A cost-effectiveness analysis of different strategies in prenatal screening for Down syndrome in Nepal: are we there yet?

Sujan Rijal

Master's Thesis

Department of Health Management and Health Economics

The Faculty of Medicine

University of Oslo

May 2021

A cost-effectiveness analysis of different strategies in prenatal screening for Down syndrome in Nepal: are we there yet?

Sujan Rijal

Supervised by- Hans Olav Melberg

Thesis submitted as a part of the Master of Philosophy Degree in Health Economics, Policy and Management

Department of Health Management and Health Economics

The Faculty of Medicine

University of Oslo

Contents

Acknowledgment

Firstly, I would like to thank my supervisor Hans Olav Melberg for taking me as his student when I struggled to find a supervisor for my thesis. He has provided me with many technical inputs, great support, and motivation throughout this journey. His witty comments and sense of humor made this journey less stressful than it would have been. He was also instrumental in introducing me to topics that I would not have got into otherwise. I would like to take this opportunity to express my gratitude towards him for all this time, effort and consideration.

Secondly, I would like to thank Mari Elisabeth Mjøen for all the assistance over the two the course of two years. To all the faculty of the Department of Health Economics and Health Management at the University of Oslo and those involved in guiding us through these tough yet wonderful two years who have played a huge role in our professional development.

From being my lecturer in the undergraduate years to a mentor for all these years, I would like to dedicate my special thanks to Prof. Dr. Archana Amatya who has been such an inspiring figure in my life. I thank you for all the love, motivation and continuous inspiration to do better things and never stop improving. Finding data on Nepal was one of the most difficult parts of this thesis. I would like to express my humble gratitude to Mr. Sushil Thapa for entertaining my questions and answering them without being bothered. I am also indebted to Dr. Suman Dhakal, who helped as if it was his thesis in finding out relevant data.

I would like to thank Ms. Bidusha Neupane for always being there and pushing me all the time. I doubt I would have been able to do it without your support and constant motivation. You always have been my backbone and pushed me to become a better version of myself. I dedicate my master's thesis to you.

Last but not least, I would like to thank my parents for providing me with all the help so that I could pursue my goal. I would like to thank all those who were directly or indirectly involved in this wonderful journey.

Thank you.

Summary

Introduction: Down syndrome (DS) is a condition when a child is born with all or part of a third copy of chromosome 21 in some or all the body cells and is the most frequent chromosomal abnormality in humans and affects anywhere between one in 400 to 1500 pregnancies. DS can be fatal or result in various kinds of disabilities. DS can be identified early in the pre-natal stage of pregnancy through various screening tests with varying accuracy which helps a woman make a decision to avoid a pregnancy with DS or not.

Objectives: The objective of this study was to evaluate the cost-effectiveness of various prenatal screening methods for DS in terms of cost per number of cases detected compared to no screening from a health care perspective.

Methods: Various screening tests like maternal serum marker tests, NIPT and nuchal translucency (NT), were compared to no screening. The total costs and the number of cases detected were calculated using a decision tree for three different age groups of singleton pregnant women (15-49 years, 35 years or above, and 40 years and above) and one special scenario (40 years and above with coinsurance). Sensitivity analyses (one-way, two-way and scenario analysis) and probabilistic analysis (PA) were conducted to validate the findings of the deterministic analysis. The results from cost-effectiveness analysis (CEA) was also supplemented by safety related outcomes like safety index and harm to benefit analysis, and budget impact analysis. Finally, a value of information (VOI) analysis through expected value of perfect information (EVPI) and expected value of partial perfect information (EVPPI) was also carried out to explore the possibility of collection of information of input parameters.

Results: The cost per detected case of DS ranged from $37661, $11794, $8940 (general population, ≥35 years and ≥40 years) for NT to $343473, $102306, $70688 (general population, ≥35 years and ≥40 years) for NIPT. The number of fetal losses due to amniocentesis varied from 12, 8, 18 (general population, ≥35 years and ≥40 years) for NIPT to 552, 173, 113 (general population, ≥35 years and ≥40 years) for QT. No screening strategy was found to be cost-effective at a threshold of $3,000 per incremental case detected regardless of the age group. NIPT was the safest strategy in terms of the number of fetal loss, but CUB had the best combination of cost, case detected, and the number of fetal loss. However, NT was found to be cost-effective in a special scenario for women 40 years or older, where the cost is decreased by 70%. One way of achieving this is introduction of coinsurance (30% for health payer-70% for the individual), with an ICER (same as CER) of $2689 per additional case detected (52%

probability of being cost-effective from PA analysis). Sensitivity analyses (both one-way and two-way) showed that incidence of DS, cost of individual, cost of amniocentesis, and respective test sensitivities affected the CER and ICER values. However, no interventions were cost-effective for any of the age-groups even in the best case-scenario at the current willingness to pay (WTP) threshold. VOI analysis for all the three age groups show that there is no value added by the collection of perfect information of all the input parameters (EVPI=0) at the current WTP threshold. However, collection of perfect information on all the parameters can be worthwhile and help in making decision with more certainty population EVPI=$124,700 per year). Similarly, collection of information on probabilities of spontaneous miscarriages and foetal loss due to amniocentesis (population EVVPI= $118,020 per year) would bring the most value in decision making followed by test input values i.e. sensitivities and false-positive rates (population EVVPI= $47,530 per year). Information on acceptance rates of screening and diagnostic tests is worthwhile if the total cost of research does not exceed $27,584 per year (population EVVPI= $27,584 per year). Lastly, NT had the lowest budget (in millions) over the course of five years i.e. $78.66, $24.63 and $16.35 (general population, $\geq$35 years and $\geq$40 years) with NIPT being the costliest intervention to implement with more than ten times the cost of NT.

Conclusion: The choice of prenatal screening strategies for Nepal at the current willingness to pay threshold is straightforward, i.e., no screening regardless of the age group. NT can be introduced as a screening strategy only if provided to women 40 years and older if the cost is decreased by 70%. One way of achieving this is introduction of coinsurance where the cost of NT is divided between the health system and individual. VOI analysis shows that the collection of further information is worthwhile for some group of parameters for this special scenario.

Abbreviations

BIA	Budget impact analysis
CBA	Cost-benefit analysis
CDC	Centre for Disease Control and Prevention
CE	Cost-effectiveness
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CEAF	Cost-effectiveness acceptability frontier
CER	Cost-effectiveness ratio
CUA	Cost-utility analysis
CUB	Combined ultrasound test
CVS	Chorionic Villi Sampling
DALY	Disability-adjusted life years
DoHS	Department of Health Services
DS	Down syndrome
DT	Double Test
EE	Economic Evaluation
EVPI	Expected value of perfect information
EVPPI	Expected value of partial perfect information
EVSI	Expected value of sample information
FPR	False positive rate
GON	Government of Nepal
ICER	Incremental cost-effectiveness ratio
LMICs	Low-middle income countries
MoHP	Ministry of Health and Population
NCD	Non-infectious disease
NGOs	Non-governmental organization
NHB	Net health benefit
NIPT	Non-invasive prenatal test
NMB	Net monetary benefit
NPV	Negative predictive value
NT	Nuchal Translucency
NTDs	Neural tube defects
PPV	Positive predictive value
PSA	Probabilistic sensitivity analysis
QALY	Quality-adjusted life years
QT	Quadruple test
SAVI	Sheffield Accelerated Value of Information
TT	Triple test
VOI	Value of information
WHO	World Health Organization
WTP	Willingness to pay threshold

Chapter I: Introduction

This chapter of the report consists of the introductory part of the thesis. Firstly, definitions, risk factors, and epidemiology of congenital disabilities or birth defects, in general, have been discussed. The second part of the chapter consists of the introduction, cause, risk factors, and Down syndrome (DS) complications. All the screening and diagnostic procedures in practice worldwide are listed, followed by explaining the socio-economic context and Nepal's health system, the study setting. The chapter ends with a brief part of the situation of DS and screening of the same in Nepal.

1.1 Congenital disabilities

Birth defects or congenital anomalies are the conditions that present at birth irrespective of the causes and lead to significant health and development complications, for instance, physical, mental, psychological, developmental, and intellectual disabilities (1, 2). They can be defined as abnormalities, be it structural or functional, and include metabolic disorders, which are seen from birth and can affect any part of the body (3). They comprise of different conditions, and each condition has its etiology. Such defects can occur as isolated events or can present themselves in the form of multiple defects. When several congenital disabilities affecting more than one organ or a system as a whole occur in individuals and families and are recurrent, they are assumed to have common underlying causes and are termed as birth defect syndrome, e.g., Apert syndrome, Down syndrome, etc. This makes it easy to better understand the baby's prognosis affected by the syndrome and helps the clinical geneticists determine the possibility of the woman having babies with similar conditions in subsequent pregnancies (4).

Unlike other diseases, birth defects affect people from all ethnicities, races, socioeconomic status, and demographic factors. However, the risk of having a baby with certain birth defects is particular to the defect and varies with the genetic and underlying factors. Most of the birth defects that occur are not identified as syndromic, so definitive etiology cannot be established (4). This leads to the concept of multifactorial cause of a defect, which refers to the idea that these conditions result from complex combinations of various genetic and/or environmental factors (5, 6). These can originate from genetic, environmental, nutritional factors and infections or combinations of these, but how these factors come together to cause the defects is not fully understood yet (7, 8). Through past research, various behavioral factors like smoking, drinking, or taking drugs in pregnancy, hereditary factors like family history with birth defects, and physiological factors like the presence of medical conditions like

hypertension, obesity, uncontrolled diabetes, and advanced maternal age have been reported to be some of the risk factors for developing birth defects in general (1, 9).

It is estimated that as high as 303,000 newborns die every year worldwide due to birth defects. Almost one-third of these deaths occur in South-East Asia alone. It was estimated that around 260,000 neonatal deaths, about 7% of the total neonatal deaths worldwide, were due to congenital anomalies (10). The Centers for Disease Control and Prevention (CDC) estimates that a baby is born with a birth defect every four and half minutes in the US (1). Moreover, they contribute even more to mortality in settings where the overall mortality of rates is lower. For instance, deaths caused due to birth defects account for almost 25% of total neonatal deaths (2).

Although other factors are also significant in causing birth defects, genetic factor is the most significant. It was reported that almost 8 million total births worldwide are born with some kind of birth defect of complete or partial genetic origin. All the causes of congenital birth defects fall into three categories: 1. Single gene-effect 2. Chromosomal abnormalities, and 3. Multi factorial influences. Genetic congenital disabilities can occur at any stage during pregnancy. The genetic makeup of the person is determined at conception. A flaw in the nuclear activity during the process of fertilization is the time when most of the congenital disabilities are determined. For instance, duplications or deletion of segments of chromosomes or entire chromosomes can occur and later be carried further into pregnancy (7).

The March of Dimes global report on birth defects also states that the five most common birth defects of genetic origin are: 1. congenital heart defects; 2. neural tube defects; 3. the hemoglobin disorders, thalassemia, and sickle cell disease; 4. Down syndrome (trisomy 21); and 5. glucose-6-phosphate dehydrogenase (G6PD) deficiency (8).

1.2 Down syndrome

Down's syndrome (DS) is a condition when a child is born with all or part of a third copy of chromosome 21 in some or all the body cells, resulting in the subsequent increase in the expression of the genes. It is the most frequent chromosomal abnormality in humans and affects anywhere between one in 400 to 1500 pregnancies and varies according to population characteristics like maternal age and health system with or without screening facilities (11). The primary health and functional difficulties with Down's syndrome are mental retardations and increased risk of various organ defects, heart disease, and Alzheimer's diseases, dysmorphic head, and neck, characteristic facial and physical features (12).

First identified by John Langdon Down in 1866, the condition was referred to Mongoloids as the children with the condition resembled Mongolian people (13). In 1959, it was finally determined that the presence of an extra copy of chromosome 21 was DS's cause (13, 14). All the DS cases are categorized into three types. The first and the most common of the three, accounting for 95% of all the DS cases, occurs when all the cells have three copies of chromosome 21, i.e., every cell has 47 chromosomes (15, 16). The second type occurs when the extra chromosome 21 is attached to another chromosome and termed translocation. The extra chromosome can be attached to chromosome 21 or chromosomes 13, 14,15, or 22. In such cases, the long arms of chromosome 21 fuse together, and the shorter arms are lost, and thus the total number of chromosomes is still 46. This type accounts for about 4% of the total cases (17). As the translocation can be inherited, this type of DS is also called familial DS. The last, "Mosaics," occurs when an individual has only some cells with trisomy 21. This is the rarest form with 1% of all the cases. The extra chromosome is of maternal origin in most cases (88%) and only 9% of paternal origin. The rest arises due to post-zygotic mitosis (11).

As discussed earlier, DS can affect people from any race, ethnicity, socio-economic status, geography, etc. The risk of having a child with DS is primarily dependent on individual characteristics. For example, nutritional intake, and use of alcohol and drugs during pregnancy, etc. (18). One of the most critical determinants of DS is maternal age (19-21). It has been well established that the risk of DS increases with maternal age. A study states the birth prevalence of DS increases from about 1 in 1350 to about 1 in 940 at age 30 compared to under 25. The birth prevalence increases as high as 1 in 85 at age 40 (20). A report from India shows similar data relating maternal age with increased incidence of DS (22). The risk of women having a second child with DS increases if their previous pregnancy was diagnosed with DS and must be taken into account with age-specific risk to calculate the woman's total risk (19).

1.3 Cost of DS

Unlike trisomy of other chromosomes like Edward and Patau's syndrome, the fetus's survival rate with DS is considerably higher, and so is the life expectancy. Over three-fourths of the fetuses, when the birth prevalence is around 1 in 150, end in miscarriage. More than half of the liveborn babies with DS have congenital structural abnormalities, the most common of which are heart defects accounting for 46 percent of the total abnormalities, followed by abnormalities of the intestinal tract, limb defects, congenital cataracts, etc. The risk of having intestinal tract abnormalities (specifically duodenal atresia) is 300 times higher in the fetuses with DS than normal fetuses. Almost 20% of the fetuses with DS die before the age of 5. However, the

likelihood of survival increases after that age. The average LE of persons with DS can be anywhere from 50 to 55 (12).

Similarly, mental retardation is substantial in persons with DS. A study reported that the mean IQ of persons with DS at the age of 21 was 42 (CI: 8-67), which was much lower than the general population's IQ (mean- 100). The mental age at the age of 21 was found to be around 5 (range of 1-8 years) (12).

There are various other medical problems associated with DS like Leukemia, Epilepsy, Alzheimer's disease, etc. Persons with DS are 43, 27, 11, and 39 times more likely to develop Leukemia, Primary congenital, Epilepsy, and Alzheimer's disease compared to the general population (12).

1.4 Screening and diagnosis of down syndrome

Screening tests include one or more markers to select screen-positive women for further tests or at risk of having a specific condition. For e.g., Pap smear is a screening test used for cervical cancer. The screening tests are generally followed by diagnostic tests or procedures used to determine the presence of the condition in the at-risk people. They are used as a confirmatory tool. For e.g., MRI, Biopsy, CT scan, ultrasound imaging, etc. For instance, ideally, if a woman is defined to be at risk of having a baby with DS by any of the screening tests like the double test, triple test, NIPT, etc., the woman goes through either amniocentesis or CVS to confirm the presence of DS.

The diagnostic tests for Down syndrome were available since the mid-1960s after discovering that extra chromosome 21 was the cause. However, there were various issues, along with the positives, with the diagnostic tests. There was a risk of fetal loss since the diagnostic tests were invasive. Similarly, those tests' high cost meant that providing these tests to all pregnant women would incur a high financial burden on the health system. The first maternal serum marker screening test was introduced in 1984 for Neural Tube Defects (NTDs) but could be used for DS as well. Since then, over four decades, various prenatal screening tests have been introduced with varying test strengths. These tests are conducted at different pregnancy stages and use maternal serum markers, ultrasound, or a combination of both, or a study of DNA fragments to classify women at risk and identify those who have pregnancies with DS. All the screening and the diagnostic tests available and used for DS are listed in the table below (11, 12).

Table 1Screening and diagnostic tests for DS

Name of the test	Weeks (Trimester)	Type of test	Procedure	Maternal serum markers examined	Year introduced
Nuchal translucency test (NT)	11 to 14 (1st)	Screening	Ultrasound	Nuchal translucency	1990s
Double test (DT)	10 to 13 (1st)	Screening	Blood sample	Pregnancy-associated plasma protein: PAPP-A free B-hCG (Beta- Human Chorionic Gonadotrophin)	1987
Combined Ultrasound and Biochemical test (CUB)	11 to 14 (1st)	Screening	Ultrasound and blood sample	Pregnancy-associated plasma protein: PAPP-A free B-hCG (Beta- Human Chorionic Gonadotrophin)	1995
Triple test (TT)	15 to 18 (2nd)	Screening	Blood sample	Alpha-fetoprotein Unconjugated estriol (uE3) free B-hCG (Beta- Human Chorionic Gonadotrophin)	1988
Quadruple test (QT)	15 to 18 (2nd)	Screening	Blood sample	Alpha-fetoprotein Unconjugated estriol (uE3) Free B-hCG (Beta- Human Chorionic Gonadotrophin) Inhibin-A	1998
Integrated test (IT)	1st and 2nd	Screening	Ultrasound and blood sample	Alpha-fetoprotein Unconjugated estriol (uE3) Free B-hCG (Beta- Human Chorionic Gonadotrophin) Inhibin-A Nuchal translucency	1999
Non-invasive prenatal test (NIPT)	After 10 weeks (1st)	Screening	Blood sample	cf-DNA	1997
Amniocentesis	15 to 20 weeks (2nd)	Diagnostic	Invasive	Amniotic fluid	Mid 1960s
Chorionic Villi Sampling (CVS)	11 to 15 weeks (2nd)	Diagnostic	Invasive	Chorionic villi	Mid 1960s

Many developed countries worldwide offer prenatal screening for genetic birth defects like Down's syndrome, Patau's syndrome, neural tube defects (23, 24). There are different screening options available and vary from country to country. In Europe, the combination of NT and first-

trimester double test (also termed as the Combined test in some countries) is generally used as the benchmark test. France was one of the first countries to implement a DS screening policy in 1997, which included NT and second-trimester triple test proposed to all the women by law. In 2010, the French High Authority of Health recommended the use of combined first-trimester screening with NT measurement and first-trimester serum screening for all the women. All the cost associated with screening was reimbursed. Amniocentesis is offered if the screening test results show that a woman is at risk and the cost reimbursed (25).

Similarly, in other countries like Belgium, Netherlands, Spain, Denmark, Finland, and Norway, the combination of NT and the double test has been used as the standard followed by either Amniocentesis or CVS as a diagnostic procedure. However, not all women are provided the tests in every country, and the re-imbursement of the cost also differs within these countries. Spain, Denmark, Finland, and Belgium provide these tests to all pregnant women, and the costs are reimbursed. On the other hand, Norway and Netherlands provide these tests for women over 38 and 36 years old at conception respectively. Pregnant women 36 years or more are provided triple test, NT alone, or/and double test in Italy. There are some countries like Austria, Sweden, Ireland, Croatia, etc., where there is no official policy regarding the screening of DS. Still, women are provided with the choice of going through various tests on an individual basis. The costs are re-imbursed if conducted in public facilities (25).

Similarly, pregnant women in Australia have access to first-trimester maternal serum screening along with NT measurement which are partly covered by the government whereas NIPT is not financed (26, 27). The provision is, however, different in the UK and the US. The United Kingdom National Screening Commission (UKNSC) recommends from 2010 that any test with a detection rate of more than 90% and screen positive rate less than 2% be used as the preferred choice. The choice of the test differs within the UK and England. Wales currently has the provision of triple test for all pregnant women while Scotland has combined test and Northern Ireland has no policy regarding this. Even within England, the CUB test is used in the northern part of the country, while second-trimester triple or quadruple test are used as the standard (25).

In the US, there is no official national policy regarding the screening and diagnostic test for DS. People generally chose the test of their choice in consultation with a genetic counselor. The private health insurer covers the cost in the general population. People with low income are covered by the free health care funded by the government (28). A survey from 2011 and 2012 showed that the most common DS screening tests in the US were second-trimester

screening, with the Quadruple test the most used, followed by integrated test and first-trimester screening respectively (29).

When it comes to Asia and prenatal screening of DS, countries like China and Taiwan have second-trimester triple test and second-trimester quadruple test as the most commonly used and funded by the National Health Insurance (30, 31).

With the introduction of NIPT, which has a higher detection rate, decreases the need for invasive procedures, and can be used to detect other chromosomal defects, the use of NIPT has increased in most countries. However, the use of NIPT is either only provided to women at high risk by the government in some countries like Norway or have to be funded by the user themselves except Belgium and the Netherlands where pregnant women have the option to choose NIPT as the primary screening test (25).

1.5 Nepal-Context

1.5.1 Socio-economic and political

The socio-economic context of a country is significant for a cost-effective analysis. For example, the willingness to pay threshold of a country depends on the country's economic condition and thus affecting the result of a CEA. Similarly, a country's social and cultural aspect affects the result and conclusion of an economic evaluation. E.g., an intervention deemed cost-effective from EE can be socially or culturally inappropriate. Thus, there is a chance that intervention being cost-effective from a health care perspective but not from a societal perspective and vice versa.

Nepal, officially the Federal Republic of Nepal, is a small sovereign country located in South-East Asia with a size of 147,181 sq. km and about 30 million people. Located between two economic power houses, i.e., India and China, it is one of the world's poorer countries with approx. 1,071 USD GDP per capita (as of 2019) (32). Previously divided into three ecological regions and five developmental regions, the country was politically divided into seven federal provinces according to Nepal's constitution, adopted in 2015. Data from 2015 suggested that over 80% of the population lived in rural areas, and over 25% of the people lived in poverty (33). Known for being a country with one of the most corrupt governments globally, it has been overwhelmed by various natural disasters that have pushed the country into further turmoil. There has been a notable increase in the economy and sectors like education, health, transportation, etc., in the last two decades. However, the quality of life of people in Nepal remains low, ranking 142nd country globally in terms of Human Development Index (34).

Agriculture, tourism, and foreign employment remain the primary driver of the economy, with almost 29.1% of the total GDP coming from remittances (34).

1.5.2 Health care system and health financing

The health care system and financing also play an important role in economic evaluations. Different countries have different ways of financing their health system. They have different allocations of budget for health, which has a massive impact on funding different health services for the people, especially prenatal screening for congenital anomalies, which is not assumed to be the primary issue. It is thus important to know how much budget can be allocated for a particular health issue. Sustainability in financing a health intervention is also important and more so for a country like Nepal. Many health programs are run from external assistance for a fixed time and cannot be continued after external funding stops. Similarly, knowing the structural and functional aspects of a health care system like infrastructure, human resources, the channel of communication, hierarchy, etc., is vital in the implementation of any health intervention as it has a direct effect on various aspects like cost, accessibility, delivery, and the effects.

The development of Nepal's health care system is still young compared to other developed countries, with the first national health policy drafted in 1991 A.D. Since then, significant progress has been made in the health sector as infectious diseases like TB, HIV, and malaria have decreased considerably in the last decade. Similarly, the mortality of infants, child, and mothers has gone down, highlighting the effort put in by governmental organizations, NGOs, health workers, and all the stakeholders. However, the quality and equity of healthcare services are still the significant hurdles of Nepal's health care system.

The health care facilities in Nepal can be categorized into private and public. Public health services are delivered under the Department of Health Services (DoHS). The DoHS, under the Ministry of Health, is accountable for the delivery and management of preventive, promotive, curative, and rehabilitative health services throughout the country. The public health service is delivered through structures like health posts, primary health care centers (PHCC), and hospitals-provincial, zonal, sub-regional, regional and central. Health posts are the first point of contact for people to access essential health services. District hospitals, now called provincial hospitals, are the primary referral points in the health system and thus play a significant role in providing outpatient, inpatient, and emergency services as close as possible to the people throughout the country. Facilities from such hospitals include childcare, maternal

care, communicable and non-communicable diseases care, pathological tests, and diagnostic services like X-ray services. All the public health facilities are funded mainly through the provincial government and also the central government.

Nepal for 2020 has allocated 6.12% (approx. USD 906 million) of the total budget for health, which is an increase of 32% compared to last year, but the majority of that is meant to manage COVID-19 and its repercussions (35). Generally, the budget in health is around 5-6% of the total budget every year (36). GoN introduced a new health policy in 2019 to accelerate universal health coverage to improve the accessibility and reach of services. One of the new health policy provision was providing, along with essential healthcare services free of cost, non-basic services through social health insurance (37). This aimed to prevent people from catastrophic health care expenditures, one of the significant barriers to healthcare service utilization as the services are primarily private and centrally located. However, social health insurance covers only a few healthcare services, mainly through public hospitals, which mostly do not include specialized services.

1.5.3 Down syndrome in Nepal

There is minimal data on disability in Nepal. The first time any data was collected on disability was in the census of 2011, which reported that 1.9% of the total population had some kind of disability (38). The majority of the disability was physical (36.3%), followed by disability related to vision. Of all the persons with disabilities, only 2.9% had an intellectual disability. Intellectual disability is used as an umbrella term for conditions like Down syndrome, Cerebral Palsy, etc. Although data on congenital anomalies is limited in Nepal, it is estimated that Down's syndrome and neural tube defects, and heart defects are the most common birth defects in Nepal. Birth defects can result in newborn and child death, chronic conditions, and disabilities. The March of Dimes Report on Birth Defects estimated that 43 727 children were born with birth defects annually in Nepal. They comprise 5767 children with defects of the cardiovascular system, 3431 with NTDs, 146 with hemoglobinopathies, 1533 with Down syndrome, and 2482 with G6PD deficiency. About 6% of the total neonatal deaths in Nepal can be attributed to congenital anomalies (8).

