Predicting the Probability of Down Syndrome

New Mexico Adventures in
Supercomputing Challenge
Final Report
April 6th, 2005

Team #56
Santa Fe High School

Team Members:
Luwen Huang

Teacher(s):
Mrs. Anita Gerlach

Project Mentor:
Dr. McClenahan

Table of Contents

Executive Summary:

In this project I attempted to efficiently predict the probability of having a child with Down Syndrome. Given a certain set of conditions, my program computes the final risk value that results from the combination of these conditions.

Down Syndrome (D.S.) is a genetic disease in which gene mutations cause the life born to be mentally deficient. It is the leading genetic disease, with over millions of people worldwide affected. Genetic testing is expensive, and this may set the stage for the clinical trials. The task was theoretically simple: an input, output program that stores numbers and at the end, using a simple probability equation, gives a single risk percentage to the user.

Body of the Report:

Introduction:

The purpose of this project is to calculate combined probabilities of D.S. that would otherwise be tedious if done manually. D.S. affects millions of people worldwide, and this program functions as an easy tool to determine your risk.

Description:

The first step was to determine which factors are to be used in calculating the final risk. There are three types of D.S., so therefore, the factors must be taken into account for each. Scouring the library and the internet, I found that only maternal age, paternal age, family history, and second child recurrence positively affect D.S. I used a single, reliable, medical journal for the data of each factor. Once my data was organized, I created a program tree to illustrate the logical progression of questions. The program is, essentially, a series of "if...else" questions, so therefore, organizing the questions logically is important. The program was programmed entirely in the Visual C++ environment.

There were certain limits to this project, the most obvious being the one of reality. Mathematically, the combined numbers should give an accurate risk value. In real life, however, other conditions can spontaneously and randomly affect the pregnancy, eschewing the predicted risk value. Also, because of privacy laws, my program could not be tested on real pregnant women. The materials used were the C++ compiler and the medical journal.

Results:

The results were expected though broad due to the lack of data. I tested inputs, and the outputs corresponded generally to the national pregnancy results. For example, a "test patient" of mine had a family history of standard trisomy 21 D.S., is 39 years old, and had a child of D.S. My program predicted a relatively high chance of having a 2^{nd} child with D.S. This matched with the national figures as pregnancies of this sort usually resulted in death, abortion, D.S., or all three. Another test subject, with age 42, no history of D.S., and one child of standard trisomy 21, was calculated to have a probability almost exact to her age-specific risk. All my results are relatively precise with each other and with the national figures.

Conclusion:

My program is simply a tool to easily calculate the risk for D.S. based on certain questions. My results were compared to documented numbers, charts, and tables. The unique aspect of this program is that the program combines D.S. factors, takes into account their weighs, and creates a single risk number. Since the information on D.S. is not very complete, this achievement had been a challenge

Recommendations:

The program could be expanded and made more accurate by adding additional algorithms to determine the genetic structure of the user. If the exact genotype(s) was known, then the program could calculate exactly the risk percentage. To achieve this, I would have to develop an extensive list of questions to ascertain the type of D.S.

involved, use this to find out the genotype, and then use the genotype along with my numbers and equations to make a precise and accurate risk figure.

The results derived fitted in consistently with the overall national results. However, they could be made more complete as research on D.S. is still far from done. There was a lack of exact information such as numbers and different sources often conflicted with each other. I believe that with more information, more definite results could be reached, but as of yet, this information is withheld from the public. From this lack of data, I would strongly recommend clinical testing even after the usage of this program.

Acknowledgements:

I owe my greatest gratitude to Dr. Chuck McClenahan, my mentor, whose aid has been invaluable and insightful. Thanks to him, I wrote my code without too much pain and suffering. This project would not have been without him.

References:

1. R.J.M. Gardner and G.R. Sutherland, “Chromosome Abnormalities and Genetic Counseling”, Oxford University Press, 200 Madison Avenue, New York, New York 10016, Copyright 1989, pages 20-26, 188-191, 147-143
2. Kaplan and Dale, “The Cytogenetic Symposia 1994”, The Association of Cytogenetic Technologists, Burbank, California, Copyright 1994, pages 3-4
3. www.ndss.org, Copyright 1995-2005, National Down Syndrome Society

Appendixes

Equations: P=P1+P2-P1*P2 P is total probability.

Computer Program: Attached.