & .02 & .15 & -.01 & -.15 & .09 & -.62 \\
\hline
\end{tabular}
\end{table}

\textit{Note.} Significant correlations (\( p < .05 \)) are in boldface type for emphasis.
who reported never having been touched underneath their clothes, did not have high body objectification, did not aspire to pursue graduate level education, had low feelings of body satisfaction, yet they endorsed lower relationship inauthenticity (i.e., were more authentic in their relationships with peers) in the 8th grade. The third group of girls who had a medium likelihood of engaging in sex by the 10th grade (Medium Group 3, n = 14; 50% engaged in intercourse) included girls who reported never having been touched underneath their clothes, did not have high body objectification, had low aspirations to pursue graduate level education, and also reported having high feelings of body satisfaction in the 8th grade.

The two subsamples of girls who had the highest likelihood of engaging in sexual intercourse by the 10th grade were determined by one of two predictors. Reflecting the most salient predictor, girls who reported ever having been touched underneath their clothes in 8th grade had a high likelihood of engaging in sex by the 10th grade (High Group 1, n = 17; 64.7% engaged in intercourse). However, among girls who did not report having been touched under their clothes in 8th grade, girls who reported having high levels of body objectification in the 8th grade had the highest likelihood of having an early coital debut (High Group 2, n = 8; 75% engaged in intercourse).

Discussion

In this two-year longitudinal study of adolescent girls, we examined multiple variables assessed in the 8th grade as potential predictors of the likelihood of girls engaging in sexual intercourse by the 10th grade. This longitudinal approach allowed us to demonstrate temporal precedence, which is one requirement when making causal inferences, giving us greater confidence in the direction of our predicted findings. By limiting our sample to only girls who were not sexually active in the 8th grade, we controlled for previous sexual experience. Given that early coital debut is associated with an increased risk for sexually unhealthy outcomes for adolescent girls, including engaging in sex with multiple sexual partners and increased risk of acquiring STIs including HIV and AIDS (Drain et al., 2004; World Health Organization, 2006), it is important to identify the strongest risk factors for early coital debut.

In a sample of early adolescent girls who had not engaged in intercourse by the 8th grade, we used RP to identify two subsamples of girls who had a low likelihood of engaging in sexual intercourse by the 10th grade, three subsamples of girls who had a medium likelihood of engaging in sex by the 10th grade, and two subsamples of girls who had a high likelihood of having an early coital debut. Each of these groups of girls was distinguished by predictors that our review of the literature on adolescent girls’ sexuality suggested would be related to early coital debut: whether a girl had been touched underneath her clothing, her level of body objectification, school aspirations, depressive symptoms, body satisfaction, and her feelings of inauthenticity in relationships.

Of particular interest are the two groups of girls with the highest likelihood of engaging in sexual intercourse by the 10th grade, which were identified by one or two predictors. Nearly two-thirds (64.7%) of the girls who reported having been touched underneath their clothes in the 8th grade had engaged in sexual intercourse by the 10th grade. In contrast, only 26.4% of the girls who had not reported having been touched underneath their clothes in the 8th grade had engaged in sexual intercourse by the 10th grade. Therefore, the most salient predictor in this sample was early experiences with physical intimacy, suggesting that coital debut may be the end of a sequence of progressively more intimate experiences. In fact, consistent with previous research (Davila et al., 2009; Longmore et al., 2004), three
measures of physical intimacy experiences were significantly related to coital debut in a bivariate sense (see bivariate correlations in Table 1), including kissing, touching someone underneath their clothes, and being touched underneath their clothes. Our findings corroborate previous studies that have suggested that sexual behaviors occur in a similar progression, such as kissing, French kissing, breast fondling, genital contact, and intercourse (Bingham & Crockett, 1996; Feldman et al., 1999). Early intervention efforts to delay coital debut may benefit from finding ways to delay other forms of physical intimacy. It is important to note, however, that we do not contend that all forms of early coital debut are problematic or risky, but that the nature of the relationship and the context should also be considered—a point to which we return shortly.

Among girls who had not experienced being touched underneath their clothes by the 8th grade, the most salient predictor of early coital debut was having high levels of body objectification. Among the girls who had not been touched underneath their clothes by the 8th grade, three-fourths of those who reported relatively higher levels of body objectification engaged in sexual intercourse by the 10th grade, whereas only 21.5% of individuals who reported lower levels of body objectification engaged in sexual intercourse by the 10th grade. These results support previous research that identifies body objectification as a predictor of negative psychological outcomes (e.g., Fredrickson & Roberts, 1997), as well as negative sexual health outcomes among adolescent girls (Hirschman et al., 2006; Impett, Schooler, & Tolman, 2006). It should be noted that body objectification was not related to coital debut in a simple bivariate sense, which could lead one to conclude (we believe, falsely) that body objectification is not an important predictor of coital debut. Our RP approach was able to identify that body objectification is a key predictor of coital debut among a subsample of girls—namely, those who had not been touched underneath their clothes by the 8th grade.