The actual birth prevalence of DS in Nepal has not been reported yet. A hospital-based study in Nepal stated that of all the chromosomal defects, the most common was Down syndrome (26.67%) (39). DS's birth prevalence could be as high as 2-3 per 1000 live births in low- and middle-income countries as women in this area conceive at advanced stages and have limited

access to counselling, family planning services, screening, and other related services (40). A 2016 report by WHO showed that DS is 4th in the most prevalent birth defects in the South-East Asia region, including Nepal. Due to the lack of the actual number in Nepal, figures from studies conducted in India are taken as references for this model (41). DS's overall birth prevalence was reported anywhere from 1 in 925 to 1 in 1230 births. The birth prevalence for women between 35-39 years of age and 40 years or above was reported to be 1 in 304 and 1 in 64 births which is a significant increase (22). These studies were conducted in 1998, and thus the birth prevalence of DS might have gone up considering the advancing maternal age in recent years.

Thus, the birth prevalence of DS in Nepal was calculated to be 1 in 1074 births which is 0.095% of all the births. The total expected births in Nepal for 2020/21 is 752,506. This equates to approx. 715 births with DS this fiscal year.

1.5.4 Screening of Down syndrome in Nepal

Mostly in low- and middle-income countries (LMICs), infectious diseases and lack of maternal health services are the major causes of neonatal and infant mortality. Genetic screening for birth defects is not perceived as having enough significance (42). Thus, the government does not allocate resources to set up such services. Disabilities may lead to increased expenditure for individuals and households (43-45). The increased cost of living may be due to health care, food, the need for special care, assistive devices, etc. (46). The presence of birth defects also may lead a woman incapable of conceiving again (47, 48). Prenatal screening and diagnosis during pregnancy can help a couple decide to either choose to terminate the pregnancy or go on with it if they have a fetus with a defect. It will help them plan the birth, which will reduce mortality. Moreover, the couple will have enough time to understand the condition better and be prepared.

There is no prenatal screening or diagnostic system for detecting birth defects during pregnancy in Nepal as a part of the national health system. There are few private hospitals and health centers that provide such services but are generally out of reach of the people due to the location and cost. The cost of such services is high and is not covered by the government. Thus, this study thus aims to evaluate the cost-effectiveness of introducing screening and diagnosis of trisomy and neural tube defects within the public health system, which will assist health policy makers in evidence-based decision making.

1.6 Objectives

1.6.1 General objective

The general objective of the thesis was to:

1. To assess the cost-effectiveness of prenatal screening for detecting DS among pregnant women of Nepal compared to no screening from a health care perspective.

1.6.2 Specific objectives

The specific objectives of the thesis were to:

1. Develop a decision-tree model to calculate the costs and outcomes of all the seven interventions in three different age groups to assess the cost-effectiveness through CERs, ICERs, NMBs and NHBs at the current WTP threshold.
2. Examine the safety of all the interventions with the calculation of the number amniocentesis and foetal loss due to amniocentesis for each intervention.
3. Determine the influence of varying values of different parameters in the final result through sensitivity analysis.
4. Identify if incorporating parameter uncertainty in the model produces different result than deterministic analysis through probabilistic analysis.
5. Develop a special scenario with introduction of coinsurance to decrease the cost of test and assess the cost-effectiveness of all the interventions.
6. Identify the parameters which affect the decision uncertainty the most and assess if collection of additional information on those parameters/group of parameters are worthwhile or give additional benefits through VOI analysis.
7. Calculate the total budget required for all the interventions to be implemented over the course of five years.

Chapter II: Theoretical background

This part of the report consists of theoretical concepts that are used in the model.

2.1 Economic evaluation

Increasing demands for health care and limited resources to fund these needs means decisions have to be made on which health programs to implement and which to forego. Clinicians are primarily concerned with the health effects of a specific healthcare intervention. These interventions are rarely cheap. In an ideal scenario of having unlimited resources, the intervention with the highest clinical efficacy should be the intervention of choice. However, in the real-world scenario, a decision has to be made on one of those. Economic evaluation tackles this problem by helping decision-makers choose the intervention that represents the best value for money or maximizes the benefit accrued within the budget limit. It helps to identify the best option from all the relevant alternative courses of action (49, 50). It is more than just mere cost minimization.

An economic evaluation measures two parameters, i.e., cost and outcomes or effects. Cost includes the values of all the tangible resources used for the intervention to take place from the capital, human resources, opportunity costs, etc. (51, 52). Drummond et al. have categorized the costs consumed in healthcare intervention into: health care costs like equipment, drugs, physician costs, etc.; costs for patient and family such as out of pocket payment for services, productivity losses due to the disease, etc.; cost for other public agencies which are affected by the disease or the intervention; and finally the opportunity cost which is the cost foregone by choosing one alterative over the other (50). The selection of costs depends upon the perspective of the economic evaluation, i.e., societal or health care perspective. On the other hand, the outcomes are measured in different capacities and can be expressed in terms of QALYs gained, the number of life-years gained, DALYs averted, etc., depending upon the type of the economic evaluation. Since two parameters i.e. costs and outcomes of the interventions are measured in an economic evaluation, conclusions are made by analyzing both the results. This means that the cheapest option is not always the best option. There might be a case where the cheapest intervention has the highest health effects. In such a case, the choice is quite straightforward and the cheapest intervention will be the most cost-effective (53).

However, the case mentioned earlier occurs very rarely. Mostly, increased health gains come with increased cost and decision has to be made by analyzing the increase in the cost relative to the health gains. If an intervention has increased costs but result in high health gain, the

intervention will be preferred. But if an intervention has increased cost with lower health gains than the comparator, it will not be preferred.

Cost-effectiveness analysis (CEA), one of the three main types of economic evaluation, is generally used for comparing interventions within patient groups with the same disease or health issues, the outcome is measured in natural units like number of life-years gained, number of cases detected, etc. Results from CEA can be expressed as the cost per unit of effect. For instance, cost per life-years gained or cost per number of cases detected, etc. CEA is most useful when there are limited information and a limited range of options with a limited budget. Although useful, the use of outcomes such as QALYs gained as measures of outcomes is thought to be better, leading to the next type of economic evaluation (54, 55).

Cost-Utility Analysis (CUA) is used when utilities are preferred to measure the outcomes or the effects of the intervention rather over natural units. The outcomes are utilities and depend upon an individual's preferences, a group, or a community. Results are primarily expressed in terms of cost per QALYs gained by undertaking one intervention over another. One advantage of CUA over CEA is that interventions that do not have the same health effects can be compared. This type of analysis also incorporates the quality of health and the quantity that is not easily accomplished in CEA (55, 56).

Lastly, Cost-Benefit Analysis is a type of economic evaluation which measures both the cost a consequence in monetary values. CBA can also compare interventions in different fields like health and education and identify the most beneficial intervention. The results are expressed as a cost-benefit ratio or the net benefit of the intervention. Although the ability to conduct cross-sectoral comparisons, the task of assigning monetary values to health benefits can be challenging (55, 57).

The time horizon for both costs and consequences of the interventions used in economic evaluation can be different according to the study and perspective. Existential time preference means there is variation in the values attached to cost and consequences now compared to past or future. Thus, it is necessary to discount both cost and consequences if the analysis involves a time horizon of more than one year. This is done using a standard discount rate: either a predetermined national rate or a rate of 5% used in most scientific literatures (50). WHO recommends using a 3% discount rate for both costs and consequences, with recommended testing from 0% to 6% for the sensitivity of the result (58).

2.2 Decision analysis

As documented earlier, decision-making is an essential part of health care. It involves choosing an alternative after the risks, benefits, and costs of all the interventions are systematically compared. In such a case, there is uncertainty associated with the decision, and the degree of uncertainty depends on the availability and validity of clinical and economic data. Decision analytic modeling is a systematic tool and approach to making a decision under uncertainty and is widely used in the economic evaluations of health care technologies and services (59).

Models in health care can be defined as "an analytic methodology that accounts for events over time and across populations, and that is based on primary and/or secondary data sources, and estimates the effects of interventions in terms of health effects and costs (60)." Decision models rely on expected values. Once all the decisions to be made are identified, all the possible sequence of events and outcomes are listed; probabilities for every sequence of event happening are assigned; values of all the outcomes are assigned, and the expected values of every possible strategy are calculated. Lastly, the assumptions or the input values are changed within a range to see the impact of the change on the results (60, 61).

Markov models and Decision trees are the most commonly used decision models in economic evaluations (50). Markov model is probably the most common among the two and uses states of diseases or health issues to represent all the intervention's possible outcomes. All the states are mutually exclusive, meaning an individual can be in only one state at a point in time. Individuals move between the states as the condition changes over time with specific transition probabilities (62, 63). The cost and outcome associated with every state for an individual are determined by the time spent in that state. The costs and outcomes are combined for a cohort of people over multiple cycles individually for all the interventions and compared at the end.

Decision trees are basically diagrammatic representations of all the possible sequences of events and outcomes. It starts with a decision node, typically represented by a square, where multiple alternatives are possible. The decision node leads to chance nodes, typically represented by a circle, which leads to various events, and the probabilities determine the chance of an event happening. The routes of all the mutually exclusive sequence of events are called pathways which end at triangular nodes. Each branch has cost and outcome associated with it (64, 65). The expected cost and outcomes of each pathway are obtained by multiplying the costs and outcomes with the probability of ending in the particular pathway (60). A simple

decision tree with decision node, chance node, pathways, and probabilities is shown in the figure below.

Decision analytic models can be both deterministic and probabilistic. In deterministic models, point estimates (fixed values) are input parameters with the models' output also measured as point estimates. However, probabilistic models use probability distributions of all the parameters as input parameters and produce output as the distributions of costs and outcomes.

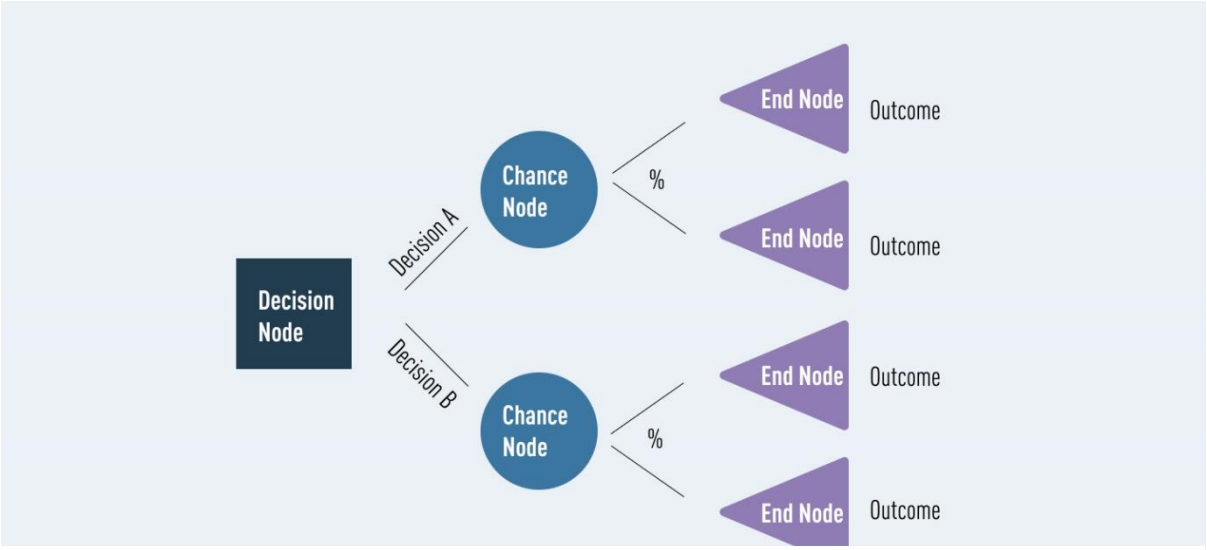


Figure 1 Simple decision tree

Doing so incorporates the uncertainty of parameters and propagates the uncertainty through the model from input to output. Different distributions (gamma, beta, Dirichlet, etc.) are used in probabilistic modeling depending on the input parameters (60). For instance, beta distribution is used for parameters that have values between zero and one. For e.g., probabilities, incidence, relative risks, etc. Similarly, for situations where three or more probabilities are involved in a single path probability, Dirichlet distribution is used instead of beta. Gamma distribution is used for the calculation of 1000 values of cost parameters.

2.3 Analysis and presentation of the result

The alternative options in economic evaluation are compared in terms of incremental costs and effects, expressed as incremental cost-effectiveness ratio (ICER), defined as additional cost per additional unit of effect from the more effective treatment (60).

Incremental Cost-Effectiveness Ratio (ICER) is a concept which expresses result of the additional cost needed for an additional unit of outcome or effect. This informs the decision-makers which intervention to choose. It is defined by the ratio of the difference in cost between two alternatives and the effects of those interventions. It can also be defined as the mean cost

associated with a one-unit additional increase in the desired effect. If the ICER is less than the willingness to pay threshold, the intervention is generally undertaken (66).

ICER can be defined mathematically as:

$$ICER= (C1-C0) / (E1-E0)$$

Where C1 and C0 are the cost of intervention of 1 and 0; E1 and E0 are the effects of intervention 1 and 0 respectively.

As discussed earlier, the results of economic evaluations are presented in terms of ICER. ICER is calculated by ranking all the interventions according to the cost incurred. When two interventions are compared, if the new treatment is compared to the benchmark incurs less cost and produces more health effects; it is considered dominant, thus more cost-effective and chosen. If the new interventions are more costly and produce more health gains, making decisions depends upon the threshold of willingness to pay. If there are more than two comparators, the intervention having higher ICER than the subsequent most costly intervention is said to be extendedly dominated. The interventions which are either strongly or extendedly dominated are excluded, and evaluation is done among the non-dominated interventions. The ICERs are presented in Cost-effectiveness (CE) plane. The line joining non-dominated alternatives is the CE frontier, and all the alternatives lying below the frontier are considered to be not cost-effective. In probabilistic analysis, there are multiple estimate plots of the costs and effects in the cost-effectiveness plane (50, 60).

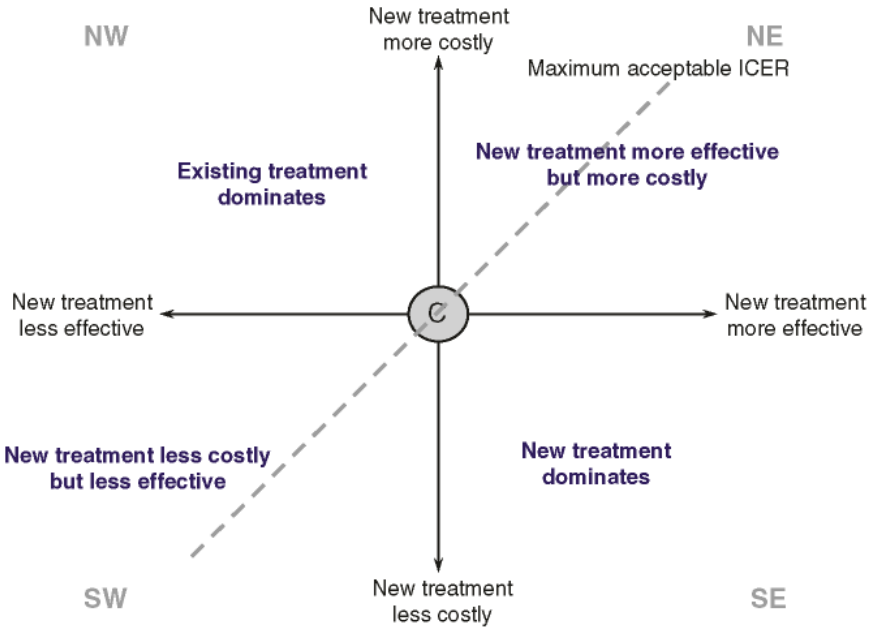


Figure 2 Cost effectiveness plane

Figure 2 represents an example of a CE plane. The X-axis represents the effects, and Y-axis represents the costs. The intervention can fall in either of the four quadrants depending upon the costs and effects. The intervention that falls in the southeast quadrant is dominant as it costs less and has more health effects. The northeast quadrant includes the more costly intervention but also with higher health effects. The intervention in the northwest quadrant are more costly and have lower health gains are not chosen. Lastly, the southwest quadrant includes the intervention that produces fewer health gains but is less costly. Generally, the interventions in SE are chosen followed by those in NE, but the decision depends on the WTP threshold (50, 60).

Another simple way deciding to choose an intervention over the other from an economic evaluation perspective without calculating ratios is the calculation of net benefits of all the interventions (67). There are two standard measures of net benefits: net monetary benefits and net health benefits.

Net monetary benefit is the net monetary gain of adopting a certain intervention in monetary terms. It has two parts: 1. health gain from the adoption of the new intervention in monetary terms and 2. monetary loss associated with the intervention (which includes opportunity loss) (67, 68). It is calculated as:

NMB= WTP threshold*health gain (from the intervention)- cost (of the intervention)

Net health benefit, on the other hand, is the net gain of adopting an intervention in health terms. Like NMB, this has two parts: 1. health gain for patients who receive intervention and 2. health loss experienced by other patients in the population which is a result of forgone opportunity to provide other care services to other patients (67, 69). It is calculated as:

NHB= Health gain from the intervention- (cost of the intervention/WTP threshold)

The intervention with the highest NMB or NHB is the most cost-effective and should be the preferred choice.

2.4 Uncertainty analysis

There is some degree of uncertainty associated with every economic evaluation or decision-analytic model. Uncertainty may be present in input parameters as the values of such parameters are drawn from multiple sources. The decision to choose an intervention is hugely dependent on the uncertainty present in the model. This makes it very important to incorporate those uncertainties in the model as much as possible. Failure to include them might mean a

wrong decision being made, and there are consequences in terms of costs and benefits foregone. Generally, there are two types of uncertainties in decision modeling: Structural or model uncertainty and parameter uncertainty. Structural uncertainty arises when there is uncertainty related to the structural assumption of the model. On the other hand, Parameter uncertainty refers to the uncertainty in the estimation of the input parameters. There are various reasons for uncertainty like methodological differences in the analysis of different studies, extrapolation, etc. (60, 70).

There are various ways of dealing with uncertainty.

2.4.1 Sensitivity analysis

One way of dealing with uncertainty is by using a concept called sensitivity analysis, where the values of input parameters are varied in order to assess how the change in the values affects study results. Sensitivity analysis can either be a one-way sensitivity analysis in which each parameter is varied individually to study its impact on the result. It can also be a multiway analysis in which the number of parameters (when two parameters are used, it is called two-way sensitivity analysis) are varied at the same time. This is done to address the fact that these parameters can be correlated and affect the results together. Scenario analysis is another method of dealing with uncertainty in which the lowest, base case, and highest values of the input parameters are chosen, and the results are assessed. Thus, a series of scenarios representing a subset of the potential multiway sensitivity analysis consisting of the base case, best-case, and worst-case scenario are constructed (50).

2.4.2 Probabilistic Analysis

Generally, the results of economic evaluations are expressed in terms of ICERs obtained through deterministic analysis. In such traditional analysis, it is assumed that all the input parameters' values are precisely known, which is mostly not the case. Thus, to reduce the uncertainty within the input parameters, the researcher may assign different distributions to the parameters and generate random draws to incorporate the uncertainty and migrate it to the outcome using the Monte-Carlo simulation. This is run many times, usually 1000 times. This technique is known as Probabilistic (Sensitivity) Analysis. ICER from every random draw is calculated and stored. All the ICERs are plotted in CE plane, and uncertainty can be summarized by how many iterations fall below or above the WTP threshold (50, 60, 70).

Uncertainty in PA can also be presented using net benefit (net health or net monetary benefit). The proportion of times a strategy has the highest net benefit among the interventions provides

evidence that the strategy is either cost-effective or not. These values of proportion, when plotted for a range of WTP thresholds, a curve known as Cost-effectiveness acceptability curve (CEAC) is obtained. This curve shows which intervention is cost-effective at a certain WTP threshold(71). A curve, known as Cost-effectiveness acceptability frontier (CEAF), can also be plotted, showing the point or value of the WTP threshold from which the decision changes (72).

2.5 Value of information analysis

Economic evaluation aims to improve decision-making and decision modeling to determine whether resources should be used for particular interventions based on available evidence by providing an estimate of expected costs and effects. Due to uncertainty associated with input parameters, the resulting distribution of expected costs and effects can also be considered decision uncertainty. In other words, it gives the probability that the given decision is the correct one or wrong one. There is cost associated with a wrong decision, and hence value can be given to the reduction of uncertainty through obtaining more information. This is based on the notion that policy changes are costly and may be difficult or impossible to reverse. There should be a value given to the option of obtaining more information that can reduce uncertainty. CEAC obtained from PSA will also give the cost of making a wrong decision by combing the cost of wrong decision and the probability of making the wrong decision. This value is known as the Expected Value of Perfect Information (EVPI). EVPI represents the maximum value of future research against which the undertaking of a particular study can be compared. If research could remove uncertainty, its value would be the cost of uncertainty. The value of information research implies that it helps decide how much research is efficient to undertake for any particular decision. Any additional research is efficient economically only if it costs less than its EVPI (50, 60).

$$\mathbf{EVPI} = E_X\left[\max \mathbf{NB}\right] - \max\left[E_X \mathbf{NB}\right] \qquad (1)$$

where the first part of the equation represents the expected value of maximum NB with perfect information about all the parameters and second part of the equation represents the maximum expected NB with the current information about the parameters.

Choosing the estimations from economic evaluations to make health care decisions might mean huge consequences if a wrong decision is made, which is possible due to all the uncertainties in the parameters measured. Thus, it is necessary to collect every bit of information possible to reduce the decision uncertainty. However, collecting more information comes at a cost in terms

of time and money. VOI analysis is one method that helps decision-makers to make decisions on adopting the new technologies and assess the need for further research simultaneously (73).

Expected Value of Perfect Information (EVPI) is the value of removing all the uncertainty from the analysis in monetary terms. This is the price that decision-makers are willing to pay to gain perfect information. EVPI is the first step in VOI analysis and tells us if further research is valuable or not (74).

Expected Value of Partial Perfect Information (EVPPI) is the value, again in terms of monetary terms, of collecting full information about particular parameters of interest. EVPI thus reduces the uncertainty but does not eliminate it. EVPI is used to assess if collecting information on certain parameters is worthwhile or not (75).

There are many ways of calculating partial EVPI. As explained by Edward Wilson in his paper titled "A practical guide to VOI analysis", the traditional way of calculating EVPPI or partial EVPI is using nested Monte-Carlo simulation consisting of two loops, i.e., the outer loop and the inner loop (76). First, a value of the target parameter or group is sampled, which represents the outer loop. This represents one of the many possible values of the target parameter. Then, using the sampled value, values of other remaining parameters are sampled from the probabilistic model. The subsequent NMB of all the intervention in the iteration is recorded, and this is repeated for many iterations (e, g, 1000-5000). This is the inner loop. The outer loop is repeated with new values of the target parameter every time and done for many iterations. EVPPI is then calculated by subtracting the maximum expected NB with current information from the expected maximum NB (maximum NB from every iteration- perfect information of target parameter assumed). This can be done for every parameter or group of parameters. EVPPI is calculated using the Monte-Carlo simulations as:

$$\mathbf{EVPPI}\ (X_i) = E_{X_i}\left[\ \max_d E_{X_{-i}|X_i}\ (\mathbf{NB}\ (d, X_i, X_{-i}))\ \right] - \max_d E_X\ (\mathbf{NB}\ (d, X))\ \quad (2)^{i}$$

where the first part of the equation is the outer loop and the second part of the equation is the inner loop as explained above. Also, *d* is the decision option, X_i is the vector of input parameters we wish to calculate EVPPI for and $\mathbf{X_{-i}}$ is the vector of remaining parameters.

However, the traditional method of calculating EVPPI is computationally intensive and requires time due to difficulties in estimating the first term in the equation above. Thus, various ways of calculating EVVPI have been developed in recent times with similar accuracy in significantly less time than the gold standard nested Monte-Carlo simulation.

One of the various methods used is the regression-based method which has been used in this study. As explained by Strong et al., the basic idea of the regression model is to avoid the calculation of the inner loop by reframing the estimation of the expectation of the model as a regression problem (77). The model output is expressed as a sum of conditional expectation and a mean zero error term for "n" simulations of PSA.

$$\mathbf{NB}\ (\boldsymbol{d}, \mathbf{X}^{(n)}) = \boldsymbol{E}_{\mathbf{X\text{-}i|Xd=xd}^{(n)}}\ (\ \mathbf{NB}\ (\mathbf{d}, \mathbf{X_i})^{(n)}, \mathbf{X_{\text{-}i}})) + \boldsymbol{\varepsilon}^{(n)} \qquad \mathbf{(3)}$$

where the first term is the conditional expectation and second term is the error term.

As the expectation term takes a different value for each value of the target parameter, it can be thought of as a function of x_i, as seen in the equation below, representing a regression problem.

$$\mathbf{NB}\ (\boldsymbol{d}, \mathbf{X}^{(n)}) = \boldsymbol{g}\ (\boldsymbol{d}, \mathbf{X_i}^{(n)}) + \boldsymbol{\varepsilon}^{(n)} \qquad \mathbf{(4)}$$

where the conditional expectation from equation iii has been expressed as a function *g* of *d* and X_i.

However, no specific assumptions are made about the function. Finally, the N model outputs where n=1, …., N from PSA can be treated as data through which the functional form of the function can be learned. This regression estimate will provide a precise estimate of the expectation, and EVPPI for a target parameter can be calculated. The final equation is shown below.

$$\mathbf{EVPI}\ (\mathbf{X_i}) = \mathbf{1}/\boldsymbol{N}\ \textstyle\sum^{N}_{\mathbf{n=1}}\ \mathbf{max_d}\ \boldsymbol{g}\ (\boldsymbol{d}, \mathbf{X_i}^{(n)}) - \mathbf{max_d}\ \mathbf{1}/\boldsymbol{N}\ \textstyle\sum^{N}_{\mathbf{n=1}}\ \boldsymbol{g}\ (\boldsymbol{d}, \mathbf{X_i}^{(n)}) \qquad \mathbf{(5)}$$

where the first term and second terms are the estimators of the first and second term from equation ii obtained from non-parametric regression.

The collection of perfect information of few parameters, let alone all, is not very realistic. Thus, information is collected from a sample study about the parameters. This additional information reduces the uncertainty of the parameters rather than eliminating it. Therefore, EVSI helps determine the optimal research design, like the optimal sample size, study population, etc., to reduce the decision uncertainty.

