

## Chapter 6

# Data Mining Approach for Predicting the Likelihood of Infertility in Nigerian Women

**Peter Adebayo Idowu**

*Obafemi Awolowo University Nigeria*

**Jeremiah Ademola Balogun**

*Obafemi Awolowo University, Nigeria*

**Olumuyiwa Bamidele Alaba**

*Tai Solarin University of Education, Ijagun, Ijebu-Ode Ogun State, Nigeria*

### ABSTRACT

*According to WHO, there are 60 - 80 million infertile couples worldwide with the highest incidence in some regions of Sub-Saharan Africa. The social stigma of infertility weighs especially heavily on women, who bear the sole blame for barren marriages in many developing countries and may face divorce as a result. Interviews were conducted with gynecologists at one of the Teaching Hospitals in Nigeria in order to identify likelihood variables for infertility. 14 risk factors were identified and data collected from 39 patients from the hospital was pre-processed and the variables used to formulate the predictive model for the likelihood of infertility in women using three different decision trees algorithms. The predictive model was simulated using WEKA environment. The results revealed that C4.5 algorithm had the highest accuracy of 74.4% while the least performance was for the random tree algorithm with a value of 53.8%. This chapter presents a predictive model which can assist gynecologists in making more objective decisions concerning infertility likelihood.*

### INTRODUCTION

The goal of data mining is to learn from data and the strategy used may vary as per the requirement (Yap et al., 2007). Data mining is an interdisciplinary field and is gaining popularity because of exploring Database technology, Information Science, Machine Learning and Neural Networks along with Statistical techniques. This study aims at applying data mining techniques to clinical data in order to predict the likelihood of infertility in Nigerian women.

DOI: 10.4018/978-1-5225-0920-2.ch006

Infertility is defined as one year of frequent, unprotected intercourse during which pregnancy has not occurred (Zegers-Hochschild et al., 2009). In another definition, infertility is the inability of a sexually active woman who is not practicing contraception to have a live birth (American Society for Reproductive Medicine (American Society for Reproductive Medicine (ASRM), 2008). Infertility can be attributed to any abnormality in the female or male reproductive system (European IVF-Monitoring Consortium, 2008). The etiology is mostly distributed fairly equally among the male and female with factors ranging from ovarian dysfunction, tubal factors amongst others. A smaller percentage of cases are attributed to endometriosis, uterine or cervical factors, or other causes. In approximately, one fourth of couples, the cause is uncertain and is referred to as unexplained infertility, while etiology is multifactorial for some couples (Puscheck & Woodad, 2009). Developed and developing societies tend to differ in prevailing assumptions about childlessness while in developed societies, voluntary childlessness is viewed as a more viable and legitimate option and women without children are often presumed to be voluntarily childfree (ASRM, 2006).

Infertility is a worldwide problem, affecting 8% – 15% of the couples in their reproductive age (Puscheck & Woodad, 2009; Boivin et al., 2009; Audu et al., 2003). WHO estimates that there are 60 - 80 million infertile couples worldwide with the highest incidence in some regions of Sub-Saharan Africa which is said to have an infertility belt warped around its center as infertility rate may reach 50% compared to 20% in Eastern Mediterranean Region and 11% in the developed world (Okonofua, 1996). In some parts of this belt, infertility is said to constitute up to 65% of gynecological consultations (Idrisa, 2005). Early exposures (e.g. in utero or in childhood) could permanently reprogram men and women for fecundity or biologic capacity (e.g. gynecologic and urologic health or gravid health during pregnancy) and fertility outcomes (e.g. multiple births or gestational age at delivery) which could affect adults later on set diseases (Skakkebaek et al., 2001). Thus, infertility could have public health implications beyond simply the inability to have children.

