STRENGTHS AND WEAKNESSES

We chose case-control design to study the association between risk factors and infertility. We selected subjects on the basis of whether they were (cases) and were not (controls) infertile. We then compared these groups with respect to the proportion having a history of an exposure or risk factor. Because of this design, the case-control study that we designed in this study offered a number of advantages for evaluation of the association between an exposure (risk factor) and disease (infertility). Because we chose subjects on the basis of their disease status, this design allowed us to identify adequate number of affected and not affected individuals. Further, we were also able to evaluate a wide range of potential etiologic exposures that might relate to a specific factor as well as the interrelationships among these factors.

Strengths

Our study has several strengths. Cases and controls were enrolled from the same catchment area (study base) representing a fairly homogenous population with minimal migration. To facilitate recruitment, we defined fertility as pregnancy within the previous two years rather than current pregnancy. Our controls were representative of those at risk of infertility. We maintained a registry and enrolled consecutive women who fulfilled inclusion criteria for infertility in our study. Thus we avoided selection bias. We avoided spectrum bias by ensuring that all women, irrespective of duration of infertility were enrolled in the study. The hospital based design was proper for our study because cases and controls were similarly sensitized
for the recall of risk factors. To reduce the effect of confounding we chose age-matched controls and performed multivariate logistic regression analysis to adjust potential confounders. Our case definition was well standardized. We increased the efficiency of our design and boosted the power of our study by recruiting one control for each case. We did a careful examination of the patients, applied well-known criteria for evaluating hirsutism and used Rotterdam criteria for classifying our patients with polycystic ovarian disease syndrome. To detect ovulation, we assessed menstrual history, and measured FSH, LH, prolactin, TSH and testosterone – tests that helped us detect most endocrinopathy associated with infertility. To evaluate tubal status, we performed SHSG, HSG and chromotubation during laparoscopy. We chose laparoscopy because it represents a better chance of documenting tubal blockage detected on HSG and identify women with high risk of tubal disease. It also helped us manage any tubal adhesive disease and managing Polycystic ovaries by drilling concomitantly- thus avoiding a second invasive procedure. We evaluated the uterus to detect developmental abnormalities such as uterine septum and fibroids or endometrial polyps that disrupt endometrial function and hamper implantation. In the past, a routine part of the infertility evaluation included an endometrial biopsy to check for a luteal phase defect to ensure that follicular and endometrial tissue development was in sync. Because studies have shown this test to be a useful discriminator between the fertile and infertile population, but it is no longer a recommended part of the infertility evaluation. We, therefore did not subjects all infertile women for dilation and curettage.

We assessed the oocyte reserve by measuring both oocyte quantity and quality. However, almost all the women enrolled in our study were aged less than 35 years- the cut point associated with rapid decline of ovarian reserve. We measured
day-3 FSH to test the ovarian reserve and used FSH level > 10 mIU/mL to detect diminished oocyte quantity and infertility. We also did antral follicle count (AFC), a measure of all 2-to10 mm follicles less than eight in the early follicular phase which has also been shown to be useful marker of ovarian reserve and predictor of fertility treatment success. A measure of AFC more than 12 in number is a marker for PCOS.

We assessed all men partners of the study participants by their reproductive history, genitourinary history, medical history and past surgical history. We examined sperm concentration, motility and morphology in all male partners- regardless of the risk factors for infertility in their spouses- by using well defined cut points to classify their fertility status. Instead of a single value for each semen measurement that presumably distinguishes between “normal” and “abnormal,” we estimated the best two values that allow for the delineation of three groups — fertile, indeterminate, and subfertile. This classification system has been shown to be clinically meaningful and is appropriate to what is, biologically, a continuous function.

Finally, we have used 24-point checklist, endorsed by major medical journals, in the recent guidelines for reporting observational studies: STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) (214) Annexture I

Limitations

Our study has several limitations. First, the definition of primary infertility used in our study differed from the definition put forth by the WHO in two main aspects. First, the age range of women in our study, 16-39 year, was different than the range used by the WHO, 15-49 year. Second, our cases may not represent the community because those who were undiagnosed, misdiagnosed, did not seek medical
consultation or consulted non-allopathic doctors might not have been included in the study. Women were enrolled if they attended the outpatient department of our teaching hospital and our results may not be generalizable to other populations. Third, we did not measure serum anti-Mullerian hormone (AMH) level, a new marker of ovarian reserve in the study. Because AMH is cycle-independent and less subjective than the AFC, it can predict fertility treatment successes much better than the other markers of ovarian reserve. Fourth, despite normal results on fertility evaluation of the female partner, there may have been unrecognized subclinical male factors contributing to the infertility of the infertile couples. In our study, sperm morphology should have been assessed by a single person with extensive training and substantial experience, and reliability should have been monitored on an ongoing basis. The application of our results in clinical laboratories required the training of technicians and the implementation of tools for continuous calibration. We could not do this and instead, relied on sperm concentration reports from different laboratories because of logistic issues. Fifth, because of case-control design of the study, it was not possible to determine temporality or causality of the associations between primary infertility and the select covariates. Finally, confirmation of the validity of these thresholds for semen measurements in an independent sample of fertile and infertile men is needed.

Although we controlled for other risk factors in our multivariate model, we cannot exclude the possibility that some residual confounding by other lifestyle and risk factors remains. For example, we did not measure genetic risk factors known to be associated with infertility. Infertile couples have been shown to have a higher prevalence of karyotype abnormalities (trisomies, mosaics, translocations, etc.) than the general population. (215) The frequency varies according to the cause of infertility and clinical history. The most common aneuploidies associated with infertility are 45,
X (Turner syndrome) in women and 47, XXY (Klinefelter syndrome) in men. We did not do chromosomal studies in our study participants. Fourth, the assessment of frequency of sexual intercourse per week was based on self-reported recall. Recall bias is known to reduce the validity of such data. We are aware that the study participants might have under-reported the alcohol consumption. We might not have obtained accurate information on socio-economic status, from both cases and controls.

**Generalizability**

Our results may not be applied to different populations, different ethnic groups and different urban settings. Most women enrolled in our study came from neighboring villages and a small city and therefore our results may not be generalizable to the more affluent settings in India in which psychosocial stress, alcohol related problems and long-distance marriages might be key factors for infertility. Our findings highlight the importance of infertility as a public health issue. These data can be used to guide future reproductive health programs in the region.