---

[i] Equation 2, 3, 4 and 5, and subsequent explanation are derived from the paper "*Estimating Multiparameter Partial Expected Value of Perfect Information from a Probabilistic Sensitivity Analysis Sample: A Nonparametric Regression Approach*" by Strong et al.

2.6 Budget impact analysis

Decisions made at present have long terms consequences. Decision-makers thus have the responsibility to make the scarce resources are utilized efficiently and effectively. They have so many decisions to make in terms of which health care services to be provided to the people within the budget limit. An intervention known to be cost-effective does not necessarily mean that it is viable for years to come as there are various factors involved. Therefore, it is vital to know the impact of all the interventions on the annual budget through Budget impact analysis (BIA) and conventional CEA (78-80).

BIA, used as a supplement to CEA, measures the cost of introducing a new technology or intervention in a particular healthcare environment for anyone in the population with a disease in a given year, regardless of how long they have had the disease. Many of the data elements and methodological criteria as shared with CEA, but the scope and the reporting vary from CEA. The size and characteristics of the affected population, cost of new and current intervention, impact on other health conditions, and treatment-related health care services are essential inputs for BIA(79). When conducting BIA, time horizon is so chosen which is relevant to the budget holder should be in line with budgeting process (78-80).

2.7 Safety Index

Safety index is defined in terms of the ratio of fetal loss due to amniocentesis associated with an individual test and the total number of cases detected. The lower the safety index, the better is the performance of a strategy in terms of safety (81, 82).

$$\textbf{Safety index} = \frac{\textbf{No. of fetal loss d/t amniocentesis (of a particular strategy)}}{\textbf{Total number of cases detected (by the same strategy)}}$$

In addition to the CEA, results like the safety index can be useful in assessing all the alternatives before making the decision. For example, some societies and health care systems might prioritize health consequences over cost-effectiveness while making decisions. Thus, the safety index provides the health consequence of the number of fetal loss due to the screening that would have otherwise not occurred.

2.8 Harm to benefit analysis

Harm to benefit analysis is another analysis concerned with the safety of the screening strategy and relates to the number of amniocenteses performed. Unlike the safety index, this does not consider the number of fetal loss but measures harm in terms of the number of amniocenteses

performed and benefits in terms of the number of cases detected. Incremental harms and benefits are calculated and plotted in an incremental harm benefit plane.

$$\textbf{Incremental harm-benefit ratio} = \frac{\textbf{Incremental harm (number of amniocenteses performed)}}{\textbf{Incremental benefit (Total number of cases detected)}}$$

The lower value of the incremental harm-benefit ratio refers that the strategy is more beneficial than harmful (83).

Similar to the safety index, this analysis also provides a different perspective of assessing the interventions. The number of amniocenteses performed has a direct effect on the cost of the screening cost per woman. It also affects the outcome in terms of the number of cases detected and the number of fetal loss per woman. Using information from ICER and incremental harm to benefit ratio could be more useful than using ICER only to choose one intervention over the other. An intervention with the lowest value of ICER and increment harm to benefit ratio would be the best choice.

Chapter III: Literature review

This chapter contains a summary of all the relevant literature that was found in the process of writing this thesis. This chapter is divided into four sections. The first section contains the premise behind selecting the topic. The methodology section contains methods and tools used to search the literature like the search engines, keywords used to search the literature, etc. The third part contains the summary of results of the studies all over the world related to the topic of interest. The fourth part contains the studies' critical analysis, followed by the last part, highlighting the gaps in the literature and providing justification for the thesis.

3.1 Background

Having worked in the field of disability in Nepal for over two years, I wanted to incorporate economic modeling and disability for my master's thesis. Disability is one area in health that is given very little time and resources with a significant focus on maternal and child health, NCDs, etc. Thus, I started looking for studies that incorporated disability and economic evaluation both in Nepal and worldwide. The search ended with hundreds of finds, and most of them were related to prenatal screening of DS. After going through a few of them, I found out that most developed countries had policies regarding the screening of DS and other congenital defects. However, there was no such policy in Nepal for any congenital defects, let alone DS. Thus, I wanted to evaluate through economic modeling if the introduction of such screening interventions would make sense in the current scenario.

3.2 Methodology

After the selection of the topic, I started a literature search on the topic. The search was mostly done using "PubMed" and occasionally using "Google scholar." The search terms used were: "Economic evaluation and Down syndrome and prenatal screening," "Cost-effectiveness and Down syndrome and prenatal screening," "(Birth prevalence or Incidence) and Down syndrome," "(Birth prevalence or Incidence) and Down syndrome and Nepal," "Fetal loss due to amniocentesis," and "Down syndrome and prenatal screening." The main search terms were "Cost-effectiveness and Down syndrome," and other searches were used to supplement the preliminary findings. Since there were hundreds of studies done, studies before the 2000s were excluded (included only if necessary) from the review unless necessary. For instance, the closest to finding data on the birth prevalence of DS in Nepal was a study conducted in 1998. The relevant articles were selected after reading the abstract.

3.3 Literature about DS and screening

3.3.1 Down syndrome

As discussed earlier, DS is one of the most prevalent congenital birth defects. DS has become more common due to few reasons: the average maternal age going up in the past decade, availability of screening and diagnosis of DS, and children with DS living longer than they used to before (84). However, the incidence has not gone up dramatically with the practice of termination of the fetus with DS as expected. Moreover, a study in England showed that the rate has been reasonably constant over the years (85). The birth prevalence of DS varies from country to country. However, it was estimated that DS affects about one in every 750-1100 live births in general (84). According to the annual report of International Clearinghouse for Birth Defects Surveillance and Research of 2012, DS occurred in approx. 1.7 per 1000 live births in the world (86). A report from the EU states that the rate increased from an average of 16 per 10,000 live births in 1993 to 23 in 2015 in Europe (87). In the US, the figure is on the higher side, with every baby born in 700 will have DS (88). However, the lack of valid and reliable data sources meant that the incidence of DS that can be generalized for low- and middle-income countries in Asia is still unknown. Countries like Taiwan, Korea, China, and Japan have definite figures related to DS. For instance, the incidence of DS in China was reported to be 1.47 per 1000 babies in 2012, which is equivalent to 23,000-25,000 new cases every year (89). Similarly, a report dating back to the late 1990s shows that the incidence of DS in India was 1 in 1000 on average (22). These were derived from tertiary level hospital data and are not population-based and technically might not represent the general population. However, these are the single largest data available for the region.

3.3.2 Risk factors

Advanced maternal age and genetic causes are the two most important risk factors for DS. Various other risk factors have been studied but are found to be very weakly associated. The incidence of having a child with DS for women of 25, 30, 35, and 40 years are estimated to be 1 in 1339, 1 in 938, 1 in 352, and 1 in 85 live births (19). So, the risk increases three times for 35 years old and 15 times for 40 years old women. A hospital-based study in India showed the increase of odds from 1 in 1198 to 1 in 64 for women over 40 years of age compared to 25-29 years (22). Down syndrome can be genetically transmitted from parents to the child but only seen in translocation. The error in meiosis generations before the baby, maternal grandmother to be specific, might manifest as DS in generations later (90). Similarly, consanguinity, the residence of parents (rural/urban), exposure of parents to chemicals, educational status of the

parents, and the mother's reproductive performance are assumed to be the possible risk factors for DS (18). However, more large-scale studies are necessary to understand how these environmental factors interact with genetic factors.

3.3.3 Screening tests: Overview of the history

The history of prenatal diagnosis for DS dates back to the 1960s when amniocentesis was used to diagnose DS directly without any prior screening (12, 91). A cut-off age of 35 to 37 years was used to provide women with amniocentesis at 16 weeks of pregnancy. The practice of prenatal screening started in the 1980s when trisomy of chromosome 21 was reported to be the cause of DS by Merkatz and his colleagues. The first screening test used was based on the single maternal serum marker, i.e., Alpha-fetoprotein (AFP), the value of which was found to be 25% lower in affected pregnancies. Similarly, it was found that levels of hCG were found to be, on average, higher in pregnancies affected by DS. Later in 1988, three maternal serum markers were used, i.e., AFP, uE3, and hCG, together with maternal age, to determine the risk of having DS-affected pregnancies, which were known as “the triple test” and conducted within 15 to 22 weeks of pregnancy. The detection rate for the triple test was 60%, while the false-positive rate was 5% (12).

A new test was introduced in 1996, which used one more maternal serum marker called inhibin-A in addition to those used in the triple test. This test had an increased detection rate of 76%, with quite a high 9% FPR than other existing tests (12).

A test with PAPP-A and free B-hCG started in 1996 which was done in the first trimester and known as “the double test.” The test had a detection rate and FPR of 62% and 5% respectively. As ultrasound was found useful in the detection of increased space in the back of the fetal neck between the spine and skin, also known as increased nuchal translucency, it was combined with serum markers from the double test and was called the combined test. This test is the most widely used globally, with a high detection rate of 85% and 5% FPR. Finally, in 1999, the first-trimester combined test and the triple test were integrated and started being known as "the integrated test” with a detection rate of 85% and FPR of 0.9%, which is more common in the US. Thus, it can be seen that the development of different screening methods was very rapid and took place within 15-20 years (12, 91).

3.4 Literature analysis

There have been multiple studies comparing the cost and effects of various screening programs in different settings. Some of the studies are discussed below:

The use of maternal age as a screening strategy for DS is not an effective strategy. The high false-positive rate (and even more for the older population) would mean a high number of invasive procedures which might not be needed at all (19). It is well established that the increased nuchal translucency measurement is a strong indication for DS with a 71.2% detection rate and 4.6% false-positive rate (92, 93). G.D. Michailidis et al. concluded that first-trimester NT measurement followed along with maternal age could be an effective strategy. Addition of second-trimester maternal serum markers would add to the detection rate. However, second-trimester screening means a delay in the diagnosis, extra visits, and costs. Thus, NT measurement and maternal serum markers measurement in the first trimester seems to be the right choice instead of the second trimester (94).

Mirjam Hoogendoorn et al. conducted a cost-effectiveness analysis of six different screening methods in a hypothetical cohort of Dutch pregnant women (n=100,000). They incorporated NTDs along with DS in their model. The interventions included ultrasound and different maternal serum markers and used a cut-off risk of 1:250 to label a woman as at risk. Termination of pregnancy after DS diagnosis was 95%. The sensitivity and specificity of amniocentesis were both assumed to be 100%. It was reported that the cost per case of DS detected was lowest for first-trimester double test and triple test, which was 100,000 euros, followed by the combined test at 176,000 euros per case detected. The figure for the double and triple test was 1.8 times lower compared to other tests. The number of cases detected were highest for invasive testing and combined testing but incurred higher cost (95).

Furthermore, the inclusion of NTDs decreased the cost per case detected to 73,000 for the triple test with similar results for other tests. The number of cases detected and the cost varied according to maternal age. As maternal ages increased, the number of cases detected went up, and the cost per case decreased. However, the number of miscarriages increased too.

A study of data from a tertiary medical center in Taiwan assessed the CE of the triple test compared to the double test and maternal age only among pregnant women. Logistic regression and a cut-off point of 1:270 were used to conduct the analysis. The authors only considered the direct costs related to screening tests and termination of pregnancy. It was found that the average cost per case detected was $14,561, $42,376, and $37,424 for maternal age screening, double test, and triple test respectively, which shows that the triple test was cost-effective than the double test. The incremental cost of adding an extra serum marker (uE3) in the double test was $15,199 per case detected. However, the ICER increased to $20,980 when only 80% of

the women who tested positive went for amniocentesis compared to 90% in the base case. Therefore, the cost of amniocentesis was very important as pregnant women had to pay it for themselves, and an increase in the cost meant fewer women were accepting the test. These two studies' findings are in line with other studies that have reported the triple test to be the better choice as a screening method for DS(96, 97).

However, some studies suggest the opposite, i.e., the double test is more cost-effective than the triple test. If we look at the report on the prenatal screening policies in Europe, most of them use the CUB test as the primary choice.

William Cusick et al. compared and reported that the first-trimester combined screening was more cost-effective than second-trimester triple screening from a health care perspective. They reported that the CUB test was associated with lower screening costs, fewer live-born DS cases, and lower total cost than triple screening. The higher sensitivity and lower false-positive rate of the CUB test meant almost all the women who had to go through amniocentesis were almost half compared to the triple test. Thus, the resources could be allocated to the patients with a higher risk of having a fetus with DS. One advantage that the triple test has over the CUB test is that it can screen NTDs, decreasing the total cost per case detected. However, the cost-effectiveness remains. The problem with tests such as CUB is the availability of professional sonographers as they can be limited in numbers(98).

A cost-effectiveness analysis of singleton pregnancies in Turkey reported that the CUB test is less costly than the triple test or NIPT for women under 35 years of age. They also reported a strategy: with CUB test offered to all the women under 35 years of age, the high-risk pregnancies considered for further screening by NIPT, and the women identified as NIPT positive offered amniocentesis was more reasonable. It would result in a lower number of invasive procedures. It would also mean that the cost of amniocentesis would be lower. However, there will extra cost associated with NIPT. NIPT alone as a screening strategy would increase the efficacy but would incur very high costs. Only direct medical costs were included in the study, and the cost associated with live-born DS cases and other societal costs have not been included(99).

B Li et al. assessed the feasibility of the CUB test in the resource-limited setting in Mainland China. The study was done with the data from 10,442 pregnant women who went through the CUB test. The women who were classified as at high risk (610 women or 5.8%) were then given a choice of either: diagnostic procedure (44.9% chose this) or NIPT (27.7% chose this).

The cost per case detected when CUB was chosen as the screening test was RMB596,686 compared to RMB 1.79 million if all the women had gone through NIPT as the primary screening test. One of the major findings reported in the study was the acceptance rate of follow-up tests, with 30% of the women electing not to, perhaps because women had to pay themselves for the follow-up tests. The acceptance rate would presumably be higher if the state provided these tests for free (100).

Although most commonly used, the combined test does not seem to be more cost-effective all the time. One of the major hurdles in issuing the test is the availability of certified sonographers. Michael Christiansen and Severin Olesen Larsen proposed a contingent strategy where all the women are provided with a first-trimester double test. Women with risk more than 1:1000 were provided with NT, and those below the risk were informed that risk was too low to warrant any further investigation. Those with high risk (>1:65) were provided with CVS or amniocentesis. Following NT, women are provided with either a diagnostic procedure (CVS or amniocentesis for total risk >1:400) or NIPT for total risk <1:400. Contingent testing was cost-effective compared to normal double tests with 53,000 pounds per DS not born and 62,000 pounds, respectively. It was further concluded that contingent screening could be a possible way to introduce first-trimester screening to all the women, especially where NT cannot be offered with optimal quality. The choice among first and second-trimester screening is not straightforward. While second-trimester screening brings more efficacy, diagnosis at later stages of pregnancy can be risky (101).

Anthony M Vintzileos et al. performed a cost-benefit analysis of the American and the British strategy for prenatal screening of DS from a societal perspective. The American strategy only includes second-trimester maternal serum screening and maternal age, followed by amniocentesis for those deemed to be at risk. The British strategy includes both first-trimester NT measurement or second-trimester maternal serum screening along with maternal age. Those at risk would then go through either CVS (first-trimester) or amniocentesis (second-trimester). Compared to no screening, the American strategy was cost-saving, with about $96 million saved. The British strategy was cost-saving only in the best-case scenario with a mere $5 million saved. The American strategy had a better benefit-to-cost ratio, i.e.1.15 compared to the British even at best-case, i.e.1.01. The British strategy was financially comparable to its counterpart only when first-trimester ultrasound had a sensitivity of 80% and FPR of 5% in detecting DS (102).

Jean Gekas et al. compared the cost-effectiveness of different strategies in Canada's context by using empirical data from trials of 110,948 pregnancies in Canada. Commonly used tests like the combined test, triple test, integrated test, etc., were used. Another strategy used was contingent screening, where first-trimester screening was used as the primary test. Women with high risk went through the diagnostic procedure, and those with moderate risk went through second-trimester screening followed by diagnostic tests if positive. They found out that contingent screening was the most CE with CE ratio of Can$26,833 with ICER of Can$3815 compared to the second most CE strategy (integrated screening). The combined test as a screening strategy had the highest CE ratio of all (Can$47,358) with the highest number of procedure-related miscarriages (103).

Gilbert et al. analyzed the CE of different strategies (Combined test, integrated test, quadruple test, NT, etc.) in the United Kingdom from a health care perspective. The integrated test was found to the most cost-effective and safest strategy of all. The ICER of the integrated test compared to NT, which was the next CE strategy (with ICER of 22,000 pounds per case detected compared to no screening), was 52,000 pounds per case detected. Regarding the safety of the strategies, all the other tests resulted in more live-born babies with DS and more miscarriages of unaffected pregnancies due to invasive procedures. The cost of amniocentesis would have to decrease by more than 50% for the quadruple test to be CE (104).

Similarly, the cost of all the other tests had to decreases by 50% for them to be more CE than the integrated test. They conclude by saying that integrated test, quadruple test, first-trimester combined test, and NT would present the best results in terms of CE, effectiveness, and safety, and the decision to choose one of them would depend on: the total budget available, the willingness to pay per case detected and the value placed on the safety of the tests. This finding in line with many other CE studies which supports integrated test as the most CE one. NIPT's cost would have to decrease by around three times for the cost per case detected to be level with first-trimester or second-trimester screening test to implement it as the first-tier screening strategy as suggested by one study (104).

NIPT as a screening test performs way better than other tests with a sensitivity of more than 99% (very close to 100%) and a false-positive rate of less than 0.1% (very close to 0%)(105-107). This means fewer women have to go through invasive procedures, resulting in very few miscarriages due to invasive diagnostic tests. Health professionals and pregnant women have greater interest due to the high accuracy and safety. This has led to many countries introducing

NIPT as a part of their screening setup. However, NIPT has been criticized due to high cost, the possible impact on the acceptance of disability, etc. In terms of cost-effectiveness, various strategies have concluded that NIPT is mostly found to be not cost-effective as a primary screening strategy.

Yan Xu et al. conducted a CE analysis of NIPT as the universal primary screening tool compared to the conventional maternal serum screening (CMSS) in a hypothetical cohort of 10,000 women in China. They found out that universal NIPT could prevent more DS cases (9.97) than CMSS (3.02). The incremental cost per case detected was $352,388 for universal NIPT compared to CMSS, and the strategy would be CE if the price of NIPT decreased to $76.92. However, the introduction of NIPT as a secondary test for women considered to at high risk from CMSS would result in 7.52 DS cases detected and cost only $20,160 compared to CMSS (31).

Similarly, N. Okun et al. studied the cost and performance of NIPT and the implementation of NIPT in the prenatal screening of DS in Canada. Compared to the integrated tests, which is the dominant screening strategy in Canada which would cost $112,919 per case diagnosed, NIPT as the primary screening test would cost $286,428, which is around 2.5 times higher. However, contingent implementation of NIPT in the screening setup produces different results. NIPT as a second-tier test to integrated test was associated with a significant reduction in the number of amniocenteses performed and decreased resulting miscarriages. Incorporating NIPT with the first-trimester combined test as the second test would reduce the cost per case detected to $69,583. Moreover, the cost would decrease to $63,383 when the detection rate and the uptake rate increased to 95% and 80%, respectively (108).

Various other studies confirm the use of NIPT as a contingent strategy like using NIPT as a second-tier test along with tests like CUB test, integrated tests as the primary test, using NIPT as tests for women over 35 years of age only, etc. result in a significant reduction in the cost per case detected and total cost as well from a health payer perspective (109-112).

However, all the studies that have concluded NIPT not being cost-effective have been analyzed from the health care or payer's perspective. If analyzed from a societal perspective as done in the study by Brandon S. Walker et al., NIPT was found to be less costly and more effective than the integrated test, with NIPT costing around $122 million less than integrated tests and $866 million less than no screening when societal costs like lifetime medical costs, education costs, and indirect costs were included in the study. While it would cost more than maternal

serum screening, it would be justified with fewer false positives and subsequent diagnostic tests. Adoption of NIPT would save approx. $277 955 for every additional case detected compared to integrated tests (110).

3.5 Critical analyses and summary

Findings from all the studies suggest that there is no single strategy that is always cost-effective than others. A screening test's cost-effectiveness depends on the detection rate, false-positive rate, uptake rate, the availability of the test, and the unit cost of the test. A test with a high detection rate and low false-positive rate like NIPT does not necessarily mean that it is cost-effective as it might be significantly more costly than others. A test with low DR and high false-positive rate despite costing less would mean more invasive procedures and subsequent fetal loss and considered not safe even if the tests with the right combination of DR and FPR values and costs might not be appreciated and accepted by the women going through the screening which would then affect the CE.

Focusing solely on the cost and effects might also be misleading when deciding on adopting certain screening tests in the health policies. An important aspect that cannot be missed is the perspective of analysis. A policy might be CE from a health care or payer's perspective but not from a societal perspective, and vice-versa. Keeping this in mind is important when making recommendations from CE studies. Similarly, it is very important to consider the perception of the government and people on disability like DS. People might put more value on life and consider termination of pregnancy as an unethical thing to do. There are countries where the termination of pregnancy is banned. So, it is critical to put all these factors into perspective before coming to any conclusions.

To conclude, a multitude of factors affect the CE of certain screening tests and might vary from one setting to another. Most of the studies analyze the CE of different prenatal screening strategies for DS based on developed countries. There are very few done in developing countries where the healthcare system, characteristics of the people, costs of tests, knowledge, and awareness on disabilities like DS, etc., are completely different from the developed countries. Moreover, it is clear from the literature review that the same strategy can be either CE or not, depending upon the setting.

To summarize, there is limited (almost no) literature on the cost-effectiveness of prenatal screening of DS in developing countries like Nepal. Most of the studies compare two or three specific DS screening methods and cater to a specific setting. This study includes all the

possible interventions present and caters specifically to Nepal. With such high uncertainty in the making the decision, there are no studies which have looked the value of collection of additional information of input parameters through VOI analysis. Thus, this study will fill this literature gap.

Chapter IV: Methodology

This chapter contains the methodology used in the study. It includes a detailed explanation of the model structure, assumptions, information about the input parameters used like cost, effect, etc., and validity and limitation of the model.

4. 1 Model

4.1.1 Model overview

In the study, decision tree modeling was built to calculate the expected costs and outcomes of all the interventions to evaluate the cost-effectiveness. Decision tree is preferred economic model for screening of conditions like DS where there is no chance of a woman to have multiple states, and go back and forth between the states which is the case in Markov model. A women either tests positive or negative in the screening and diagnostic tests and therefore best represented by a decision tree as done in most of the studies that have assessed the cost-effectiveness of screening interventions of DS worldwide and thus chosen for the study (29, 31, 95, 99).

The model begins when a woman is screened in her gestational period (everyone assumed to be in 10^{th} weeks of pregnancy) and spans till the end of pregnancy irrespective of the outcomes, i.e., birth (with or without defects), abortion, or miscarriage. All the cost and outcome were calculated only for the mentioned time period, which means that the cost and quality of life associated with live birth with the defect and miscarriage will not be included in the study for feasibility reasons. MS Excel was used for the calculation of outcomes in a decision tree.

Hypothetical cohorts of 753,506 (reproductive age group mentioned as general population), 231,308 (pregnant women 35 years or above) and 139,948 (pregnant women 40 years and above) pregnant Nepalese women were modeled separately. Only singleton pregnancies are considered in the model. A test-based approach with a sequential diagnostic method was used in the model (113). The evaluation is from the healthcare provider's perspective, and thus only the direct medical cost, i.e., cost of screening, diagnosis, genetic consultation, etc., were taken into account. Similarly, the outcome of interest in the model was the number of DS cases detected. The cohort of women have a 0.092% (general population), 0.3289% (women 35 years or above) and 1.5625% (women 40 years or above) probability of developing pregnancy with DS based on the age group and thus enter the model with the same probability. The corresponding birth prevalence of particular age group was used to calculate test characteristics like Positive Predictive Value (PPV), the proportion of total positive cases from a particular

test (T+), etc. Since the model's time frame was one year, no discounting of costs or outcome was needed.

4.1.2 Model structure

A simplified version of the decision tree is shown in figure 2.

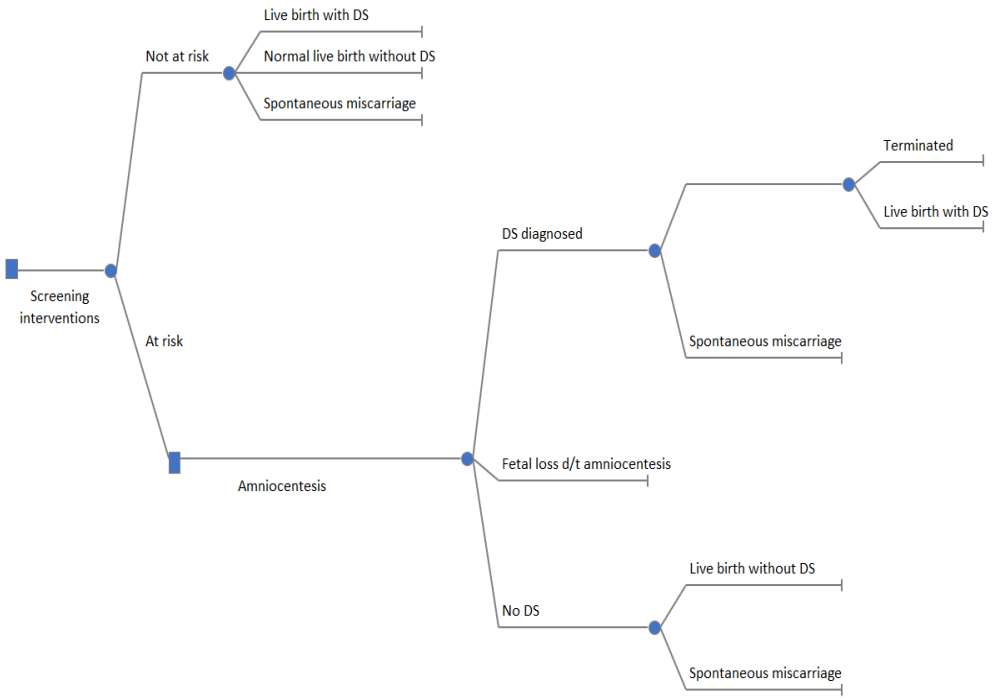


Figure 3 Simplified version of decision tree

At the start of the model, the cohort of women go through the intervention (one intervention at a time for all the listed interventions). All the women are assumed to undertake the intervention. Depending upon the sensitivity and false-positive rate of the particular screening test and provided that the women agree to take on the screening test, women will be divided into test positives and test negatives. A risk cut-off value of 1:250 was taken for categorizing women at risk. The women who test positive after accepting screening interventions go through amniocentesis, chosen as the diagnostic procedure. There are three possible outcomes for those women who test positive in amniocentesis i.e. fetus with DS, No DS, and fetal loss due to amniocentesis as shown in the figure above. Amniocentesis confirms the presence or absence of DS. Women have the chance of terminating the pregnancy or keeping the baby if they have a positive fetus. Women with negative amniocentesis go through normal live birth without DS. There is also a possibility of spontaneous miscarriage for women with or without positive fetus. The women with negative test results do not undergo any further process and end with any of the three possible outcomes i.e. live birth with DS, normal live birth without DS and

spontaneous miscarriage as shown in the figure above. The outcome from these two pathways only was considered as having utility weights of one. Utility weights for the rest of the pathways were considered zero. The same cohort of women are considered for all the interventions for analysis if one age group and go through the model one at a time for all the interventions separately. The results obtained from the model are then compared.