Whereas early physical intimacy experiences and body objectification were identified as factors that increased the likelihood that girls had engaged in sex by the 10th grade, having aspirations to attend graduate-level education and low levels of depressive symptoms were identified as factors that delayed a girl’s coital debut. Previous research has shown that having future education plans significantly delayed Latino adolescents’ coital debut (Marchi & Guendelman, 1994), as well the age at which Norwegian adolescents first engaged in sexual intercourse (Valle et al., 2009). Having the long-term goal of pursuing graduate-level education in the 8th grade may reflect two other potential factors that may protect against early coital debut: having a future time perspective and coming from a higher socioeconomic status. Duangpatra, Bradley, and Glendon (2009) found that future time perspective was inversely related to reckless sexual behaviors among emerging adults, and multiple studies have found lower socioeconomic status to relate to an increased likelihood of engaging in first intercourse at younger ages among diverse samples (Magadi & Agwandaf, 2009; Valle et al., 2009). In fact, level of maternal education was significantly related to coital debut in this sample in a bivariate sense (see bivariate correlations in Table 1).

At first glance, the finding that among a certain group of girls, having high body dissatisfaction was related to a lower likelihood of engaging in sexual intercourse by the 10th grade might seem both surprising and somewhat counterintuitive, especially given the findings from one study linking body satisfaction with earlier ages of coital debut (Valle et al., 2009). At the same time, however, this finding supports research by Schooler et al. (2005) in which women who felt dissatisfied with their bodies had less overall sexual experience (perhaps due to heightened concerns about how their bodies look to others) than girls who were more highly satisfied with their bodies. Future research should focus on mechanisms that might account for these findings such as body shame or ruminative thoughts about one’s body (Grabe, Hyde, & Lindberg, 2007).

Another somewhat surprising finding was that higher relationship inauthenticity (i.e., greater self-silencing) was related to a lower likelihood of engaging in sexual intercourse by the 10th grade for a particular sample of girls. This association is peculiar given that relationship inauthenticity has generally been associated with both negative mental health and sexual health outcomes (e.g., Impett et al., 2011; Tolman et al., 2004). The measure of relationship inauthenticity used in this study assesses the tendency for an adolescent girl to conceal her thoughts and emotions when engaging in conflict (Tolman & Porche, 2000), and conceptually similar measures have been shown to be related to depression in samples of both adult women and adolescent girls (Impett et al., 2011; Harper & Welsh, 2007; Jack, 1991; Uebelacker, Courтnage, & Whisman, 2003). It may be that inauthentic girls’ fears of rejection and desires to avoid conflict lead them to avoid sexual interactions altogether—particularly with their potential for expressing vulnerability. It will be especially important that these findings be cross-validated on an independent sample (Strobl et al., 2009) to see if this particular result replicates.

Applications of RP

To our knowledge, this is the first study to use tree-based models in the prediction of coital debut in a sample of adolescent girls. We hope that our application of RP demonstrates the utility of this approach to the field of sexuality research. Here, we briefly summarize the benefits of RP and review how we believe RP can be used effectively in this field. One of the greatest
benefits of RP is that it can be used to efficiently examine a large number of predictors even in the face of modest sample sizes. Therefore, future research should not be limited to examining a small handful of predictors as has typically been done when data analysts rely on more traditional statistical techniques. Perhaps the second greatest benefit of RP is that it creates an intuitive, comprehensible decision tree that can be used easily by practitioners, such as school counselors. Unlike other approaches of identifying subsamples in a population (i.e., latent class and mixture models), once a decision tree is created by RP, practitioners have no need to conduct complicated statistical analyses to identify subsamples.

Given that our application only represents one potential use of RP, we briefly discuss how RP can be used in future research. Once a decision tree is created in an exploratory sample using RP, as in this study, one can cross-validate that decision tree in an independent sample (test/validation sample) to determine whether the subgroups created in the exploratory sample are valid. To achieve this end, one would examine the predictive value of the classification tree, including examining the decision tree’s specificity (e.g., proportion of participants who do not debut and are not predicted to debut) and sensitivity (e.g., proportion of participants who debut and are predicted to debut). This process may be likened to the use of exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) to examine factor structure. If the factor structure is not validated by a CFA, then the researcher reserves the ability to reexamine the data with an EFA. In using RP, if a decision tree fails to replicate in a validation sample, a researcher would be able to reexamine the data with an exploratory RP analysis. Once a decision tree has been validated across samples, and shown to be relatively stable, additional predictors can be examined that predict belonging to any particular subgroup identified by the decision tree. For example, based on our data, future researchers may prioritize examining predictors of early physical intimacy given these experiences were the strongest predictors of coital debut.