In general, an infertility evaluation is initiated after 12 months of unprotected intercourse during which pregnancy has not been achieved. Earlier investigation may be considered when historical factors, such as previous pelvic inflammatory disease or amenorrhea suggest infertility, although physicians should be aware that earlier evaluation may lead to unnecessary testing and treatment in some cases. Evaluation can be initiated earlier if the female partner is older than 35 years, because fertility rates decrease and spontaneous miscarriage and chromosomal abnormality rates increase with advancing maternal age (Puscheck & Woodad, 2009). Partners can be evaluated together or separately which can be due to the fact that each person may want to reveal information which their partner may be unaware of, such as previous pregnancy or sexually transmitted disease.

The risk factors for infertility can be classified into: genital, endocrinal, developmental and general factors. Pelvic inflammatory disease (PID) due to sexually transmitted diseases, unsafe abortion, or puerperal infection are the main causes of tubal infertility caused mainly by chlamydial infection. Polycystic ovarian syndrome (PCOS) is thought to be the commonest cause of an ovulatory infertility. Several lifestyle factors may affect reproduction, including habits of diet, clothing, exercise, and the use of alcohol, tobacco and recreational drugs. Exposure to textile dyes, lead, mercury and cadmium, volatile organic solvents and pesticides has been also associated with infertility. Estimates of the proportion of infertility cases attributable to male or female specific factors in developed countries were derived in the 1980s by the WHO: 8% of infertility cases were attributable to male factors, 37% to female factors, 35% to both the male and female, and 5% to an unknown cause (the remaining 15% became pregnant) (Zegers-Hochschild et al., 2009).

25 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/data-mining-approach-for-predicting-the-likelihood-of-infertility-in-nigerian-women/163822](http://www.igi-global.com/chapter/data-mining-approach-for-predicting-the-likelihood-of-infertility-in-nigerian-women/163822)

## Related Content

---

### Towards Semantic Interoperability in Health Data Management Facilitating Process Mining

Barbara Traxler, Emmanuel Helm, Oliver Krauss, Andreas Schulerand Josef Kueng (2018). *International Journal of Privacy and Health Information Management* (pp. 1-12).

[www.irma-international.org/article/towards-semantic-interoperability-in-health-data-management-facilitating-process-mining/211973](http://www.irma-international.org/article/towards-semantic-interoperability-in-health-data-management-facilitating-process-mining/211973)

### Design and Implementation of Digital Asthma Diagnosis System

Qinghua Yaoand Xiantao Yang (2019). *International Journal of Healthcare Information Systems and Informatics* (pp. 1-15).

[www.irma-international.org/article/design-and-implementation-of-digital-asthma-diagnosis-system/234316](http://www.irma-international.org/article/design-and-implementation-of-digital-asthma-diagnosis-system/234316)

### Workarounds: Risk or Benevolence for Patient Safety?

Nada Nadhrahand Vaughan Michell (2014). *Handbook of Research on Patient Safety and Quality Care through Health Informatics* (pp. 76-113).

[www.irma-international.org/chapter/workarounds/104075](http://www.irma-international.org/chapter/workarounds/104075)

### Predicting Seminal Quality and its Dependence on Life Style Factors Through Ensemble Learning

Satya Ranjan Dashand Ratula Ray (2020). *International Journal of E-Health and Medical Communications* (pp. 78-95).

[www.irma-international.org/article/predicting-seminal-quality-and-its-dependence-on-life-style-factors-through-ensemble-learning/246079](http://www.irma-international.org/article/predicting-seminal-quality-and-its-dependence-on-life-style-factors-through-ensemble-learning/246079)

### A Roadmap to the Introduction of Pervasive Information Systems in Healthcare

Fotis Kitsios, Thanos Papadopoulosand Spyros Angelopoulos (2011). *Wireless Technologies for Ambient Assisted Living and Healthcare: Systems and Applications* (pp. 1-13).

[www.irma-international.org/chapter/roadmap-introduction-pervasive-information-systems/47117](http://www.irma-international.org/chapter/roadmap-introduction-pervasive-information-systems/47117)