All the possible list of outcomes in the model are listed below (13):

1. Down's syndrome detected and terminated
2. Live birth with Down's syndrome (after positive amniocentesis)
3. Live birth with Down's syndrome (after the negative result from screening test)
4. Fetal loss due to amniocentesis
5. Spontaneous miscarriage (1^{st} and 2^{nd} trimester)
6. Live birth without Down's syndrome

4.1.3 Model assumptions

Several assumptions were made during the process of building the model. All pregnant women from all age groups are assumed to be willing to go through the first layer of the screening process due to the lack of data regarding the acceptance of these tests, which would require a survey (*not possible right now due to COVID*). It is also assumed that all women will consent for amniocentesis if they are deemed at risk from primary screening. The cohort is representative of pregnant women in Nepal. Since specific values of the test properties like sensitivities and false-positive rates are not available for Nepal, the most reported values in the scientific literature have been used. According to The Right to Safe Motherhood and Reproductive Health Act, 2075 B.S (2018 A.D) of Nepal, women can opt for termination of the fetus up to 28 weeks if a licensed doctor suggests that the health of the pregnant woman is in danger or an infant with a disability might be born if the abortion is not performed (114). Thus, all the DS cases will be eligible for termination since all the tests involved will be done within 24 weeks of pregnancy. Abortion service is free of cost in Nepal. Thus, it is assumed that all the women will be advised to choose public hospitals for termination and opt for it.

4.2 Interventions evaluated

Seven risk screening (prenatal) methods were evaluated one at a time in the study. The base case of benchmark for the study is "do nothing" or "no screening." Although various private facilities provide the tests that can pay for it, there is no official policy for screening DS, and the government does not provide any sort of re-imbursement. The tests were:

1. Ultrasonography (Nuchal translucency)
2. Double test (DT)
3. Combined Ultrasound and Blood test (CUB) test
4. Triple test (TT)
5. Quadruple test (QT)
6. Non-invasive prenatal test (NIPT)
7. Contingent test (CUB test as first-tier and NIPT as second-tier tests)

All the positive cases from the risk screening tests were provided with Amniocentesis as a diagnostic procedure to confirm the presence or absence of DS.

8. Amniocentesis

There are few other screening methods like the Integrated test, but it was not included in the study as the integrated test is not practiced in Nepal. Only those screening tests which are available in Nepal were included. Similarly, amniocentesis was the sole diagnostic test selected as CVS is not available in Nepal to date.

4.3 Perspective, WTP threshold time horizon, and discounting

The CEA was done from a health care perspective for Nepal, which is the Ministry of Health of Nepal. The WTP threshold per additional case detected was assumed to be $3,000. WHO recommends the use of threshold (per QALY gained or DALY averted) equal to 1-3 times the GDP per-capita of the nation (115, 116). However, no recommendation was found on the threshold per additional case detected. Thus, the method proposed by WHO was used for the calculation of WTP as $3,000 per additional case detected. Similar approximation and figures have been used by other studies (104, 117). A time frame of one year was used for the model. Hypothetical cohorts of women according to the age group is the expected number of pregnant women in Nepal for 2020/21 A.D. All the screening and diagnostics provided will be within one year. Similarly, BIA was calculated from the provider's perspective. A time horizon of 10 years was selected for BIA, and an annual budget for individual strategy was calculated. Costs in BIA were not discounted. A five-year time horizon was adopted for VOI analysis and a discount rate of 3.5% was used for discounting the population affected over five years.

4.4 Data

4.4.1 Epidemiological data

There was no data related to birth prevalence or incidence of DS in the Nepalese population through a literature search. Therefore, studies conducted in countries like India and Bangladesh

Table 2 Input parameters

Variable		Baseline value	Range	Distribution
Incidence	For women of all age group (%)	0.0092[22]	0.076-0.114[b]	Beta
	For women ≥35 years (%)	0.3289[22]	0.256-0.38[b]	Beta
	For women ≥40 years (%)	1.5625[22]	1.2-1.8[b]	Beta
NT	Sensitivity	0.7[95]	0.66-0.76[95] [95]	Beta
	False-positive rate	0.05[95]	0.0192-0.0288[95]	Beta
	Positive predictive value	0.051[b]	0.0216-0.0324[b]	Beta
	Proportion of Test positives	0.0127[b]	0.0197-0.0296[b]	Beta
Double test	Sensitivity	0.62[95]	0.55-0.74[95]	Beta
	False-positive rate	0.05[95]	0.04-0.06[95]	Beta
	Positive predictive value	0.0113[b]	0.0093-0.014[b]	Beta
	Proportion of Test positives	0.051[b]	0.04-0.06[b]	Beta
CUB test	Sensitivity	0.818[95]	0.77-0.87[95]	Beta
	False-positive rate	0.021[95]	0.0168-0.0252[95]	Beta
	Positive predictive value	0.034627143[b]	0.0286-0.0429[b]	Beta
	Proportion of Test positives	0.0218[b]	0.0174-0.0261[b]	Beta
Triple test	Sensitivity	0.651[95]	0.52-0.78[95]	Beta
	False-positive rate	0.047[95]	0.037-0.056[95]	Beta
	Positive predictive value	0.012594079[b]	0.0104-0.0156[b]	Beta
	Proportion of Test positives	0.0476[b]	0.038-0.057[b]	Beta
Quadruple test	Sensitivity	0.76[82]	0.608-0.912[b]	Beta
	False-positive rate	0.09[82]	0.072-0.108[b]	Beta
	Positive predictive value	0.0077[b]	0.0064-0.0096[b]	Beta
	Proportion of Test positives	0.091[b]	0.072-0.108[b]	Beta
NIPT	Sensitivity	0.99[31]	0.947-1[31]	Beta
	False-positive rate	0.001[31]	0-0.002[31]	Beta
	Positive predictive value	0.4768886[b]	0.3879-0.5819[b]	Beta
	Proportion of Test positives	0.002[b]	0.0016-0.0023[b]	Beta
Amniocentesis	Sensitivity	0.993[31]	0.947-1[31]	Beta
	False-positive rate	0.0014[31]	0-0.0017[31]	Beta

b- From calculation

*positive predictive values for 0.0092% incidence of DS

were searched. India has a similar socio-economic, demographic and cultural situation to Nepal. Moreover, India too does not have a screening policy for DS. Thus, birth prevalence of

based on the age group was taken as the estimate for the study conducted in India mentioned earlier (22).

The sensitivities and false-positive rates of individual tests were reported from various studies. However, the values of PPV and T+ are for the general population as their values depend on the incidence. The values of PPV and T+ of individual tests for the other two groups in the analysis are reported in the appendix. The input values of the tests are listed in the table below:

4.4.2 Cost

Since the benchmark or base case in this analysis was no screening or no screening and the analysis was done from a health care perspective, there was no direct cost associated with no screening. There is no guideline on the cost or re-imbursement of the listed screening test from GoN or MoH, Nepal. Thus, expert advice on the cost of all the tests was sought.

Table 3 Cost values

Intervention	Parts of intervention	Cost ($)	Total cost ($)	Range	Distribution
Nuchal translucency	Ultrasound	8[a]	8	9.6-6.4[c]	Gamma
Double test	Serum test	25.44[a]	25.44	30.528-20.352 [c]	Gamma
Combined Ultrasound and Biochemical test	Ultrasound	8[a]	33.44	40.8-27.2 [c]	Gamma
	Serum test	25.44[a]			Gamma
Triple test	Serum test	39[a]	39	46.8-31.2 [c]	Gamma
Quadruple test	Serum test	51[a]	51	61.2-40.8 [c]	Gamma
Non-invasive prenatal test (NIPT)	cf-DNA test	250[a]	250	300-200 [c]	Gamma
Amniocentesis		260[a]	260	312-208 [c]	Gamma
Termination	Termination of affected pregnancy	0[b]	0	0	Gamma
Delivery	At public hospitals	0[b]	0	0	Gamma
Consultation with physician[b]		6.5[b]	6.5	7.8-5.2 [c]	Gamma

a- Expert opinion

b- Ministry of Health, Government of Nepal

c- From self-calculation

The cost of the tests represents the amount charged on average by the laboratories that have these services available. All the cost has been converted into US dollars for uniformity and comparability. The cost of medical consultation is based on the guideline provided by the MoH. There is no cost related to termination as abortion services are free of cost in public facilities in Nepal. Lastly, women who opt for keeping the babies with DS were assumed to choose to deliver their babies in a public hospital, and thus there is no charge for delivery as it is almost free of cost in public facilities. All the costs associated with the screening and diagnostic services have been listed in the table below:

No discounting is needed as the analysis has a time frame of one year.

4.4.3 Effect (Outcome)

The effectiveness of the interventions was measured by the number of Downs Syndrome detected. This includes both the groups: those who terminate the fetus and those who choose to keep the baby with DS. Both the outcomes were assigned with a utility of one. All the other possible outcomes were assigned with the utility of zero.

Table 4 Outcome values

Pathways	Utility weight
DS cases detected (terminated or not terminated)	1
All other pathways	0

There are other outcomes possible in the model. The list of possible outcomes and their values are provided in the table below:

Table 5 Input parameters

Variables	Baseline value	Range	Distribution
Fetal loss due to amniocentesis	0.0081[(95)]	0.005-0.01[(95)]	Beta
Spontaneous miscarriage for DS pregnancy			Beta
From first-trimester screening onwards	0.35[(95)]	0.31-0.75[(95)]	Beta
From second-trimester screening onwards	0.20[(95)]	0.12-0.50[(95)]	Beta
Overall spontaneous miscarriage risk from the first-trimester	0.027[(95)]	0.02-0.04[(95)]	Beta
Overall spontaneous miscarriage risk from the second-trimester	0.012[(95)]	0.01-0.02[(95)]	Beta

As for the costs, discounting the number of cases was unnecessary as the time frame is one year.

4.5 Uncertainty analysis

Uncertainty analysis was done using two methods. One-way and two-way sensitivity analysis was done as a part of deterministic uncertainty analysis, and 1000 Monte-Carlo simulations were done for probabilistic uncertainty analysis. Both uncertainty analyses were carried out for the cost-effectiveness ratio of all the interventions and ICER of the most cost-effective interventions within the willingness to pay threshold.

4.5.1 Deterministic sensitivity analysis

One-way and two-way sensitivity analyses were performed for both CERs and ICERs.

Birth prevalence/Incidence of DS, maternal age, and cost of amniocentesis were used in one-way sensitivity analysis for CER. A two-way sensitivity analysis was carried for CER using scenario analysis where CER was calculated for the most optimistic and pessimistic input parameters one at a time and reported. The result of scenario analysis is presented in the form of a tornado diagram.

One-way sensitivity analysis of ICERs of all interventions was carried out by varying the input parameters' values from maximum to minimum. The ICER thus calculated was reported in the form of a table.

4.5.2 Probabilistic analysis

Probabilistic analysis was carried out using Monte-Carlo simulation for 1000 iterations. PSA was carried out for CER for all the interventions individually and also for ICER for the intervention with the lowest ICER compared to the next less costly intervention. The results from both PSA were presented in a CEP. Similarly, the results from ICER were also used to calculate CEAC and CEAF using a willingness to pay threshold of 3,000 USD and presented in figures. Gamma distribution was used for the probabilistic analysis of cost values, and beta distribution was used for all other input variables.

4.6 Value of Information analysis

The expected value of perfect information was calculated using the results from PSA and validated using SAVI. EVPI was calculated using the results from PSA subtracting the maximum expected net monetary benefit (the maximum NMB that can be achieved with

current information) from the expected maximum NMB (the maximum NMB that can be gained with the collection of perfect information on all the parameters.)

Individual and population EVPI at a threshold of $3,000 per incremental case detected were calculated and reported. Individual EVPI was also presented in the form of an EVPI graph along with CEAC and CEAF. There was no need to calculate EVPPI as the value of EVPI was zero at the current willingness to pay threshold for every age group.

A special scenario was developed for women 40 years older with the cost of tests shared between the health care payer and the individual (30% individual and 70% health payer). EVPI was calculated at the same threshold and reported in the forms of table and CEAC+EVPI graph. Since there was a positive value of EVPI, partial EVPI or EVPPI was calculated using the SAVI framework, which uses regression-based methods for calculation instead of traditional Monte-Carlo simulations. Finally, individual EVPPI and group EVPPI were calculated.

4.7 Budget impact analysis

A budget impact analysis was done using the cost obtained from the analysis and projecting the annual cost to five years to evaluate the total budget needed to realize every strategy and assess whether it is viable from the long-term perspective. No discounting of cost has been done in BIA in this study.

4.8 Safety index

Safety index was calculated after obtaining the data related to fetal loss due to amniocentesis of every strategy and the total number of cases detected from the analysis as a ratio of former over the latter. This was used as a measure of the safety of each strategy and as a supplement for the results obtained from CEA. The result of the safety index has been presented in the form of a table.

4.9 Harm to benefit analysis

Harm to benefit analysis was also calculated to support the CEA. As discussed in the theoretical section of the report, harm was measured in terms of the number of amniocenteses performed for every test, and the benefit was measured in terms of the total number of cases of DS detected. This is another measure of the safety of the interventions. The result is presented in a table and incremental harm-benefit plane. All the strategies were arranged according to their harms in ascending order, and incremental harm and benefits were calculated compared to the strategy with the lowest harm. If a strategy had higher harm and lower benefits, then the

strategy was assumed to be dominated by the one with lower harm and/or higher benefit and was excluded from the calculation.

Chapter V: Results

This chapter includes the results of the decision-analytic model. All the results from the deterministic analysis, sensitivity analysis: deterministic and probabilistic analysis, VOI analysis, harm to benefit analysis, scenario analysis, etc., are described in detail in this section using appropriate tables and figures.

5.1 Cost-effectiveness analysis

Table 6 shows the total cost of introduction of different screening tests for all the pregnant women (752,506), the total number of DS cases detected and the cost per case detected in dollars. Similarly, the ICER of each strategy compared to no screening and the next costly strategy is also presented in the table.

Table 6 Deterministic results

Strategy	Total Cost ($)	No. of cases detected	Fetal loss d/t amniocentesis	CER	ICER	ICER (Com. to DN)	Remarks
No screening	0	0	0	#NA			
Nuchal Translucency	15732622	418	150	37661	37661	37661	
Double test	33923565	394	308	86057	-772769	86057	SD
Contingent screening	34724215	460	5	75501	450,289	75501	ED
CUB test	34733313	482	133	72027	294647	72027	
Triple test	43546911	409	290	106399	-120821	106399	SD
Quadruple test	61002268	507	552	120229	1044070	120229	
NIPT	193397265	563	12	343473	2377837	343473	

SD- Strongly dominated

ED-Extendedly dominated

All the strategies have been arranged in ascending order according to the total cost. The current strategy, which is no screening for DS, does not have any cost from a health care perspective, and there are no DS cases detected. The next costly strategy in terms of the total cost is NT which costs approx. $ 15.73 million. The total cost for DT and the contingent test is almost the same, i.e., $33.92 million and $34.72 million. There is a gradual increase in the total cost from the CUB test to TT and to QT. The cost is around $34.73, $43.54, and $61 million, respectively. However, the total cost of a screening strategy for NIPT is more than three times the next lowest intervention which is QT. This is due to the very high cost associated with NIPT, which is more than 4-8 times higher than other interventions.

In terms of the number of DS cases detected, NIPT is the most effective as it detects 563 cases of DS. The next best strategy is QT, followed by the CUB test with 507 and 482 cases detected, respectively. The TT detects 409 DS cases compared to 460 by the contingent test. NT detects about 418 cases of DS, whereas DT which is the least effective of all with 394 cases.

Of all the strategies, NT has the lowest cost per case detected, i.e., $ 37,661, which is approx. half less than the next best strategy, the CUB test. The CUB test has a CER of $72,027 per case detected. The contingent test and DT screening have a similar cost per case detected, with costs of $75,501 and $86,057 per case detected, respectively. Lastly, the three most costly strategies: TT, QT, and NIPT, have an estimated cost of $106,399, $120,229, and $343,473 per case detected in respective order. Although the DT and contingent test cost roughly the same as CUB, they are far less effective in terms of case detected and thus have higher CER. TT costs more than CUB but detects fewer cases. QT results in the detection of more cases than CUB but has almost double the cost. Finally, the high effect of NIPT does not justify the huge cost and has by far the highest CER of all the interventions.

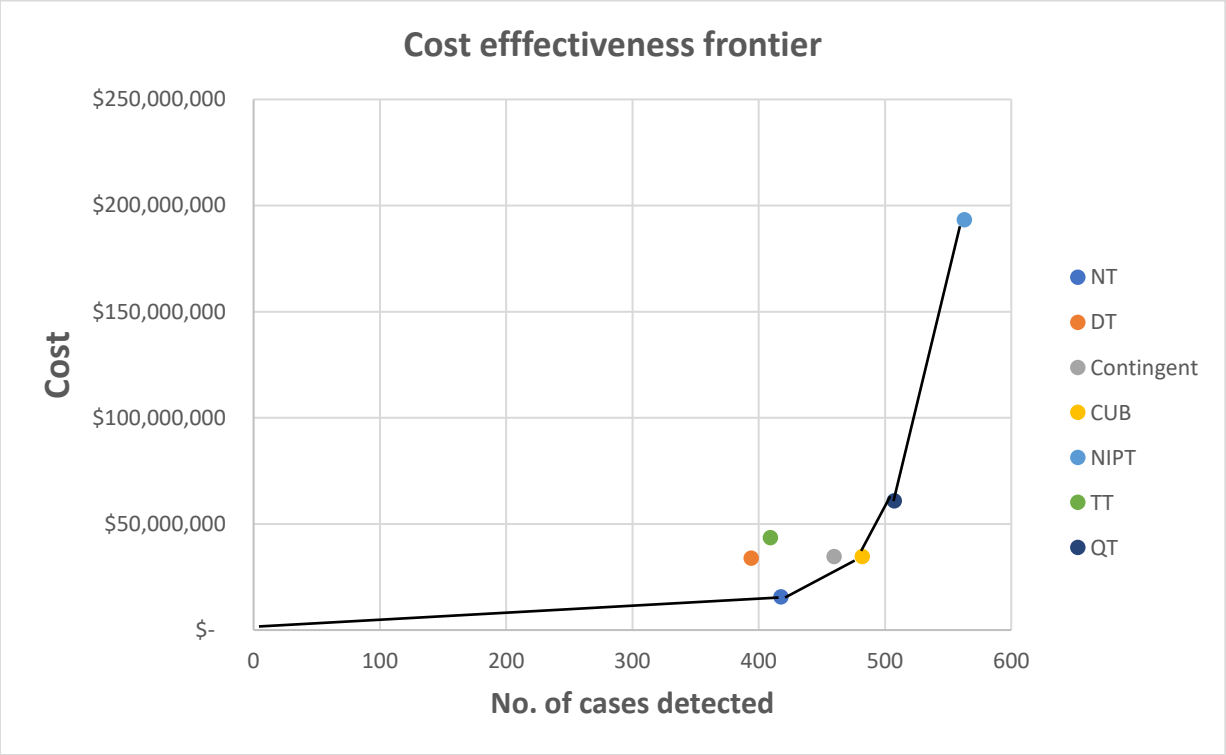


Figure 4 Cost-effectiveness frontier (general population)

The cost-effectiveness plane of all the strategies evaluated is presented in the figure. The x-axis represents the number of DS cases detected, while the y-axis represents the total cost of

the strategies. The line joining the points is the cost-effectiveness frontier. As seen in the figure, the most effective strategy compared to no screening is NT, followed by CUB. The contingent test is very close to CUB in terms of cost-effectiveness. The steep rise in the slope of the line from CUB to QT and NIPT suggests that there is a considerably high increase in the ICER of those two tests. DT, TT, and contingent test all lie inside on the left side of the frontier, which means they have been dominated (either costs more or/and have a lower effect).

ICER, in this case, is the cost associated with the detection of extra cases of DS. Two ICERs have been reported in the table. 1)ICER when compared to the strategy with the next lowest cost and 2) ICER when compared to no screening.

ICER of all the strategies when compared to no screening is equal to the cost per case detected or CER as both the costs and effects of no screening are zero.

Thus, the most cost-effective strategy when compared to no screening is NT with ICER of $37,661, followed by CUB with ICER of $ $72,027. Similarly, the cost of additional cases detected when compared to no screening for contingent screening is $ 61,135, which is close to CUB's. The next two cost-effective strategies are DT and TT with ICER of $86,057 and $106399. QT and NIPT, with ICER of $120,229 and $343473, are the last two strategies in terms of ICER compared to no screening.

However, when multiple interventions are evaluated at a time, interventions are compared with another intervention having the next lowest cost instead of no screening. The results are different when ICER is calculated using a strategy with the next lowest cost as a comparator. NT when compared to no screening, has an ICER of $ 37,661. DT is strongly dominated as it costs more and detects fewer DS cases than NT. Similarly, the contingent test is extendedly dominated by the CUB test as it costs only fractionally more and results in a large incremental effect. Thus, the ICER of CUB compared to NT is $294,647. TT costs more with lower effect than CUB and therefore is strongly dominated. Thus, the ICER of QT compared to CUB and NIPT compared to QT are $1,044,070 and $2,377,837 per additional case detected.

The results of the analysis where the tests were provided to pregnant women who were 35 and over and 40 or over separately at the time of conception are presented in the table below. CER and ICERs of all the strategies in both the scenarios have been reported.

Table 7 Deterministic results (for two age groups)

Strategy	CER	ICER	ICER (Com. to DN)	CER	ICER	ICER (Com. to DN)
	≥35 years			≥40 years		
No screening						
Nuchal Translucency	11794	11794	11794	8940	8940	8940
Double test	27829		27829	20277		20277
CUB test	22171	85365	22171	16071	58608	16071
Contingent screening	22863		22863	17211		17211
Triple test	34101		34101	24617		24617
Quadruple test	40078		40078	29189		29189
NIPT	102306	508448	102306	70688	333889	70688

The overall results for both the scenarios are similar to the result for the general population. NT is the most cost-effective strategy, followed by the CUB test. However, the CER and ICER of all the strategies in both the scenarios change drastically. The CER (ICER is the same as CER for NT) of NT falls from $37,661 to $11,794 and $8940 per case detected for women 35 years and above and 40 years and above. Similarly, the CER of the next cost-effective strategy, i.e., CUB, is $22,171 and $16,071 per case detected for the two age groups compared to $72,027 for the general population. ICER of CUB compared to NT for both the age groups also decreases from $294,647 for the general population to $85,365 and $58,608 per additional case detected. However, QT no longer remains the next cost-effective strategy as it was for the general population because it costs more and results in fewer cases detected than the CUB test. NIPT is the most effective and most costly strategy for both the age groups with CER (≥35 years -$102,306 and ≥40 years-$70,688 per case detected) and ICER (≥35 years- $508,448 and ≥40 years-$333,889), which is significantly lower than the values for the general population.

Apart from the costs and the number of cases detected by all the strategies, the total fetal loss as a result of the invasive procedure in all the screening strategies is also important from an analytical point of view. The results for all the age groups are presented in the table below:

Table 8 Total fetal loss d/t amniocentesis

Interventions	Total fetal loss d/t amniocentesis		
	General population	≥35 years	≥40 years
NT	150	49	39
DT	308	97	66
Contingent	5	5	14
CUB	133	44	37
TT	290	92	64
QT	552	173	113
NIPT	12	8	18

Across all the age groups, NIPT and contingent tests are the two most effective and safe tests in terms of the number of fetal loss due to invasive procedures. The contingent test, which includes two tiers of screening test before amniocentesis, results in the lowest number of fetal loss, i.e., 5, 5, and 14 (general population, ≥35 years, and ≥40 years). NIPT results in a slightly higher number with 12,8 and 18. The next best strategy is CUB tests with 133, 44, and 37, followed by NT with 150,49 and 39 fetal loss, respectively. DT, TT, and QT result in a very high number of fetal loss compared to other strategies, as can be seen from the table above.

The reason for such results is that NIPT has very high sensitivity and specificity and can detect the true positive and negative better than any other tests. Similarly, contingent screening includes NIPT as the second-tier screening combined with CUB, which already has relatively higher sensitivity and low FPR. CUB test, although it has lower sensitivity and higher FPR than NIPT, has higher sensitivity and lower FPR than DT, TT, and QT. Thus, the number of false positives or women being falsely categorized as at-risk is lower for these tests. Lower false positives mean lower amniocentesis and a low number of resulting fetal loss.

5.2 Net Monetary Benefits and Net Health Benefits

The results of economic analysis can also be expressed in terms of net monetary benefits or net health benefits. NMB refers to the monetary gain attached with the intervention for the population, and NHB refers to the health gains for the population. The NMB and NHB of all the interventions for different scenarios are presented in table 9.