Limitations and Future Directions

This study had a number of limitations that must be noted. First, RP is essentially an exploratory, stepwise technique, which can be vulnerable to random variations in a sample. Therefore, the results of this study should be considered preliminary, and subsequent research is needed to determine if the effects demonstrated here can be replicated in a new sample of adolescent girls, and ideally in new samples that include greater representation across all ranges of the key predictors (e.g., race and ethnicity, class or social status). It is important to note that all predictors were selected from a review of the literature; therefore, although our analyses were exploratory, theory played a key role in discerning the important and relevant predictors for this population and provided an important explanatory mechanism once the RP tree was complete.

We also had a limited ability to examine possible racial and ethnic group differences in coital debut in this sample. Although the sample included moderate numbers of both White and Latina (mostly Dominican) participants, girls from other racial and ethnic groups were underrepresented. Previous research indicates significant racial and ethnic group differences in several of the predictor variables (e.g., body objectification; Hirschman et al., 2006; Richter, 2000), and when analyzed separately, some interesting racial and ethnic differences in coital debut have emerged. For example, in one study, higher parental education predicted an older age at first intercourse among White girls, but not among Black girls, whereas higher parental educational expectations predicted an older age of first intercourse among Black girls, but not among White girls (Schvaneveldt, Miller, Berry, & Lee, 2001). Given our limited sample size, we were unable to examine tree models separately for White and Latina girls to determine whether these factors operate similarly or differently in these two racial and ethnic groups of girls. However, it must be noted that racial and ethnic group dummy codes were entered as potential predictors of coital debut, but no salient differences were found, corroborating other existing studies that did not find any racial and ethnic group differences in coital debut (e.g., Jordahl & Lohman, 2009).

Another limitation is that we are unable to make strong causal inferences from these data. Although collecting data across a two-year longitudinal period from early to middle adolescence demonstrates at least one requirement of making causal inferences (i.e., temporal precedence), we can only speculate as to the causal mechanisms involved in shaping adolescent girls’ coital debut. Future theory development is needed to make stronger causal inferences. In addition, changes in these constructs over time could be modeled as predictors of subsequent change in coital debut through the use of latent growth curve modeling (Duncan, Duncan, & Strycker, 2006). Combining theory and the use of other sophisticated statistical techniques such as latent growth curve modeling would allow stronger causal inferences than can be gleaned from the particular statistical analyses used in this study.

Another important direction for future research will be to assess the nature and quality of an adolescent girl’s relationship with her sexual partner. Although early coital debut has been linked with having sex with multiple partners (World Health Organization, 2006), as well with other problematic sexual health outcomes in adulthood (Magnusson & Trost, 2006), not all sexual episodes among early adolescents are inherently risky. It is possible that the health of a girl’s first sexual experiences may depend more heavily on the quality of her
relationship with her sexual partner than the timing of her first experience. Girls who are involved in relationships characterized by mutual trust and respect may engage in sex at younger ages, yet it is possible that these sexual episodes may be more likely to be protected. Ultimately, the use of condoms is the most important factor that can protect sexually active girls from acquiring STIs. Future research would benefit from applying a RP analysis to the study of adolescent girls’ condom use in addition to the study of early coital debut.

Strengths and Practical Implications

Despite these noted limitations, this study makes a number of unique contributions to our understanding of the factors that lead to early coital debut among adolescent girls and has important implications for prevention. We identified six variables that interactively predicted early coital debut among adolescent girls, including a girl’s experiences with physical intimacy, body objectification, school aspirations, depression, body satisfaction, and relationship inauthenticity. Using these predictors, we created a decision tree that can easily be used to identify 8th grade girls who have a low, medium, or high likelihood of engaging in sexual intercourse by the 10th grade. The decision tree showed that the factors that confer risk of early coital debut differ across subsamples of adolescent girls. From a prevention perspective, these findings suggest that a one-size-fits-all approach to the prevention of early coital debut is less than ideal. In fact, one could potentially tailor prevention programs to individuals based on the subsample to which they belong. Although the stability of this decision tree across samples has not yet been determined, it provides a good starting point for future research and theory development. As is the case with all statistical tests, stronger effects are more likely to replicate than weaker effects. With a decision tree, this fact means that predictors at the root of the tree (i.e., at the top) are more likely to replicate than predictors toward the bottom of the decision tree. For example, in our model, physical intimacy experience and body objectification are more likely to be stable predictors of coital debut than body satisfaction and relationship inauthenticity. In terms of tailored prevention, it is not likely pragmatic for schools or similar settings to implement several different programs targeted at each of the seven subsamples identified in our data. However, one could tailor different types of prevention messages to girls with and without early physical intimacy experience. Perhaps the program developed for girls without physical intimacy experience by the 8th grade could be directed at decreasing body objectification, as body objectification was identified as the strongest predictor of early coital debut among this subsample of girls.

In sum, the way that RP splits the data is ideal for identification of homogeneous subgroups in a sample, which is intuitively appealing given the present focus on tailored interventions. Thus, we believe that RP lends itself to identifying subgroups that can be used to tailor interventions. Unlike other methods of detecting latent subgroups in a populations (i.e., latent class and mixture models), RP can be used by practitioners quite easily and does not depend on complicated statistical calculations that are unavailable to practitioners.

References