Table 9 NMBs and NHBs

Interventions	Net Monetary Benefit			Net Health Benefit		
	General population	≥35 years	≥40 years	General population	≥35 years	≥40 years
No screening	0	0.0	0.00	0	0	0
NT	-14479405	-3674023	-15.53	-4826	-1225	-0.005177234
DT	-32740968	-9372255	-40.19	-10914	-3124	-0.013395038
Contingent	-33344468	-9325068	-42.86	-11115	-3153	-0.014285234
CUB	-33286637	-9458383	-39.91	-11096	-3108	-0.013302833
TT	-42319077	-12282555	-52.75	-14106	-4094	-0.017584836
QT	-59480111	-17431386	-74.93	-19827	-5810	-0.024976589
NIPT	-191708072	-57833828	-249.56	-63903	-19278	-0.083185509

As seen in the table, NMBs of all the interventions for the general population, ≥35 years and ≥40 years are negative. This means that there is no monetary gain attached to these interventions at the WTP threshold of $3,000. Moreover, they result in loss, which means that they cost more than decision-makers are willing to pay. Alternatively, in other words, the costs outweigh the costs (expressed in dollars).

Similarly, NHBs of all the interventions at the current WTP threshold for all the age groups are negative except NT in the special scenario. This basically means adopting all the interventions would bring negative health effects in the population, or the health benefits do not outweigh the health losses to the population and thus would not be the correct choices.

These negative values of NHBs and NMBs in all the cases can be attributed to the very low WTP threshold of the setting. Thus, this suggests decision-makers decide not to adopt the interventions in the current scenario.

5.3 Uncertainty analysis-

Uncertainty analysis of the results includes two sections. The first section includes the deterministic sensitivity analysis and the second section includes the probabilistic analysis of the results.

5.3.1 Deterministic sensitivity analysis

The deterministic sensitivity analysis has been further divided into two parts: sensitivity analysis of 1. CER of all the strategies and 2. ICER of the two most cost-effective strategies. A number of one-way and two-way sensitivity analysis was carried out to assess the effect of different input variables individually and combination on the results.

5.3.1.1 Cost-effectiveness ratio

One-way sensitivity analysis of the CER by varying birth prevalence of DS, maternal age, and cost of amniocentesis, which are the common variables for all the strategies, are presented in figures.

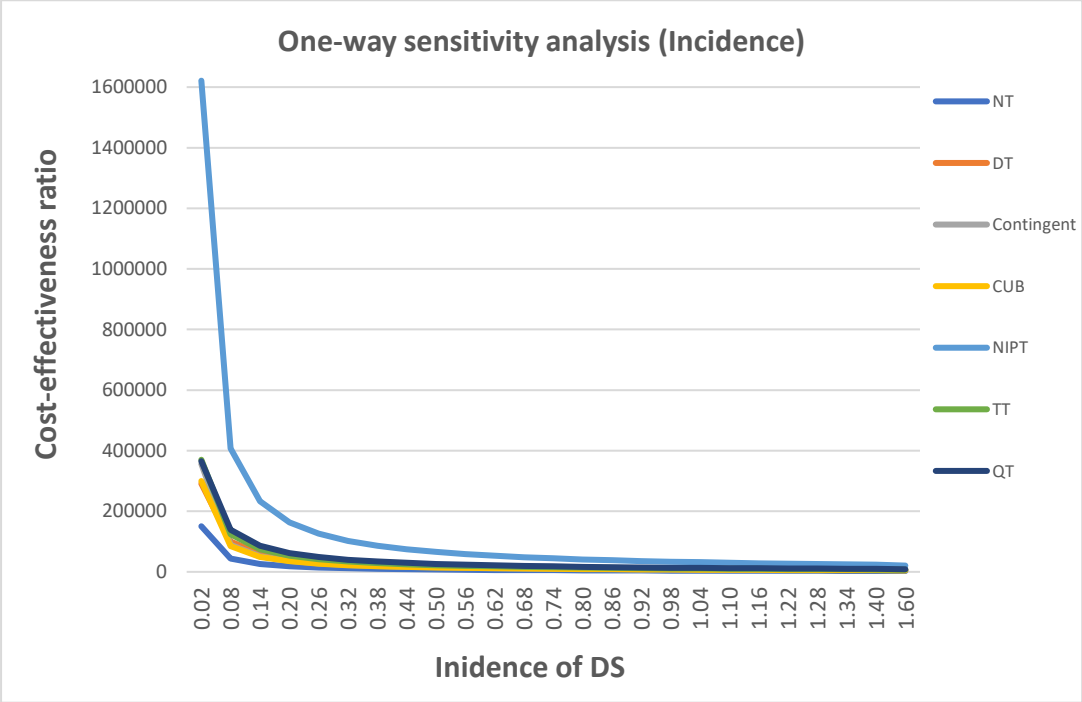


Figure 5 One-way sensitivity analysis (Incidence)

The x-axis in the figure ranges from 0.02% to 1.6% to depict the reality where the incidence can range from such low value to as high has 1.6% in the general population. As seen in the figure, the CER of all the strategies decreases sharply at first when the incidence/birth prevalence of DS increases. After a certain point, the decrease is gradual and constant at higher values of incidence. The CER of all the strategies is very high compared to the base case. Even NT has a CER of $150,446 per case detected at 0.02% incidence. The CER of all the strategies reduces to a lower level at around 0.5% and stays more or less constant or reduces very minimally. The CER of NIPT shows the most drop from the lowest value of incidence to the

highest values. It starts with the highest similar CER to others when the incidence is around 1.6%.

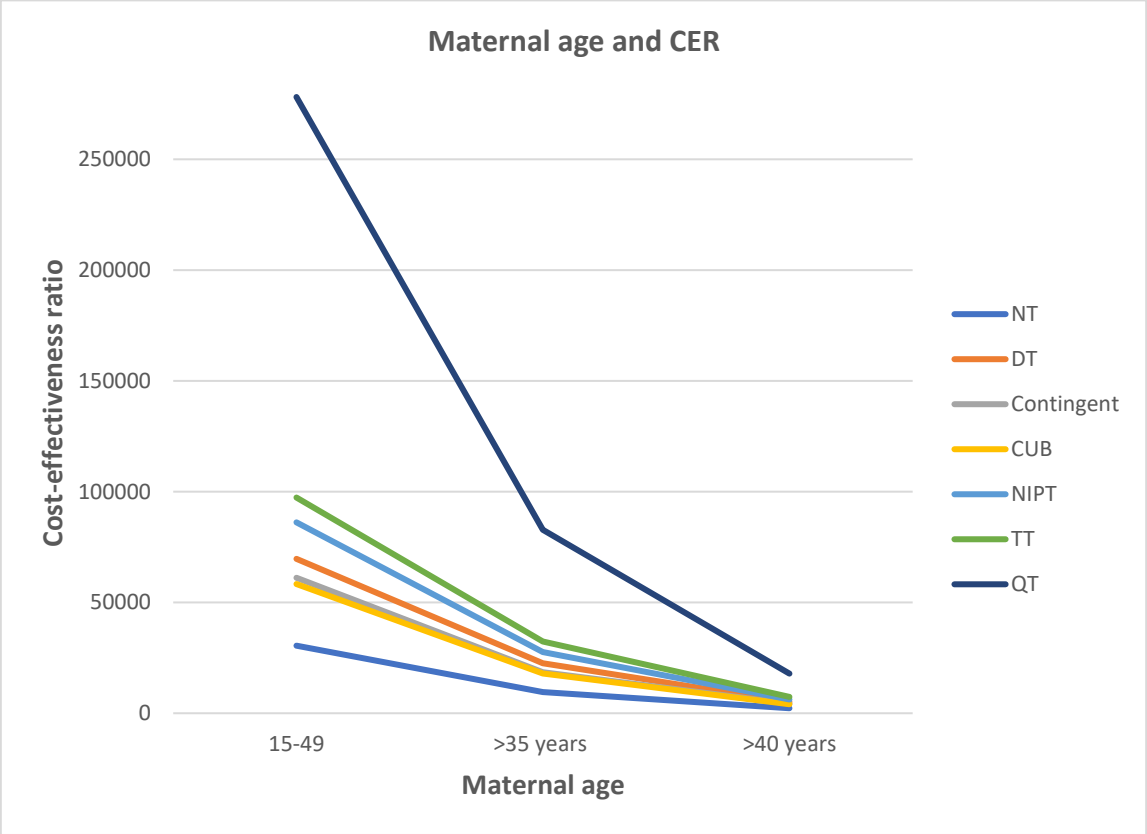


Figure 6 One-way sensitivity analysis (maternal age)

Figure 5 shows the change in CER with the change in age group selected for analysis. As seen in the figure above, CER for all the strategies is higher for the general population compared to other age groups. The slope of the decrease is high from general to ≥35 years which means a significant decrease in the CER. Similarly, there is a significant decrease for the age group ≥40 years compared to ≥35 years but not as high as the drop from the general population to ≥35 years. Thus, all the strategies are more cost-effective if women of higher maternal ages are chosen for screening.

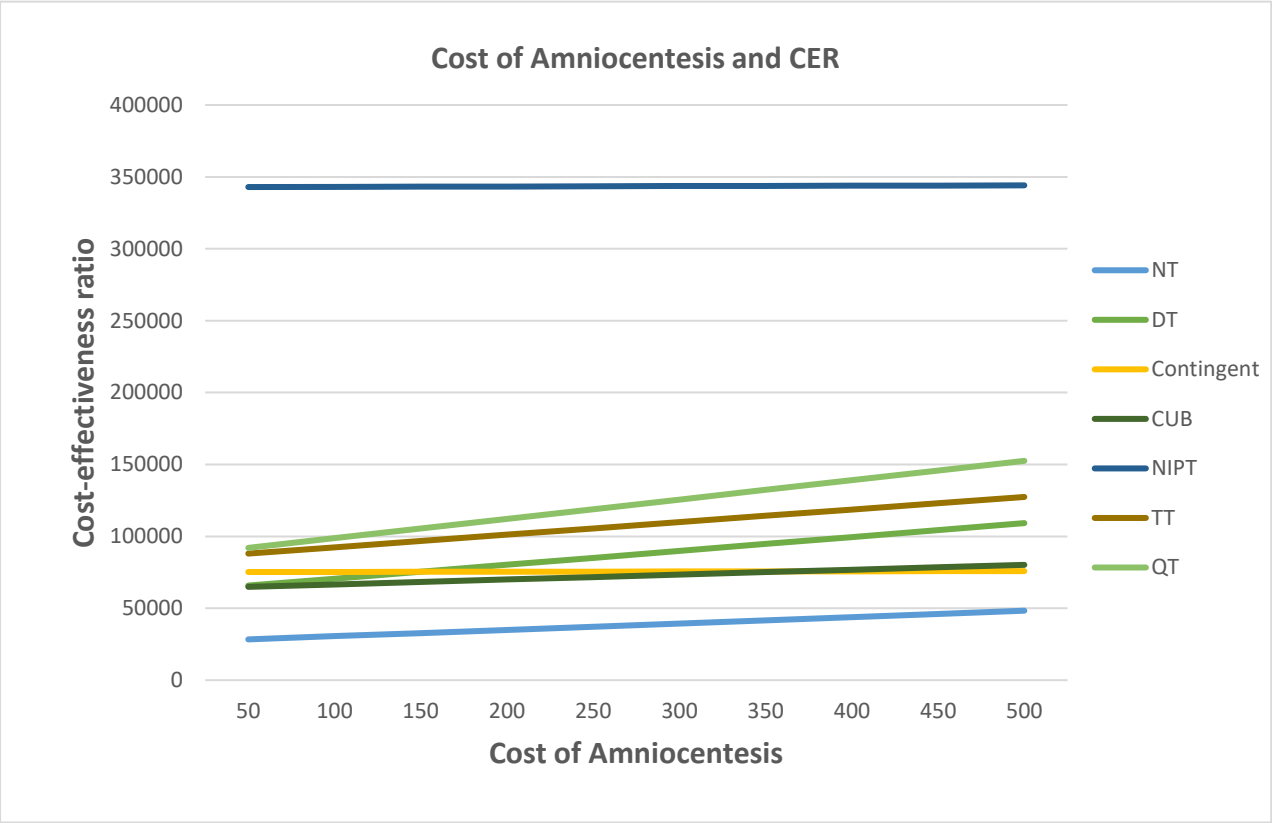


Figure 7 One-way sensitivity analysis (Cost of amniocentesis)

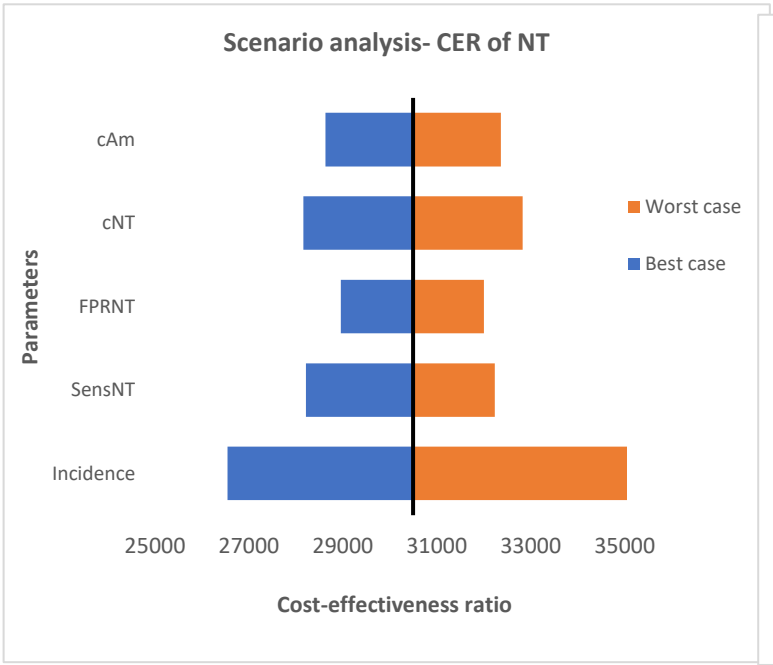
One-way sensitivity analysis of the cost of amniocentesis and the CERs of the strategies are shown in the figure. The line plot shows that an increase in the cost of amniocentesis increases the CERs of almost all the strategies except NIPT and contingent screening. The increase is comparatively higher for DT, TT, and QT as they have a higher proportion of false positives than other tests. The increase is gradual for CUB and NT. Lastly, there is almost no increase in the CER of contingent strategy and NIPT with the increase in the cost of amniocentesis as they have a relatively lower number of pregnant women going through amniocentesis.

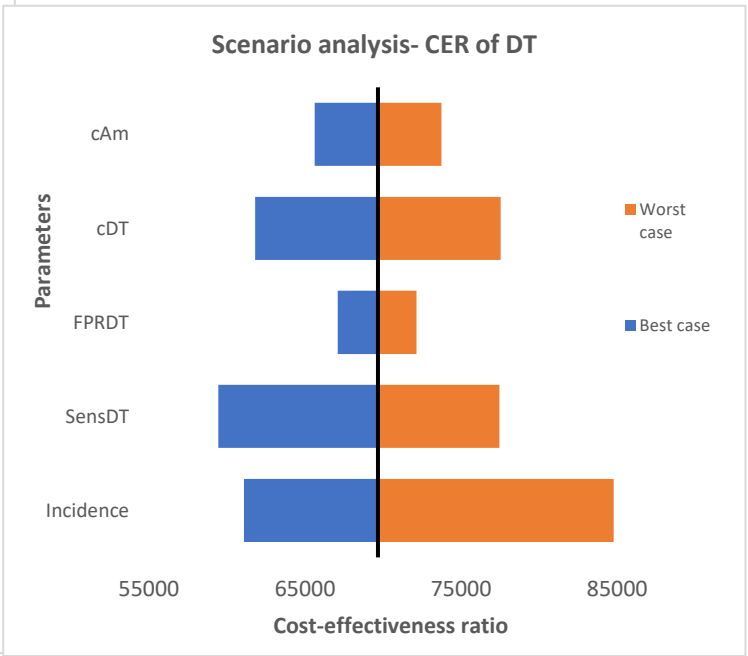
Two-way sensitivity analysis of the CERs of all the interventions was done in the form of scenario analysis. Two scenarios, i.e., Best case (most optimistic values of all the parameters) and worst-case (most pessimistic values of the parameters), were defined to see the performance of all the strategies in the most optimal or the worst case. This analysis shows which parameters are the most responsible for the change of individual CERs. The results are shown in the form of a tornado diagram where the width of the bar signifying the level of influence. For e.g., the widest bar meaning the most influential parameter. The left-hand side

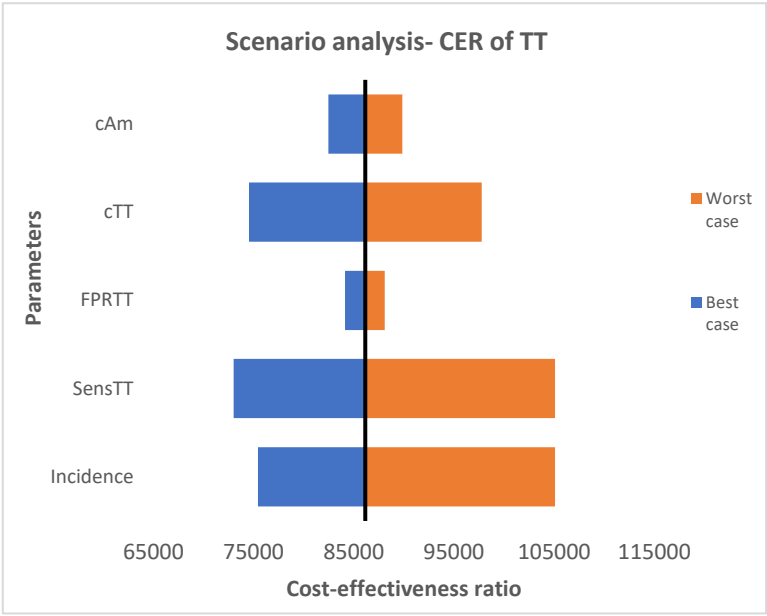
is the best case, and the right hand is the worst-case scenario, with the black line in the middle showing the base case.

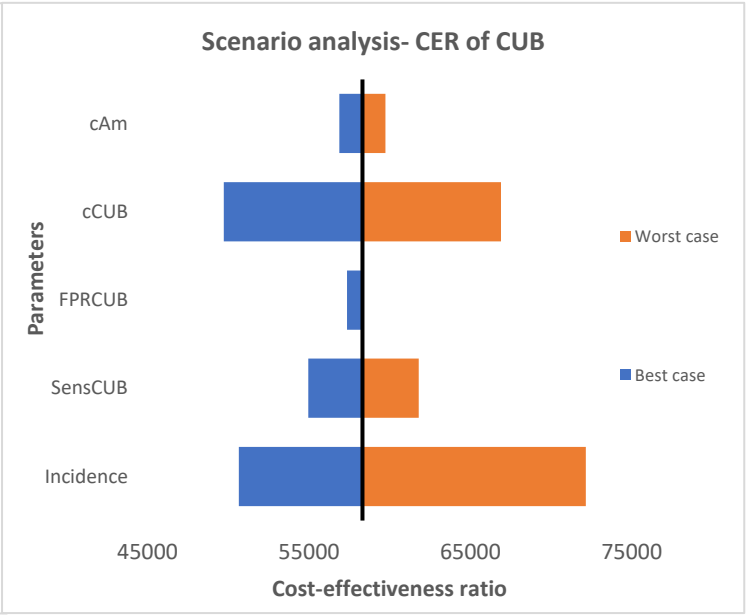
As we can see from the figures, the incidence/birth prevalence is one of the most influential parameters on CERs of all the strategies, with one of the widest bars for most strategies. Similarly, the cost of the individual test also has a huge influence on its own CER, as apparent from the figures. The cost of amniocentesis is a parameter that has varying effects in the CERs of different strategies. The cost of amniocentesis has a similar influence in the CERs of NT, DT, and QT, with the bars considerably shorter than ones for incidence and cost of individual tests. For the CUB test and TT, amniocentesis cost has very little effect on the CER. Lastly, there is no visible bar when the cost of amniocentesis is varied from best to worst for NIPT and contingent test, which means negligible or no effect.

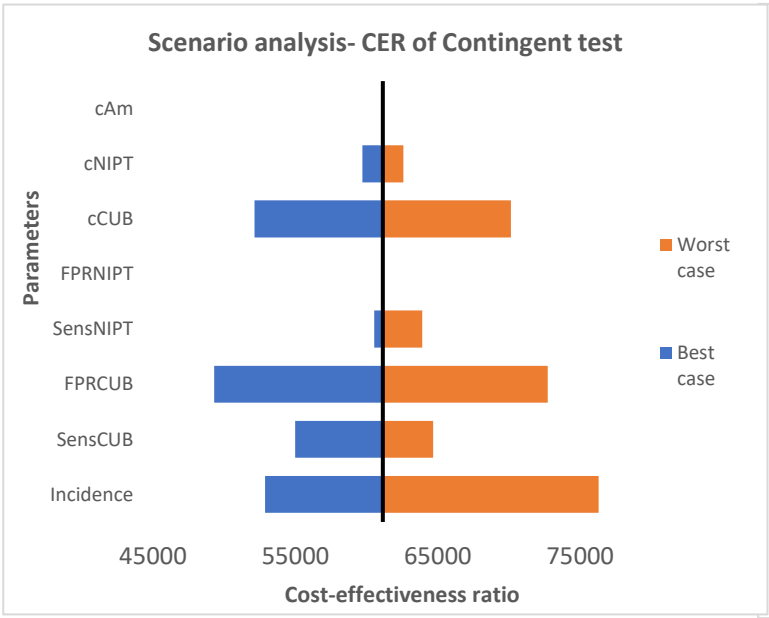
The effects of sensitivity and FPR of individual tests on the CERs are interesting to analyses with strong influences in some and negligible on others. For instance, the sensitivity of DT, TT, and QT varies considerably from the best case to worst case and thus has a strong influence on the CERs, whereas the influence is not as strong for NT, contingent test, and CUB test as the range of values from best to the worst case is smaller for these tests. The sensitivity of NIPT has a very low influence on its CER as the differential value of sensitivity of NIPT across scenarios is very small.

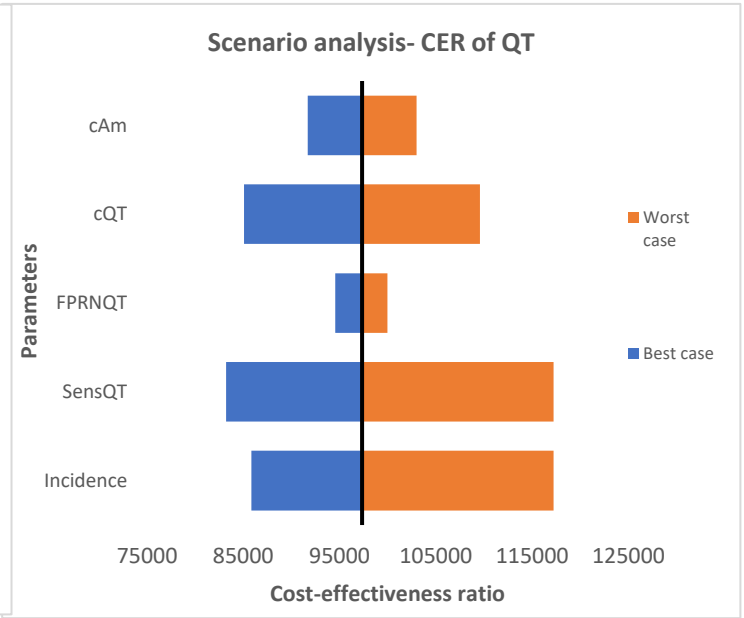





Scenario analysis- CER of TT
cAm
cTT
FPRTT
SensTT
Incidence
Parameters
Worst case
Best case
65000
75000
85000
95000
105000
115000
Cost-effectiveness ratio


Scenario analysis- CER of CUB
cAm
cCUB
FPRCUB
SensCUB
Incidence
Parameters
Worst case
Best case
45000
55000
65000
75000
Cost-effectiveness ratio


Scenario analysis- CER of Contingent test
cAm
cNIPT
cCUB
FPRNIPT
SensNIPT
FPRCUB
SensCUB
Incidence
Parameters
Worst case
Best case
45000
55000
65000
75000
Cost-effectiveness ratio


Scenario analysis- CER of QT
cAm
cQT
FPRNQT
SensQT
Incidence
Parameters
Worst case
Best case
75000
85000
95000
105000
115000
125000
Cost-effectiveness ratio

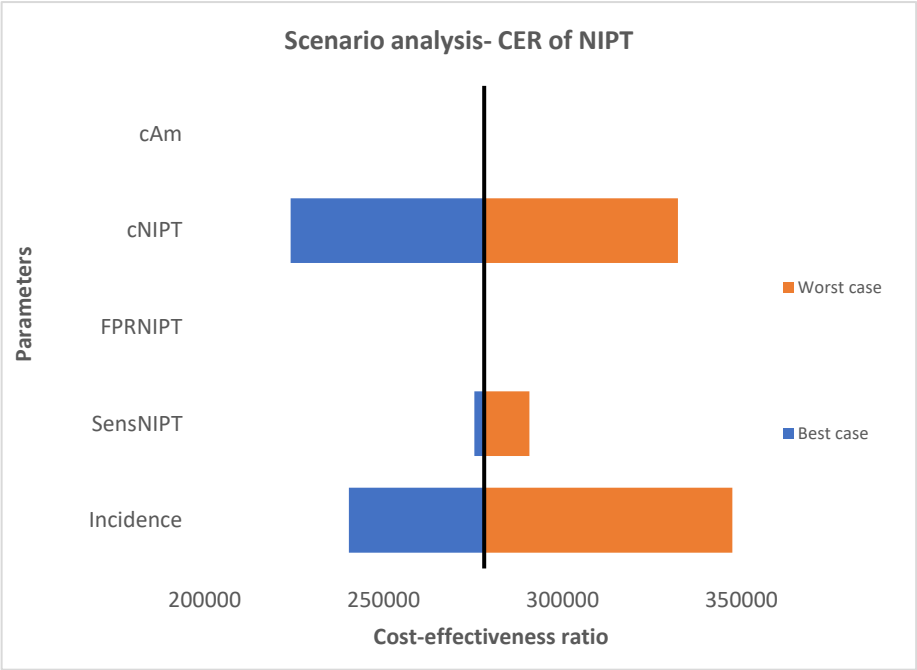


Figure 8 Scenario analysis

The FPRs of tests are the least influential parameter in the CERs of the tests apart from contingent tests and result in interesting findings. The only strategy where the FPR of a test has very influence in the CER is the contingent test, where the FPR of the first tier CUB test has a bigger influence than the costs or the sensitivities. Apart from that, the FPR has very little (for NT, DT, TT, QT) or negligible (CUB) or no (NIPT) impact on the CERs of the strategies.

5.3.1.2 Incremental cost-effectiveness ratio

Similar to CER, the effect of input variables on the ICERs all the strategies were analyzed through one-way sensitivity analysis.

Input parameters like sensitivities, costs of all tests, cost of amniocentesis, maternal age, and birth prevalence of DS were included in the one-way analysis. The analysis was performed such that the values of all the input parameters were varied: once the highest and once the lowest, and the ICER was reported. The result of the one-way analysis is presented in the table below:

Table 10 One-way sensitivity analysis (ICER)

Parameter varied	Base estimate	Range	ICER (Base case)						
		Max-Min	NT (37,661)	DT (-772769)	Contingent (450289)	CUB (294647)	TT (-120821)	QT (1044070)	NIPT (2377837)
Incidence	0.095	0.114	31690	-557362	348102	244003	-95806	1414989	1618340
		0.076	46439	-1259217	638338	371866	-163422	827287	4481330
Sensitivity-NT	0.7	0.76	34846	-315545	2342577	624404	-	-	-
		0.66	39808	-22088312	292777	217977	-	-	-
Sensitivity-DT	0.62	0.74	-	408294	-	-	-	-	-
		0.55	-	-287203	-	-	-	-	-
Sensitivity-CUB	0.82	0.87	-	-	-4	202201	-85910	-6008004	-
		0.77	-	-	1249141	798	-193103	501301	-
Sensitivity-TT	0.65	0.7812	-	-	-	-	8896585	-	-
		0.5208	-	-	-	-	-59837	-	-
Sensitivity-QT	0.76	0.912	-	-	-	-	-	235889	-4319747
		0.608	-	-	-	-	-	-429044	1962718
Sensitivity-NIPT	0.99	1	-	-	-	-	-	-	1833907
		0.947	-	-	-	-	-	-	2812095
Cost of NT	8.00	9.6	40544	-721622	421742	275976	-	-	-
		6.4	34779	-823917	478836	313318	-	-	-
Cost of DT	25.44	30.528	-	-34297	-	-	-	-	-
		20.352	-	-610120	-	-	-	-	-
Cost of CUB	34	40.8	-	-	571614	373998	-50674	840691	
		27.2	-	-	328964	215296	-190968	1247449	-
Cost of TT	39	46.8	-	-	-	-	-201284	-	-
		31.2	-	-	-	-	-40359	-	-
Cost of QT	51	61.2	-	-	-	-	-	1349138	2239983
		40.8	-	-	-	-	-	739002	2515692
Cost of NIPT	250	300	-	-	-36285	-	-	-	3053594
		200	-	-	430880	-	-	-	1702080
Cost of amniocentesis	260	312	39970	-3597	428160	292896	-134670	1151195	2315502
		208	35353	-729717	-36367	296397	-106973	936945	2440172

The result from the sensitivity analysis supports the results from the deterministic analysis. Although the values of ICER change with the change in the parameters, NT always has the

lowest ICER compared to no screening. DT is strongly dominated in most cases as it is associated with more cost and less outcome than NT. The contingent is extendedly dominated by the CUB test in almost all cases. The strategy that can get closest to NT in terms of ICER is CUB, but even in that case where the sensitivity is set to maximum, its ICER is $202,201 per additional case detected. No change in the sensitivities or cost of any other test results in a different decision being taken. A decrease in the cost of an individual test decreases the ICER of that particular test, but there is no change in the most cost-effective strategies.

5.3.2 Probabilistic analysis

5.3.2.1 Cost-effectiveness ratio

The result from the probabilistic analysis of 1000 Monte-Carlo simulations of CERs of all the strategies is presented in the cost-effectiveness planes below. All the plots represent 1000 values of CERs of each strategy from the probabilistic model. There are three planes with CERs of strategies in the general population and pregnant women who are $\geq$35 years and $\geq$40 years.

When the population under analysis are all the women of reproductive age, NT lies at the most bottom. The elongated plot shows that the cost of NT remained similar, but the number of cases detected differed in 1000 iterations. DT, CUB test, and contingent test cost similar, but the number of cases detected seems to higher for CUB, and the scatterplot of CUB densely populated than the other two, which are more scattered, which means the number of cases detected by CUB was less spread than contingent and DT. Scatterplots of QT and TT lie close to each other with similar costs and outcomes. However, NIPT lies at the top with the highest cost and outcome, which is the least cost-effective of all.

The result is similar for both the other age groups in the sense that the scatterplot of CERs of NT lies at the bottom with that of contingent, CUB and DT lying close to each other above NT. The only difference between these two groups compared to the general population is that the scatterplots of NIPT lie below than in the plane of the general population. The reason being the population of both groups are smaller than the general population and thus less cost associated. Increased incidence should intuitively lead to more cases detected, but the fact that there are more spontaneous miscarriages associated with DS pregnancy in $\geq$40 years old age group, the number of cases detected does not differ much and rather decreases.

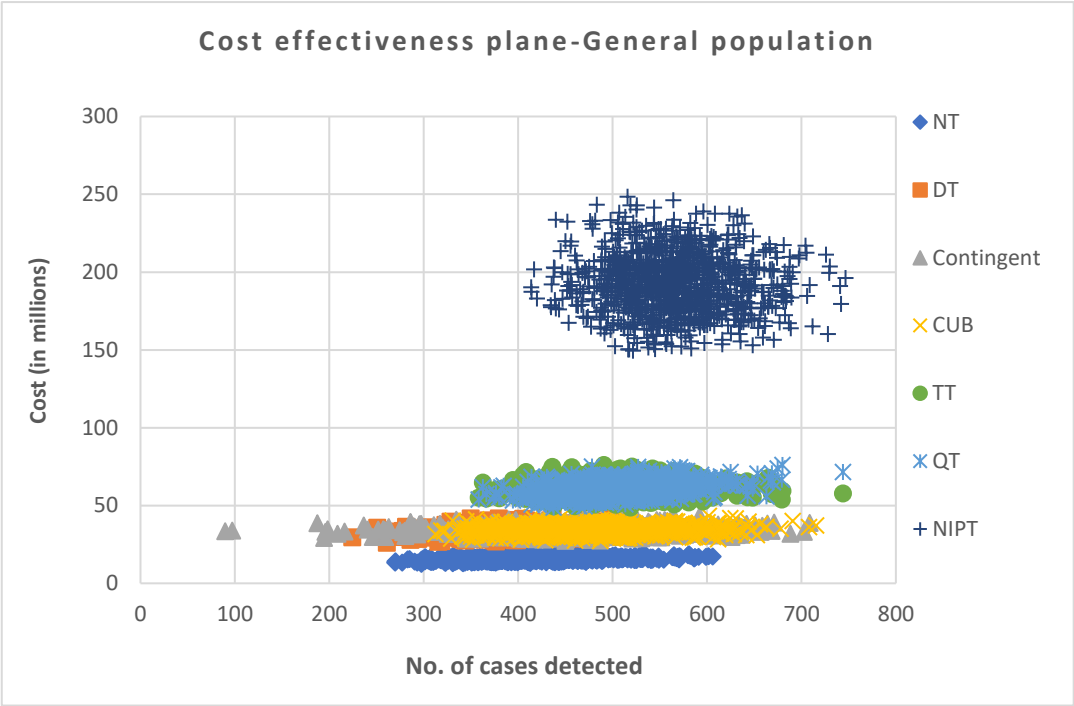


Figure 9 PSA of CER (general population)

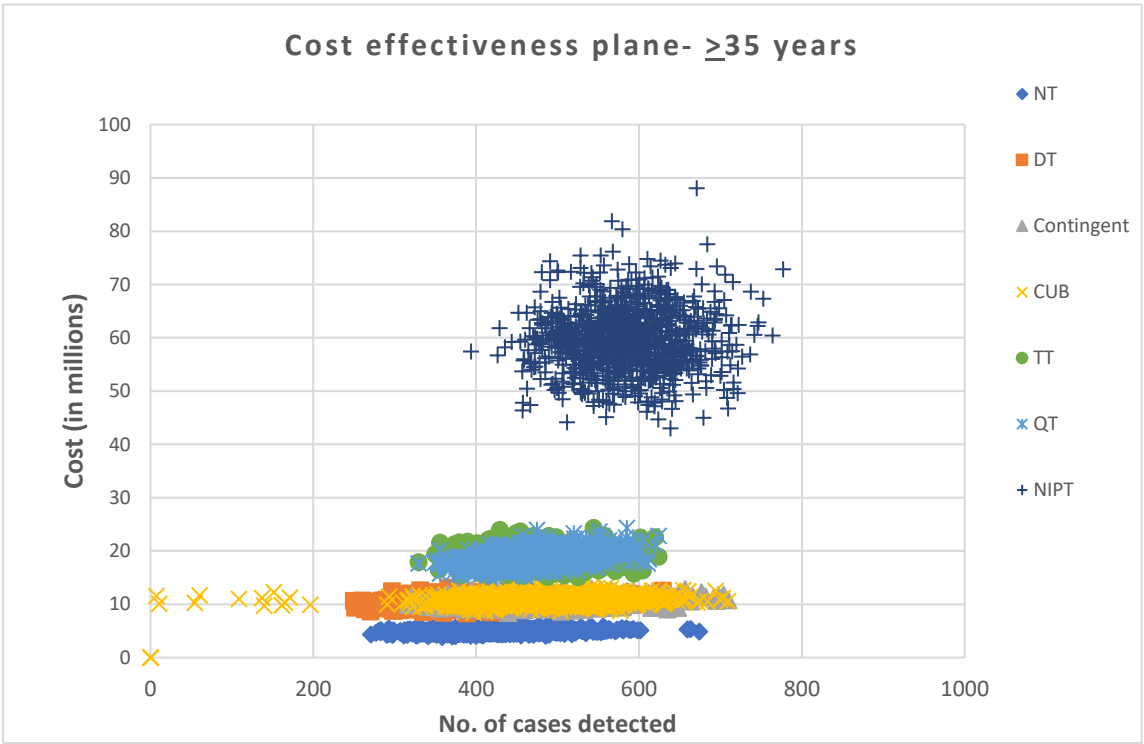


Figure 10 PSA o CER (≥35 years)

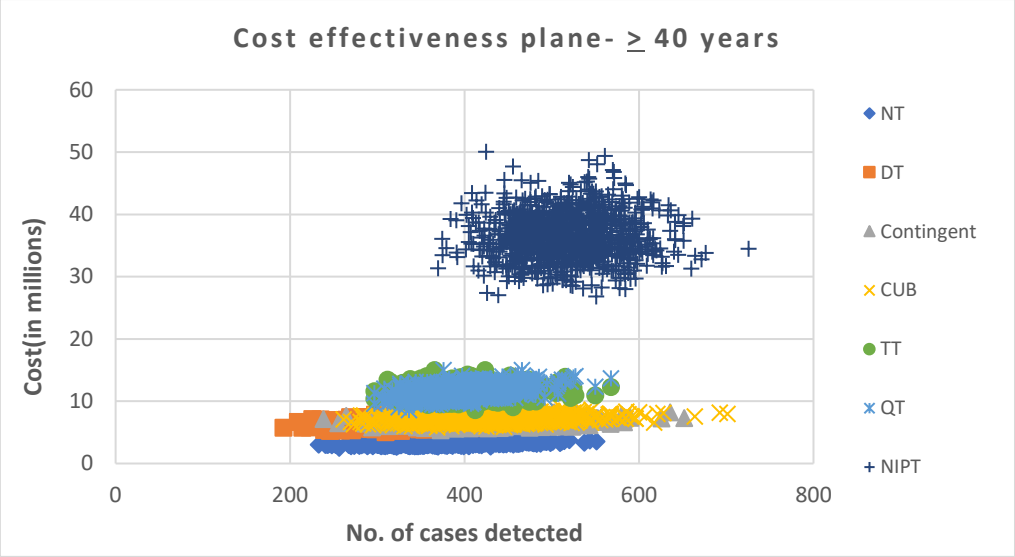


Figure 11 PSA of CER (≥40years)

5.2.2.2 Incremental cost-effectiveness ratio (ICER)

The result from 1000 Monte-Carlo simulations of the values of ICERs is presented in the figure below in the cost-effectiveness plane where the x-axis represents the incremental cases detected, and the y-axis represents an incremental cost. The red line represents the willingness to pay threshold, which is assumed to be $3,000 dollars per case detected. Two interventions: NT and the CUB test have been included in the simulation as they have been found to be the most cost-effective of all the strategies from the deterministic analysis.

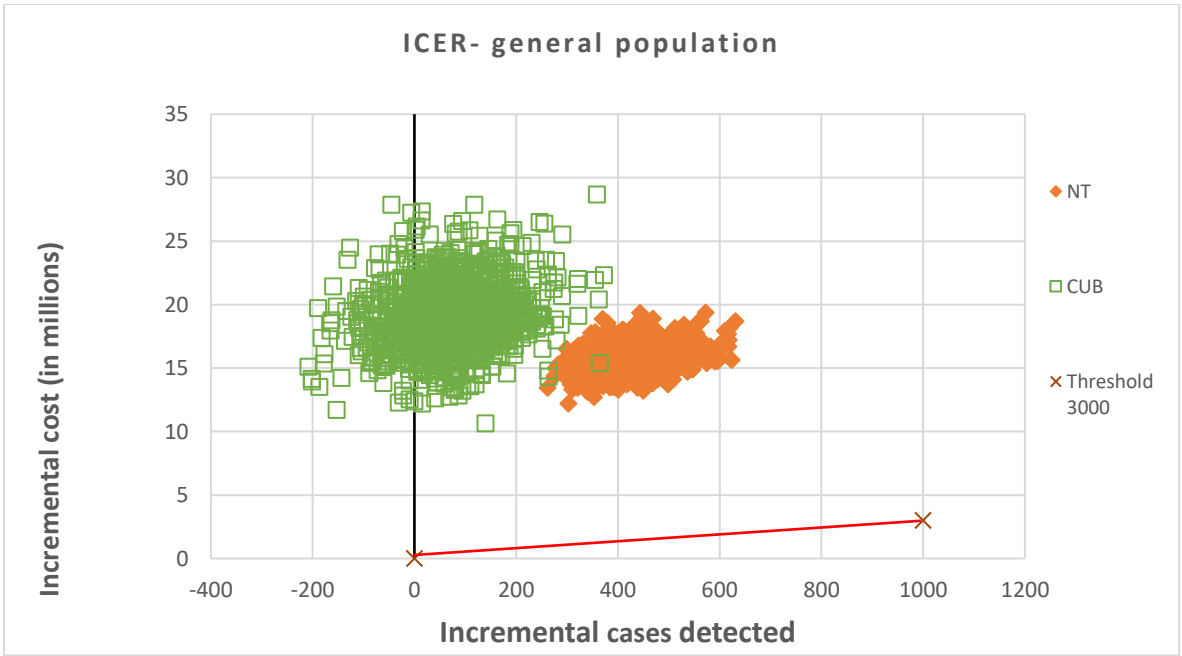


Figure 12 PSA- ICER (General population)

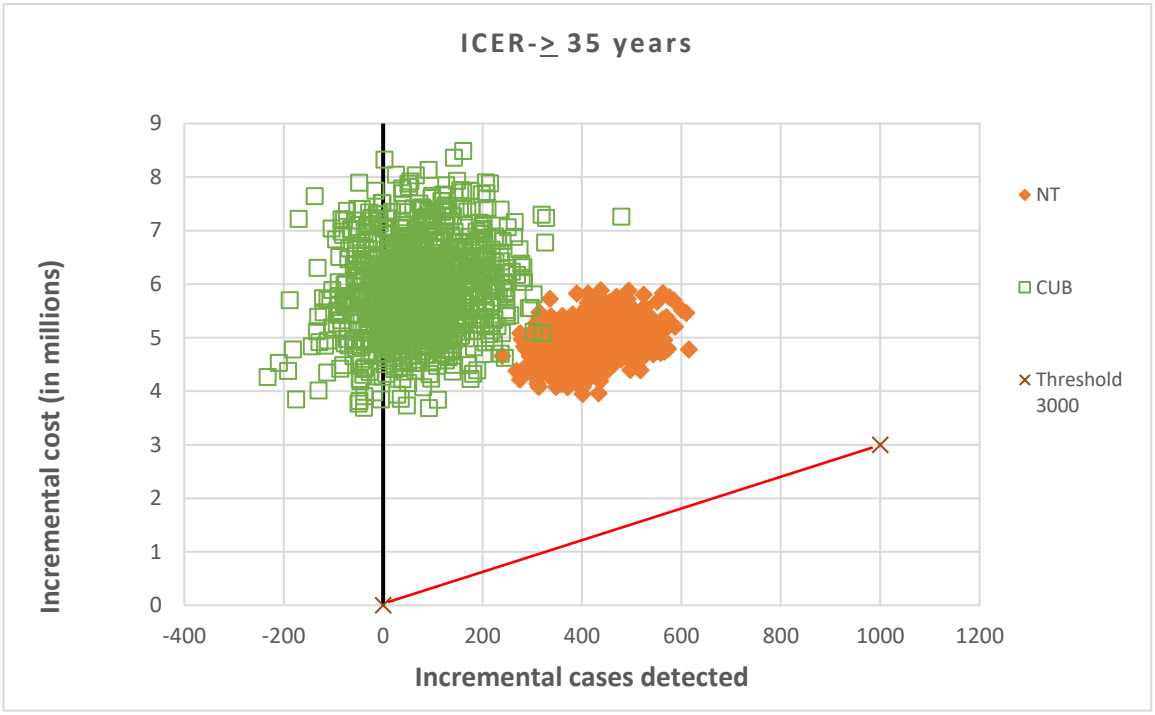


Figure 13 PSA- ICER (≥35 years)

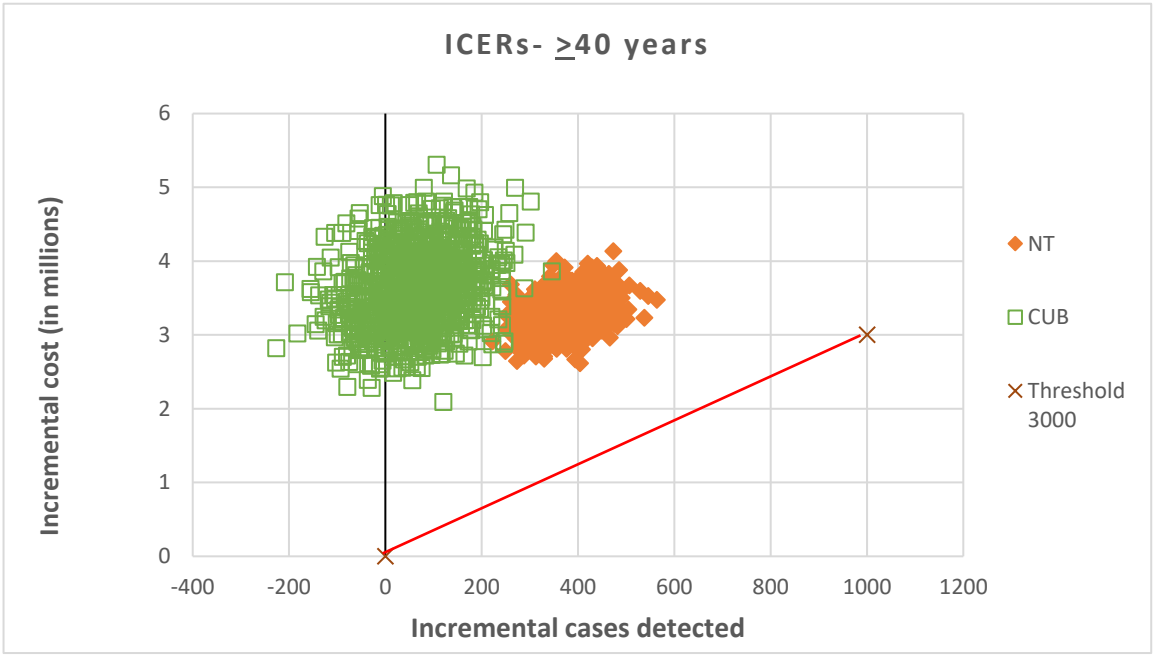


Figure 14 PSA- ICER (≥40 years)

As seen in figure 12, the ICER of NT from 1000 Monte-Carlo simulations lies in the northeast quadrant of the plane for the general population. This means that NT is more costly and has more effects than no screening, which makes obvious sense as zero cost and outcomes are associated with no screening. From the 1000 ICER plots of the CUB test compared to NT, it can be seen that approx. 70% of the plots lie in the northeast quadrant, which means it results in more effects at a higher cost 70% of the time. However, approx. 30% of the time, the ICER of CUB lies in the northwest quadrant. This means that 30% of the time, CUB is more costly with fewer effects and is not cost-effective compared to NT at any threshold. Lastly, all the scatterplots lie way above the willingness to pay threshold. Thus, at a threshold of $3,000 per case detected, none of the two strategies are cost-effective.

Similarly, scatterplots of ICERs from 1000 Monte-Carlo simulations for women $\geq$ 35 years and $\geq$40 years can be seen in figures. The pattern of the scatterplot is similar to the general population for both age groups. For women $\geq$ 35 years, the ICERs get closer to the threshold, but neither of them is cost-effective at the threshold. Similarly, for $\geq$40 years and older women, the ICER of both NT and CUB lie well above the willingness to pay threshold. The total cost associated with the test decrease sharply without an increase in the number of cases detected. Thus, no strategy is cost-effective compared to no screening at a threshold of $3,000 per additional case detected.

The result from the probabilistic analysis has been used to produce CEAC and CEAF curves which are presented in figures 15 and 16. The x-axis represents the willingness to pay threshold, and the y-axis represents the probability of cost-effectiveness. In the CEAC, it can be seen that no screening is the only preferred strategy till about the threshold of $25,000 per incremental case detected. From that point, the probability of cost-effectiveness of no screening decreases, and that of NT increases. At around a threshold of $38,000, both no screening and NT are equally cost-effective, and beyond that point, NT takes over as the more cost-effective strategy. The probability for NT increases to a maximum of 1 at a threshold of $60,000, at which CUB starts emerging as an alternative. The probability of the CUB test increases gradually with an increase in threshold, whereas that of NT decreases at a similar rate. At a threshold of $280,000, both NT and CUB are equally cost-effective, and CUB takes over beyond that point. This also can be seen from the CEAF.

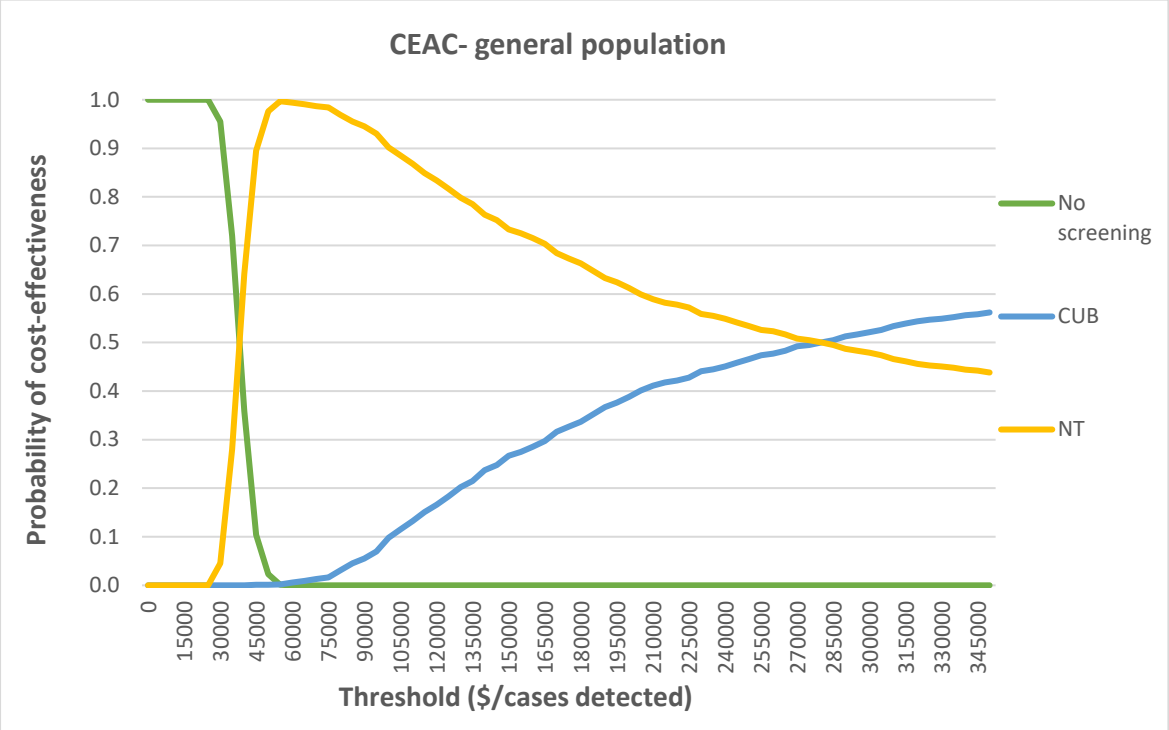


Figure 13 CEAC (General population)

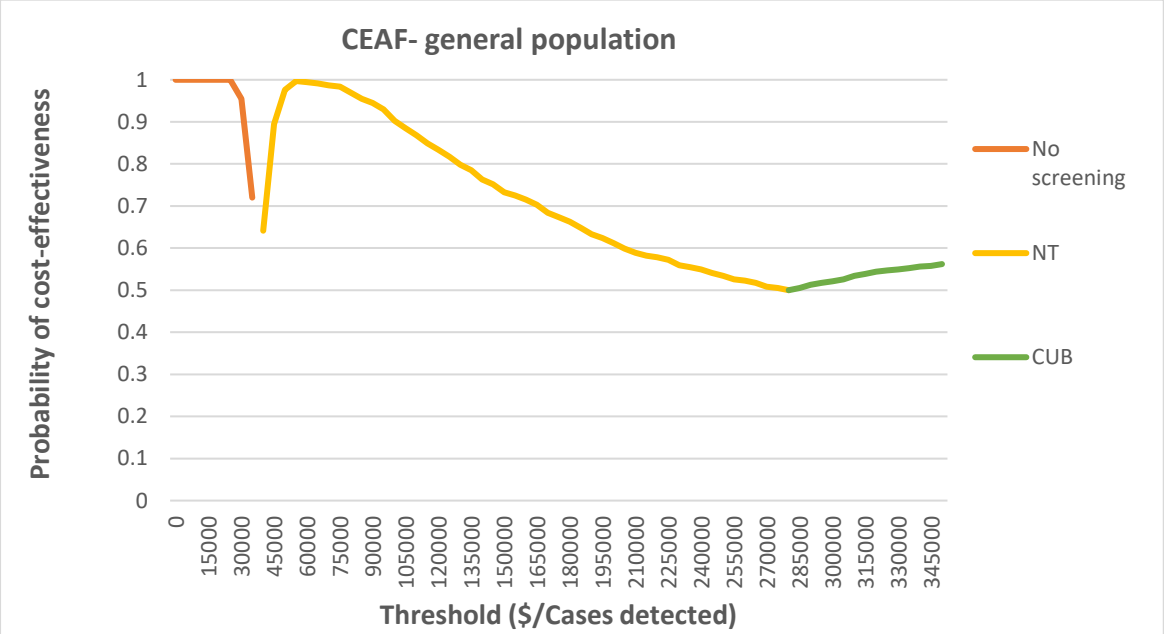


Figure 14 CEAF (General population)

The CEAC curves when women ≥35 years and ≥40 years were considered for analysis are presented in figures 16 and 17, respectively. Compared to the general population, where the

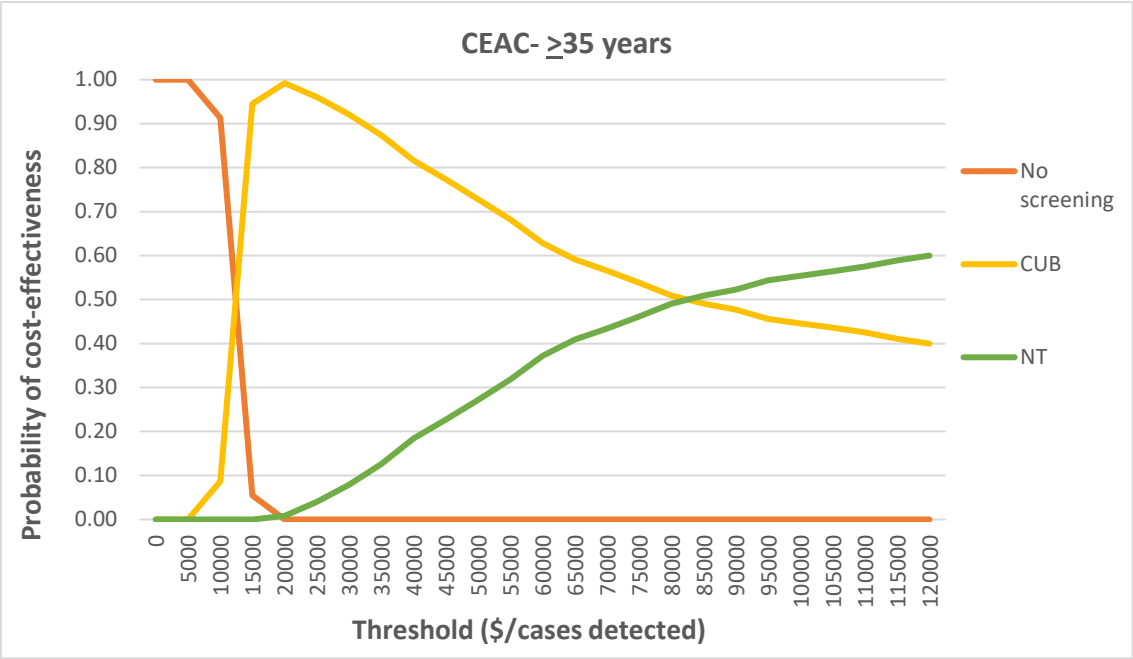


Figure 15 CEAC (≥35 years)

switch from no screening to NT happened beyond the threshold of $38,000, the switch for women ≥35 years happens at a threshold of around $12,000 per additional case detected, which is still higher than the threshold assumed for the study. At a threshold of $85,000 and beyond, CUB takes over as the cost-effective strategy from NT.

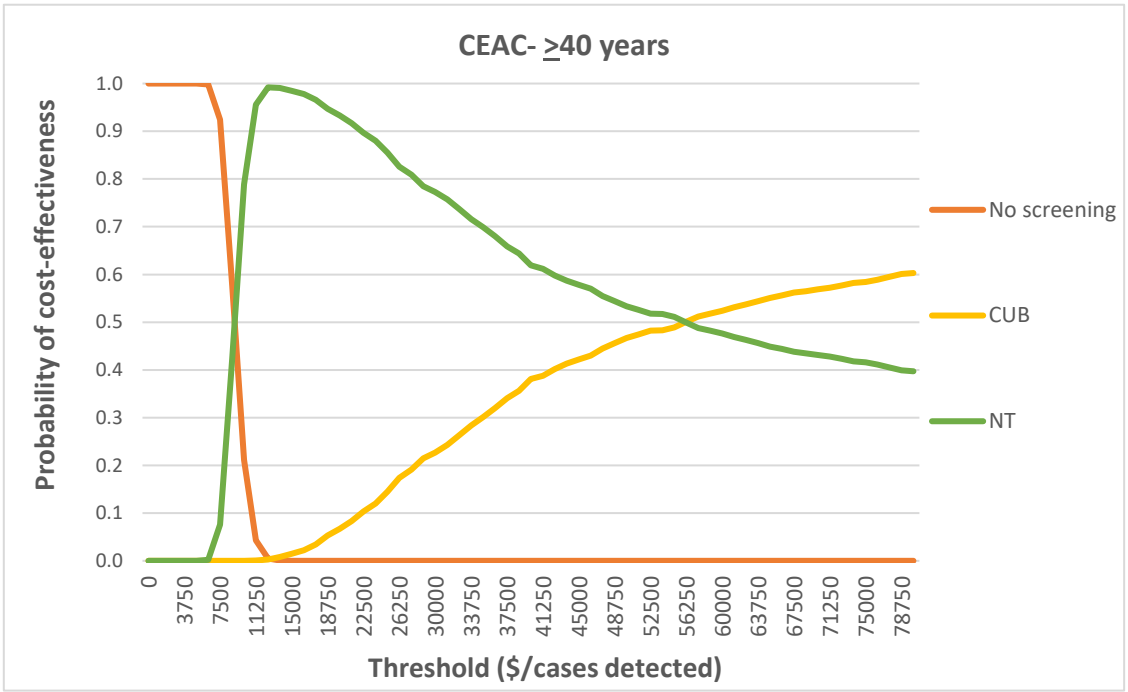


Figure 16 CEAC (≥40 years)

Similarly, the CEAC for women ≥40 years shows the probability of cost-effectiveness of the three strategies at different thresholds. In this case, the probability of cost-effectiveness of NT is more than 0.5 at a threshold somewhere between $9,000. This means that NT is still not cost-effective at the threshold assumed for the study and no screening is the preferred choice for women ≥40 years in the present context. The probability of cost-effectiveness of NT peaks at 0.93 at a threshold of $13,750, from where the probability of cost-effectiveness of CUB increases. At a threshold between $56,250, the CUB test takes over as the more cost-effective strategy.

5.4 EVPI

As apparent from the CEAC and CEAF curves, there is uncertainty in choosing one strategy over another along with different values of the WTP threshold. These uncertainties can be presented in the form of a combined CEAF and EVPI curve shown in the figure.

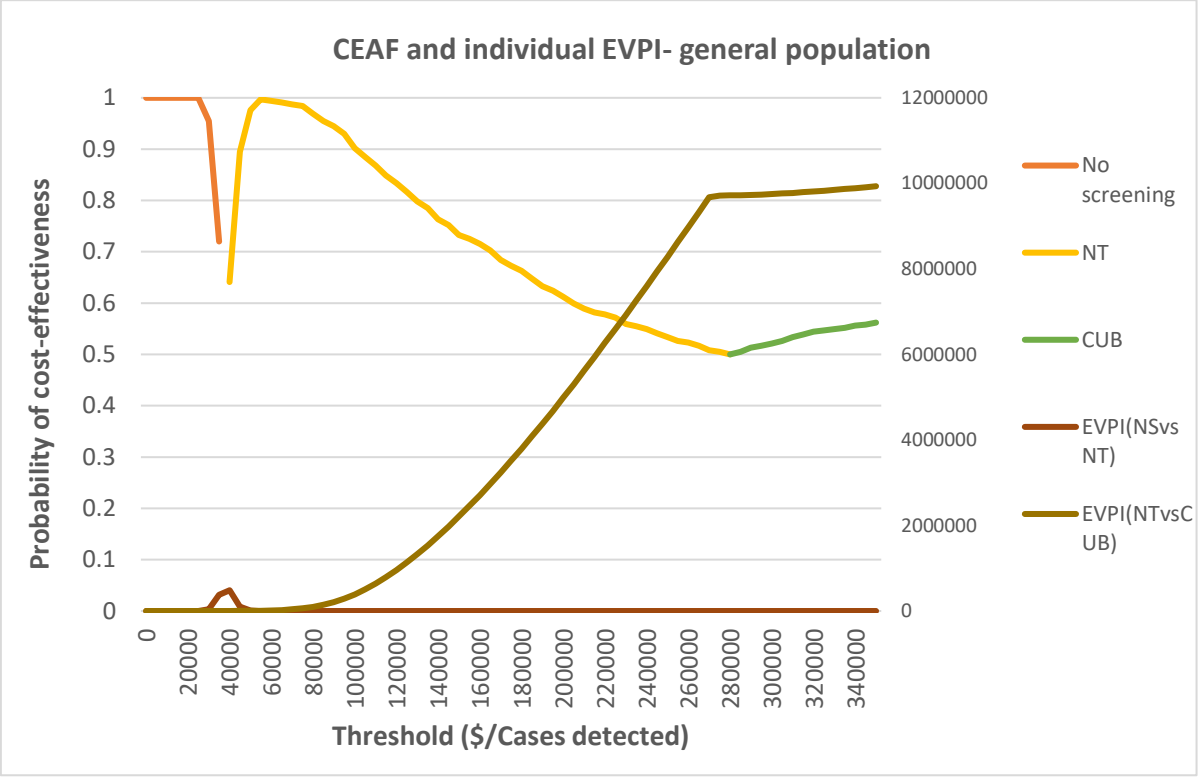


Figure 17 CEAF and Individual EVPI (general population)

The figure above shows the EVPI and CEAF for the general population. At a threshold of $38,000 and $280,000 per additional case detected, there is a switch from no screening to NT and NT to CUB as the preferred strategy. The probability of cost-effectiveness of all the strategies changes along the threshold. The value of probability determines the decision uncertainty associated with making the switch. There is no or very low decision uncertainty

when one has higher cost-effectiveness than the other one. For e.g., up to a threshold of $20,000, NT is 100% cost-effective than NT, and thus the decision can be made with total certainty. However, beyond that threshold, the probability of NT being cost-effective increases, and so does the decision uncertainty. At a threshold of $38,000, there is maximum decision uncertainty as both NT and no screening are equally cost-effective, and making a decision on either strategy can be wrong. The value of EVPI reaches its peak of 737790. This means that the collection of perfect information about all the parameters is most worthwhile at this threshold. The value of EVPI decreases from that point as NT is the preferred choice.

Similarly, the value of EVPI decreases to a minimum when the probability of NT being cost-effective almost reaches one. From that point, CUB comes into the picture as the probability of CUB being cost-effective increases, and that of NT decreases. This results in an increase in the value of EVPI. The value of EVPI increases quite steeply up to a threshold of $275,000, where NT and CUB have an equal probability of being cost-effective. The value of EVPI is 9,714,344. The value of EVPI increases gradually after that point as the probability of cost-effectiveness of CUB increases likewise, which means there is an increase in decision uncertainty after that point.

Thus, for the general population at a threshold of $3,000 per additional case detected, the collection of perfect information of all the parameters is not worthwhile, and the decision is made based on the available information.

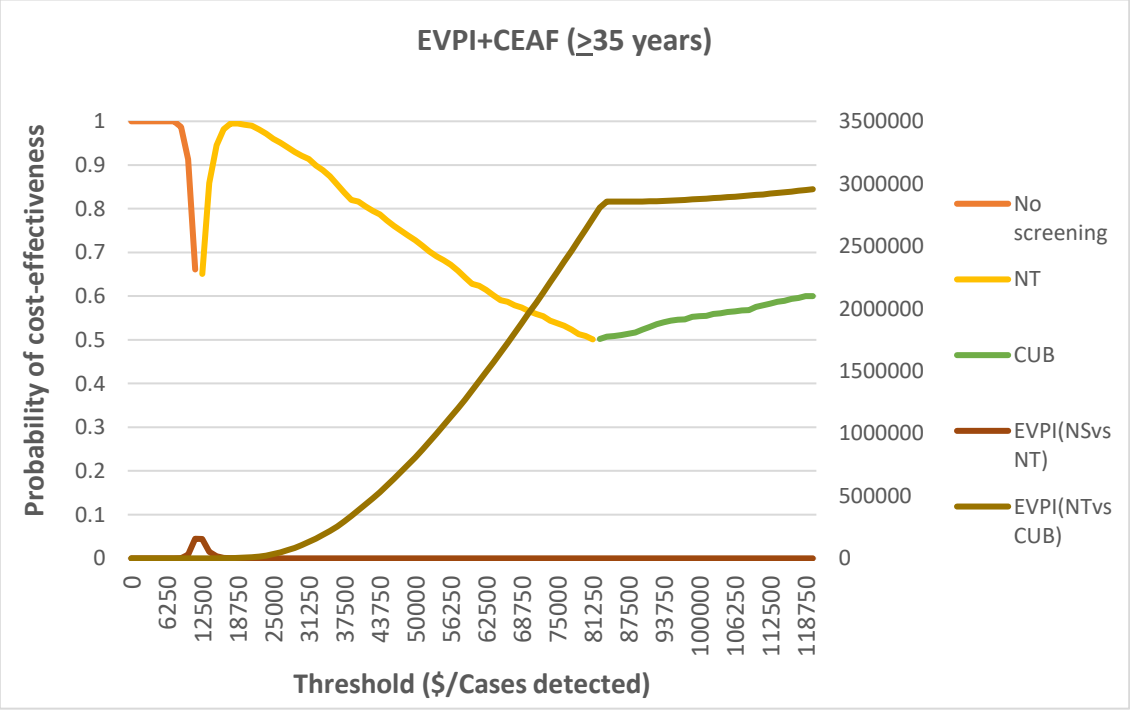


Figure 18 EVPI and CEAF (≥35 years)

The figure above shows CEAF and EVPI when women ≥35 years are taken as the subjects of the study. Similar to the previous analysis, the collection of perfect information is not worthwhile at a threshold of $3,000 per additional case detected. There is some value attached to the collection of information only beyond the threshold of $8,750 up to around $18,750, from where the value of EVPI is very close to zero. The value of EVPI increases again beyond that point, meaning that the collection of effect information of all parameters is worthwhile beyond the threshold of $18,750 as the probability of CUB being a cost-effective strategy increases.

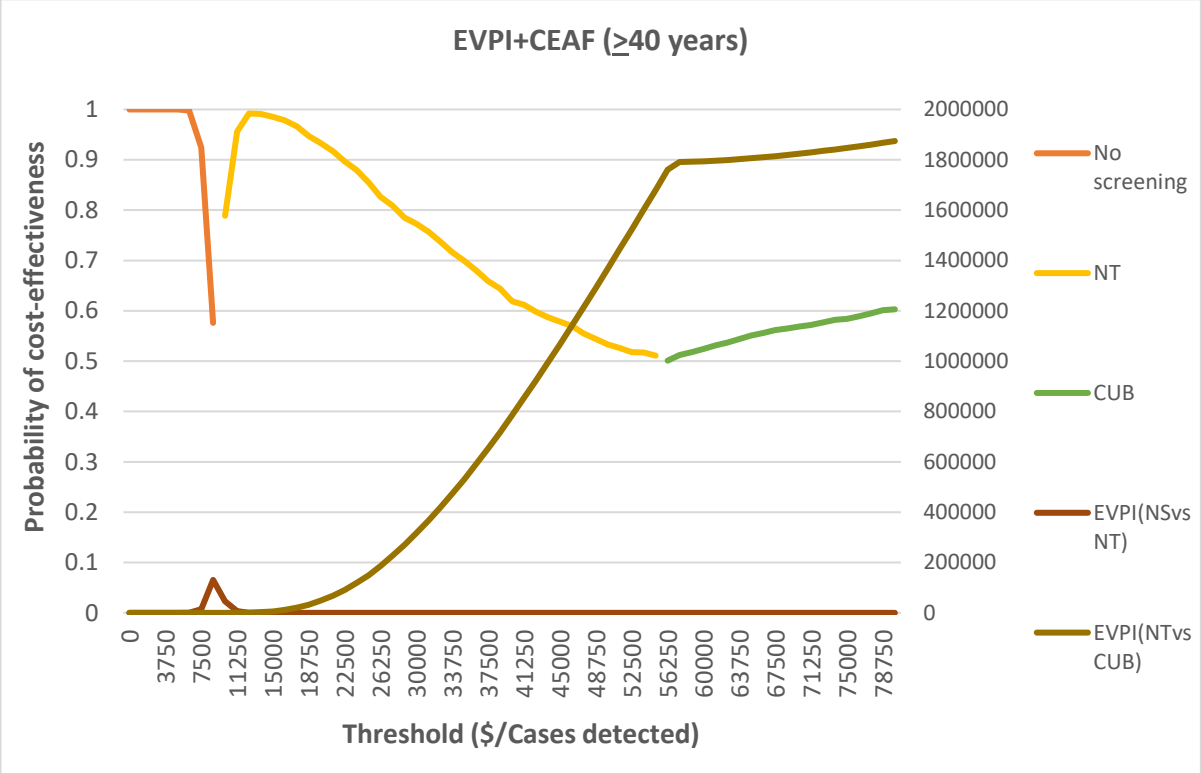


Figure 19 EVPI and CEAF (≥40 years)

The case when pregnant women ≥40 years are taken as subjects of the study is no different compared to previous analyses. The individual EVPI increase to a peak of 46,758 at a threshold of $10,000 per additional case detected as both NT and CUB have an equal probability of being cost-effective. Beyond that point, as NT emerges as more cost-effective than NT, the value of EVPI decreases towards zero. However, decision uncertainty increases as the probability of CUB being cost-effective increases. The increase is steep up to the point where there NT and CUB have equal probability at a threshold of around $57,000, and the increase becomes gradual after that point. This means that there is always decision uncertainty when the choice is between

NT and CUB, and perfect information is worthwhile. Even for a population of 40 years and older, no study on the collection of perfect information of parameters is worthwhile.

5.5 Budget impact analysis

The table shows the budget of all the strategies for the general population, women 35 years or older, and 40 years and older (*blue, orange, and grey bars*).

Table 11 Budget impact analysis

Interventions	Total budget (5 years- in $)		
	General population	≥35 years	≥40 years
Nuchal translucency	78663109	24637115	16357402
Double test	169617825	52523385	33001838
Contingent screening	173621073	53921485	34335052
CUB test	173666567	54434652	36318204
Triple test	217734556	67336683	42037300
Quadruple test	305011338	94208788	58437387
NIPT	966986325	297904867	182364210

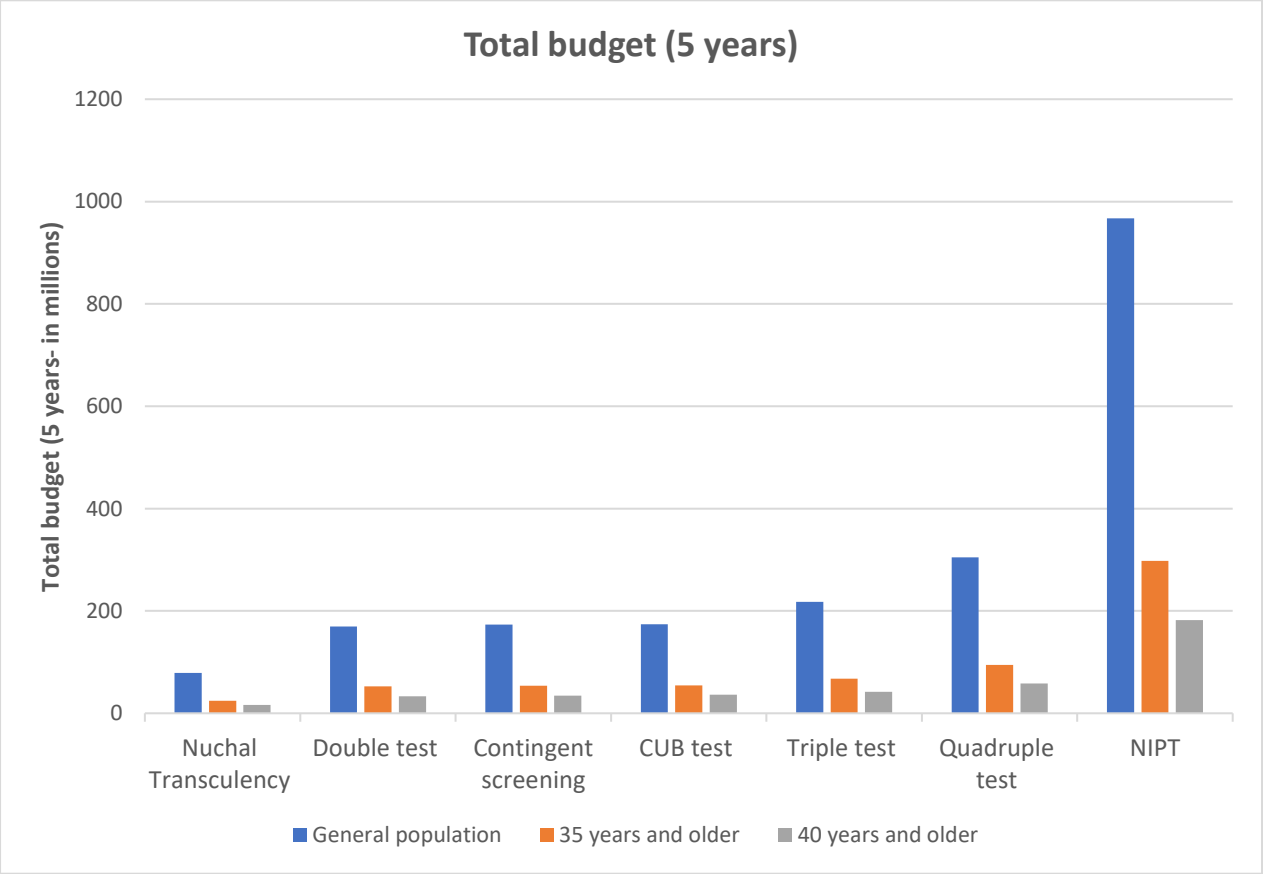


Figure 20 Total budget (5 years)

NT has the lowest budget for all age groups, around $78.66 million, $24.63 million, and $16.35 million for the general population, women 35 years or older, and 40 years and older, respectively. DT, contingent test, and CUB test have a similar 5-year budget with the highest for CUB among the three: $173.66 million, $54.43 million, and $36.31 million (general population, women 35 years or older, and 40 years and older, respectively). NIPT has the highest total budget, which is $966.98 million, $297.90 million, and $182.36 million (general population, women 35 years or older, and 40 years and older, respectively).

Figure 21 also shows the total budget required over the course of 5 years for all the interventions. The overall budget increases gradually from left to right. Similarly, the budget for every intervention is lower for women ≥35 years and the lowest for ≥40 years as there are considerably lower numbers of women in the two age groups compared to the general population. The difference between the three groups is higher for NIPT, as seen in the figure.

5.6 Safety index

In this section, the safety of every intervention in terms of total fetal loss due to amniocentesis is presented. All the interventions will be analyzed using the safety index, which is a ratio of the total number of fetal loss due to amniocentesis as a result of the test and the number of cases detected by the test. A strategy to be better has to have a lower safety index than others.

Table 12 Safety Index

Intervention	**Safety index**		
	General population	**≥35 years**	**≥40 years**
Nuchal translucency	0.35956	0.117327	0.105753
Double test	0.7815	0.257229	0.203896
Contingent screening	0.01047	0.010279	0.032684
CUB test	0.27501	0.090713	0.087426
Triple test	0.70851	0.232145	0.186073
Quadruple test	1.08883	0.367223	0.283252
NIPT	0.021	0.013399	0.034789

As seen in the table, the safest intervention is contingent screening, with a safety index of 0.0104. 0.0102 and 0.03 (general population, ≥35 years and ≥40 years) for three groups, which

has two tiers of screening test and thus, very few women are going through amniocentesis. The next safe intervention is NIPT with a safety index very close to contingent; 0.021 (general population), 0.013 (≥35 years), and 0.03 (≥40 years). CUB and NT are the next safe strategy with similar safety index values. QT is the least safe strategy with the highest safety index values of 1.088 (general population), 0.36 (≥35 years), and 0.28 (≥40 years).

5.7 Harm to Benefit Analysis

Table 12 shows the deterministic values of harm to benefit analysis. In this analysis, harm is measured in terms of the number of amniocenteses performed and benefit in terms of the number of cases detected. The incremental benefit harm ratio is calculated for each strategy. The strategy with the lowest harm to benefit ratio is the preferable strategy.

Table 13 Harm to benefit analysis

Interventions	Harm	Benefit	Incremental harm	Incremental benefit	Incremental harm to benefit ratio	Remarks
Contingent screening	595	460	595	460	1.29	
NIPT	1460	563	865	103	8.38	
CUB test	16372	482	14913	-81	-184.48	Dominated
Nuchal Translucency	18543	418	17084	-145	-117.56	Dominated
Triple test	35800	409	34340	-154	-223.30	Dominated
Double test	38033	394	36573	-169	-216.58	Dominated
Quadruple test	68205	507	66745	-56	-1198.75	Dominated

Contingent screening requires the least number of amniocentesis (595) followed by NIPT (1460) and CUB (16,372). QT, DT, and TT have the highest number of amniocentesis (in descending order). Contingent screening is the most dominant, with a ratio of 1.29, which means contingent only requires 1.29 amniocentesis per case detected, followed by NIPT with 8.38 (8.38 amniocenteses performed per case detected). All the other tests are dominated as they have increased harm with lower benefits than NIPT.

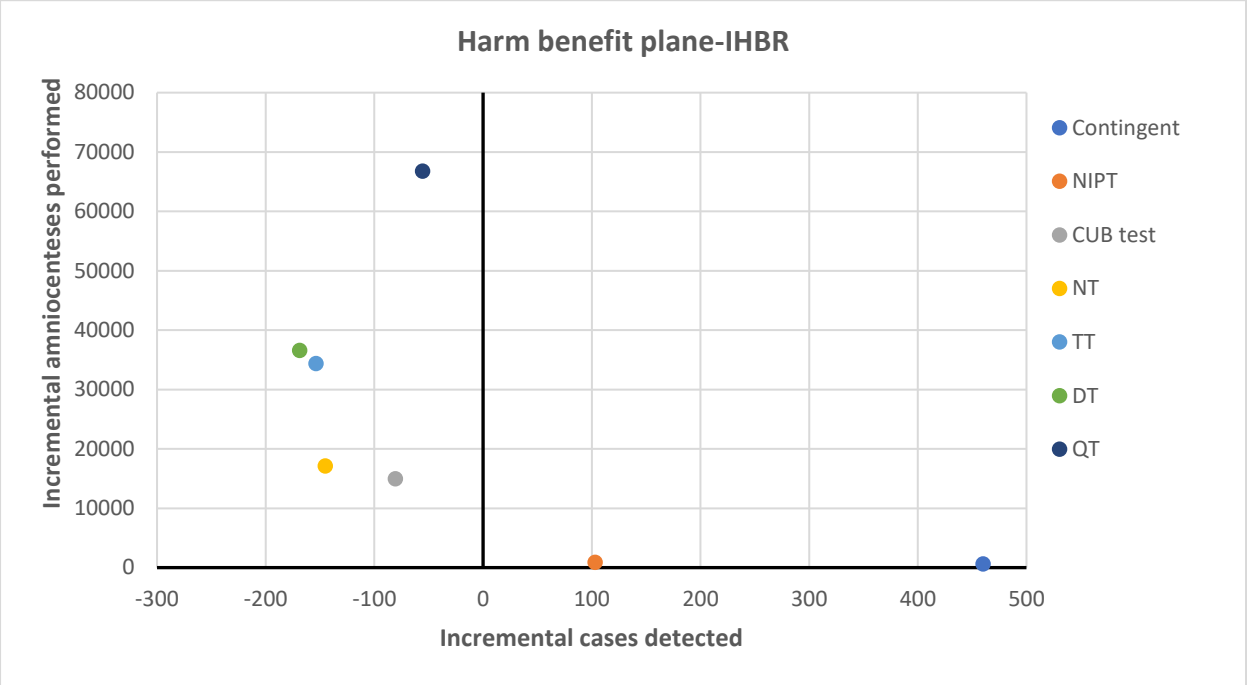


Figure 21 Incremental harm benefit plane

The result is also presented in an incremental harm benefit plane where the x-axis represents an incremental benefit, and the y-axis represents incremental harm. Both contingent and NIPT lie almost on the horizontal axis in the northeast quadrant. However, the incremental cases detected for contingent is higher than NIPT and thus lies far-right to NIPT (as contingent is compared to no screening and NIPT is compared to contingent). All the other strategies fall in the northwest quadrant, which means they result in greater incremental harm but lower incremental benefit.

The result of harm to benefit analysis is similar when women ≥35 years and ≥40 years are taken as subjects of the study i.e. contingent screening and NIPT are the best compared to other strategies.

Chapter VI: Special scenario

A special scenario was developed to assess the conditions under which one of the strategies could be cost-effective at the current WTP threshold.

6.1 Scenario description

Since all the strategies were found to be not cost-effective at the current WTP threshold of $3,000 per additional case detected, a special scenario was assumed for women who were 40 years and older where all the other parameters were estimated to the same as the normal scenario for the age group. This special scenario was created to see under what conditions would any of the interventions would possibly be cost-effective at the current WTP threshold. As seen in the scenario analysis in previous sections, no strategies are cost-effective, even in the best case. Thus, the only way to possibly have one of the interventions cost-effective was to reduce the cost of intervention for health care payers by as much as 68%, i.e., the government of Nepal in this case. The most reasonable and sustainable way of achieving this desired reduction of the cost would be the introduction of coinsurance for these tests, which would be paid by the individual themselves. Among various combinations, 30%-70% was selected where 30% of the total cost would be reimbursed by the health system and the remaining 70% of the medical cost would be the responsibility of the people themselves. The best possible combination of coinsurance percentage was found to be 32%-68%, but the previous combination was adopted for simplicity.

6.2 Deterministic results

The table below represents the CERs and ICERs of all the strategies when there is coinsurance, where 30% of the total cost is reimbursed by the health payer.

Strategy	Cost	No. of cases detected	Total fetal loss	CER	ICER	ICER (Com. to DN)	Remarks
No screening	0	0		#NA			
Nuchal Translucency	6	0.00212	0.00022	2,873	2,873	2,873	
Double test	12	0.00188	0.00038	6,547		6,547	SD
CUB test	13	0.00247	0.00022	5,269	19,565	5,269	
Contingent screening	14	0.00181	0.00006	7,796		7,796	SD
Triple test	16	0.00198	0.00037	8,013		8,013	SD
Quadruple test	22	0.00232	0.00066	9,433		9,433	SD
NIPT	70	0.00299	0.00010	23,531	111,537	23,531	

As seen from the table, NT has the lowest CER of $2873 per case detected, followed by CUB, which has $5269 per case detected. All the other strategies have higher CERs, similar to previous analyses. Similarly, NT has the lowest ICER of all the strategies. The ICER of NT compared to no screening is $2873 per additional case detected, which is lower than the WTP threshold of $3,000 per additional case detected and is thus cost-effective. The next strategy with the lowest ICER is CUB, with a value of 19,565, which is wat higher than the WTP threshold. All other strategies except NIPT are strongly dominated and are excluded from the analysis. NIPT has a very high value of ICER, which is more than 30 times higher than the WTP threshold.

Table 14 NMBs and NHBs (special scenario)

Interventions	Net Monetary Benefits	Net Health Benefits
No screening	0.00	0.00000
NT	0.27	0.00009
DT	-6.68	-0.00223
Contingent	-8.68	-0.00289
CUB	-5.61	-0.00187
TT	-9.91	-0.00330
QT	-14.91	-0.00497
NIPT	-61.31	-0.02044

The NMBs and NHBs of all the interventions in the special scenario are presented in the table above. At the current WTP threshold, the NMBs and NHBs of all the interventions except NT are negative. Negative NMBs mean that the cost of interventions is more than the benefit in monetary terms. Similarly, the negative NHBs of the interventions mean that the health benefits due to the interventions, i.e., the number of cases detected in this case is outweighed by the health losses for the population. Either way, the interventions are not adopted. However, NT has a positive value of NMB, i.e., 0.27, which means that adopting NT as the benefit of screening strategy outweighs the costs by $0.27. Similarly, the NHB of NT in this scenario is positive with a value of 0.00009. This small but positive value indicates that the overall health of the population would increase by 0.00009 if NT is adopted as the preferred choice at the current WTP threshold.

6.3 PSA results

The results from PSA of ICERs of NT and CUB, the two most cost-effective strategies for the current scenario, have been presented in the figure below.

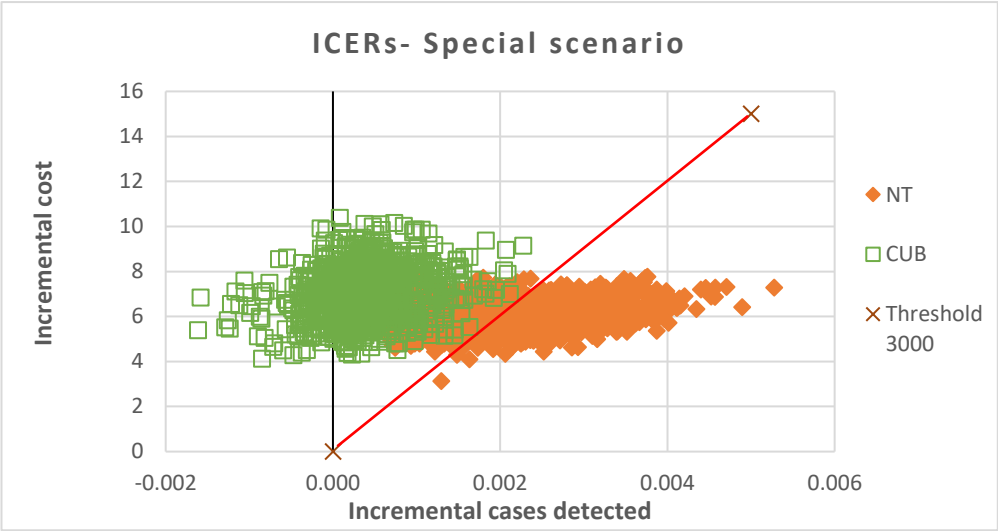


Figure 22 ICER- Special scenario

As can be seen from the figure, almost 52% of the values of ICERs of NT compared to no screening are below the WTP threshold line. This means that NT is cost-effective 52% of the time for the current WTP and should be the preferred strategy. However, 100% of the values of ICER of CUB compared to NT lie above the WTP threshold and thus is not cost-effective.

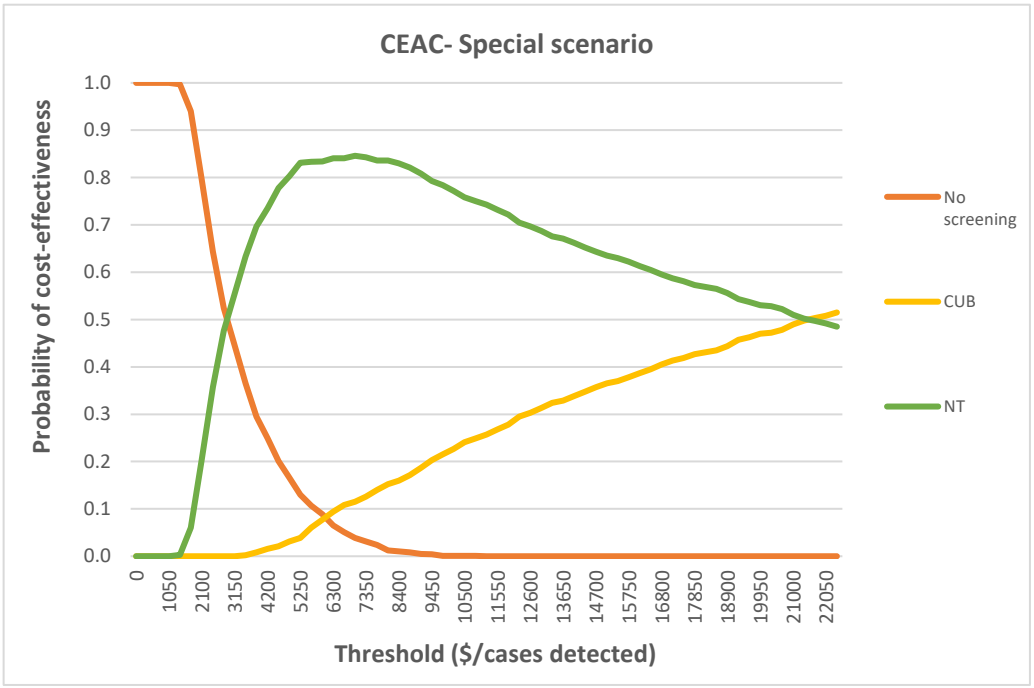


Figure 23 CEAC- Special scenario

The results from PSA can also be presented in the form of CEAC. As apparent from both figures, no screening is the sole preferred strategy till the WTP threshold of about $1400 per additional cases detected. After that point, NT starts emerging as a possible strategy and takes

over as the cost-effective after a threshold of around $2900 per additional cases detected. CUB emerges as a possible strategy after a threshold of $22,500 but is not cost-effective even at a high value of the WTP threshold and thus is not relevant in the present context.

6.4 Value of Information analysis

VOI analysis was also conducted for this special scenario to inform about the decision uncertainty. VOI analysis was done using the results from the PSA analysis. Firstly, EVPI was calculated to see if any further collection of perfect information on the parameters was worthwhile. Since there was some monetary benefit attached to the collection of perfect information, EVPPI was then calculated to explore which parameters affected the decision uncertainty most. A 1000 values of all the parameters were also calculated using Monte-Carlo simulation for calculation of partial EVPI or EVPPI using SAVI.

6.4.1 EVPI

There is some decision uncertainty when the probability of cost-effectiveness of no screening decreases and that of NT increases with the increase in the threshold. Figure 23 shows CEAC and individual EVPI for different values of WTP threshold, and table 13 shows the EVPI values per person, for population per year and over several years.

The value of EVPI peaks with a value of 0.8909 dollars per person at a threshold of 3000, which is the current WTP and is slightly higher than 0.868 at $2850 around where the switch from no screening to NT as the preferred strategy takes place as seen from the table too.

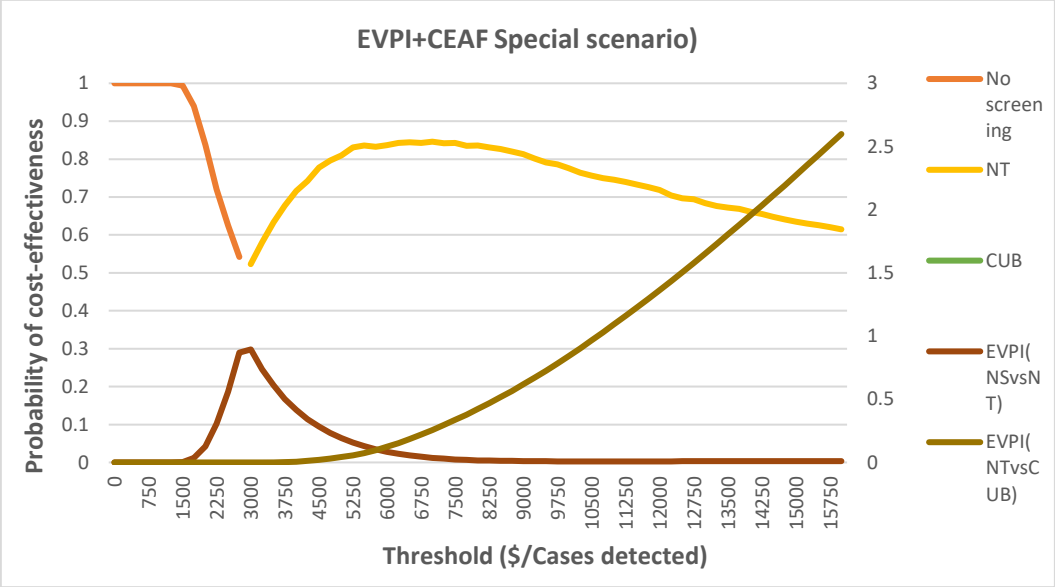


Figure 24 EVPI and CEAF (special scenario)

Table 15 Overall EVPI

	Overall EVPI ($)	Overall EVPI (No. of cases detected)
Per Person Affected by the Decision	0.8909	0.000297
Per Year in Nepal Assuming 139948 Persons Affected per Year	124700	41.56
Over 5 Years	623400	207.8
Over 10 Years	1247000	415.6

*Source-SAVI

This is because at a threshold of $3,000, NT is cost-effective only 52% of the time, and thus there is decision uncertainty as choosing NT over no screening would mean wrong decision would be made 42% of the time. This means that the collection of perfect information on all the parameters would gain a net monetary benefit of 0.8909 dollars per person at the current threshold. Assuming that the population being affected by the decision is 139,938 per year, the population EVPI is 124,700 dollars per year and 623,400 dollars over five years. This means that the cost of collection of perfect information on all the parameters, if chosen to do so, should not exceed $124,700 (per year) and $623,400 (over five years). If the cost exceeds those values, the collection of information is not worthwhile. Since there is some monetary gain, even though it is very small, attached to the collection of complete information of all the parameters, further analysis in the form of EVPPI was conducted.

6.4.2 EVPPI

Expected value of partial perfect information analysis was conducted for the current scenario using the SAVI framework. Partial EVPI was calculated for various parameters used in the model individually and in groups. Parameters were grouped into two categories according to the type, i.e., test input parameters, cost parameters, and acceptance rates. The EVPI values of all the parameters (individual and grouped) are presented in the tables below. EVPPI values of a group of parameters are more useful than individual values as we get information on a group of parameters from a study than individual parameters.

As can be seen in the tables above, the values of individual partial EVPI for most of the parameters are zero. This means that there is no gain in terms of monetary benefit associated with the collection of perfect information of these parameters listed. However, there are some values attached to eight parameters.

Table 16 EVPPI- Individual parameters

Parameters	Per Person EVPPI ($)	EVPPI for Nepal Per Year ($)	EVPPI for Nepal over 5 years ($)	Parameters	Per Person EVPPI ($)	EVPPI for Nepal Per Year ($)	EVPPI for Nepal over 5 years ($)
SensNT	0	0	0	cDT	0	0	0
FPRNT	0	0	0	cCUB	0	0	0
T_NT	0.01	1279	6393	cTT	0	0	0
PPVNT	0	0	0	cQT	0	0	0
SensDT	0	0	0	cNIPT	0	0	0
FPRDT	0	0	0	cAm	0.01	1737	8686
T_DT	0	0	0	cConsultation	0.01	874.9	4375
PPVDT	0	0	0	Incidence	0	339.5	1697
SensCUB	0	0	0	T_AmNT	0.17	24100	120500
FPRCUB	0	0	0	T_AmDT	0	0	0
T_CUB	0	173.7	868.4	T_AmCUB	0	0	0
PPVCUB	0	0	0	T_AmTT	0	0	0
SensTT	0	0.6685	3.343	T_AmQT	0	0	0
FPRTT	0	0	0	T_AmNIPT	0	0	0
T_TT	0	0	0	pFLAmn	0	0	0
PPVTT	0	0	0	pSM1	0	0	0
SensQT	0	608.8	3044	pSM2	0	0	0
FPRQT	0	0	0	pSMDS2	0.84	117900	589700
T_QT	0	0	0	aNT	0	0	0
PPVQT	0	0	0	aDT	0	0	0
SensNIPT	0	0	0	aCUB	0.01	918.7	4594
FPRNIPT	0	0	0	aTT	0	0	0
T_NIPT	0	0	0	aQT	0	0	0
PPVNIPT	0	71.56	357.8	aNIPT	0	0	0
cNT	0.02	2991	14960	aAm	0.02	2802	14010

*Source-SAVI

Among them, pSMD2 (probability of spontaneous miscarriage among DS pregnancy in the second trimester) has the highest value and contributes most to the EVPI. The proportion of test positives from NT and Amniocentesis with NT as primary test are the other three parameters with values of EVPPI more than zero. These are related to the sensitivities and FPRs of the tests and thus are the primary source of uncertainty. However, these values are considerably low and thus do not have a significant impact on the decision uncertainty. Similarly, the cost of amniocentesis, consultation, and NT have very low EVPPI value. Thus, the collection of information on pSMDS2 would mean the highest monetary gain of $551,300 over five years for the whole population, and we can invest within this amount to decrease the decision uncertainty caused by this parameter. Similarly, the maximum monetary benefit that can be gained from the collection of perfect information on the proportion of test positives from NT and subsequent amniocentesis is $120,500 for the population. There are some parameters

for which the individual EVPPI is zero, but the population EVPI is more than zero. This is because the individual values are not zero but very close to zero and thus reported as zero. However, when the number is multiplied with a population of 139,948, it has some value.

However, the results might be different when the parameters are grouped as there might be possible interaction between them, resulting in positive EVPPI values in a group, even for those parameters that had zero individual EVPPI values.

Table 17 EVPPI (group of parameters)

Parameters	Per Person EVPPI ($)	EVPPI for Nepal Per Year ($)	EVPPI for Nepal over 5 years ($)
Acceptance rates	0.197	27584.438	137922.190
Costs	0.087	12106.626	60533.130
Test input parameters	0.340	47530.970	237654.848
Probabilities- spontaneous miscarriages and fetal loss d/t amniocentesis	0.843	118020.097	590100.484

*Source- Sheffield Accelerated Value of Information

Table 17 shows the EVPPI values of group of parameters. The parameters were grouped in accordance with the type. For e.g., all the parameters like sensitivities and FPRs of the test were grouped once. All the costs associated with the test were grouped next. The probabilities of fetal loss and miscarriages were also grouped. Lastly, all the parameters were grouped to cross-verify the result from EVPI. Lastly, the acceptance rates were grouped in one.

The fourth group in the table has the highest group EVPPI value of $0.843 per person and contributes the most in the decision uncertainty, followed by the test input parameters group ($0.340), acceptance rates ($0.197), and cost group ($0.087) in descending order. Thus, the group of parameters that return the most benefit and assist in decision certainty through the collection of more information is the fetal loss and miscarriage probabilities group, followed by test input parameters, acceptance rates, and finally, cost-related data. Thus, we should start with the collection of information on fetal loss and miscarriages (followed by the group with the next highest EVPPI values) if we have enough budget and the budget expenses are such that it is below the EVPPI for five years value of the group. For instance, the maximum one should be spent in the collection of information on probabilities of spontaneous miscarriages and fetal loss for the population per year and over five years is $118020.097 and $590100.484.

6.5 Strengths and Limitations

This scenario gives us an insight that there is a possible way of implementing prenatal screening of DS among pregnant women in Nepal, which is cost-effective in the present context. This could pave the way for further research into the acceptance of such screening methods for DS at different levels of coinsurance. This scenario also makes sense in the present context as there is no health financing on the government's part for such screening services. Thus, even reimbursement of 30% of all the cost could entice women, who are at very high risk of having pregnancies with DS, to utilize the services. This also would decrease the out-of-pocket payments for the individuals, which is very high for such services. Results from VOI analysis present us with the possibility of future studies for the collection of information on various parameters to decrease decision uncertainty and gives us an insight on which parameters to focus on.

As much as this scenario helps us to assess under which conditions one of the strategies is cost-effective at the current WTP threshold level, there are few limitations to this scenario. It can be tough to convince people to spend 70%, and the payer would reimburse the remaining expenses, and also might be deemed unethical. Moreover, there would be some administration and management expenses pertaining to the insurance that could slightly increase the cost for health care providers, which has not been included in the study. The values of acceptance rates are based on the assumption of the author with opinions from experts, and thus conclusion should be made keeping that in mind.

Chapter VII: Discussion

This chapter includes the discussion and interpretation of the results of the study. The findings from relevant pieces of literature have been used to compare the result from this study. The strengths and limitations of the study are also discussed. Finally, the need for further research in the field of prenatal diagnosis of DS and such other congenital anomalies has been discussed as well.

7.1 Results of the study

There is very limited awareness and knowledge about congenital anomalies in Nepal. The disability sector in health has long been neglected by health policymakers of Nepal. It was not until 2017 that Nepal had a disability act which ensured basic health, education, and livelihood for persons with disability in Nepal (118). The census showed that there were 1.9% of the total population with some form of disability in Nepal, but there is a general consensus that those figures might be the tip of the iceberg. Similarly, with very limited resources, there is always some kind of compromise that the health sector has to bear in terms of resource allocation from the government. With issues like infectious diseases, NCDs, and maternal-child health given the most importance, very little progress has been made in the prevention or management of disability. Thus, this study has aimed to assess the cost-effectiveness of introducing prenatal diagnosis of DS (one of the most common congenital anomalies in Nepal) through the public health system and explore if such interventions are feasible in the present context.

The results from CEA showed that among all the seven strategies that were considered in the analysis, NT had the lowest cost per case detected. Several other studies support NT as the most cost-effective strategy for DS in terms of cost per case detected and cost per incremental cost (94, 104, 119). One of the important factors in this is the cost of conducting an NT test which is comparatively low compared to other tests. However, along with a lower cost, NT also meant a lower number of cases DS cases detected compared to some other strategies like CUB, QT, and NIPT. The strategy with the second-lowest-cost per case detected and ICER for all the age group was the CUB test which is one of the most commonly used primary screening tests in the world. CUB (under different names like first-trimester combined test) has been suggested as the most cost-effective strategy by several studies (98-100, 120, 121). However, there are studies that refute this claim and suggest it is time to move on to other screening strategies (102, 103). Some studies suggest different combination of first-trimester maternal serum markers and NT as the most cost-effective strategy (101). Second-trimester tests like TT and QT are shown to cost more and result in a lower number of cases detected (98, 99).

They also pose some risk as they are conducted later in the pregnancy. Some other studies show that TT and QT can be cost-effective and do not have limitations of CUB and NT (97, 104, 108). NIPT, which has come up as a new strategy with a very high detection rate and very low FPR, was found to be too costly and did not justify the high number of cases detected, which resonates with the findings of several studies (31, 108). However, NIPT, if combined as a second-tier screening test with a maternal serum marker test, can be cost-effective, as seen in this study with contingent screening almost as cost-effective as the CUB test (109, 111, 112). It can also be cost-effective from a societal perspective as denoted by one study (110).

Very few studies have conducted a cost-effective analysis with the strategies chosen for this study. There were only a few studies that included most of the interventions and calculated ICER values. One stated that NT had the lowest ICER compared to no screening, which is exactly the case in this study too. It also reported that integrated test was the next cost-effective strategy which was not included in the study. If the integrated test was removed from the analysis, CUB was the next most cost-effective, which supports the finding from this study (104). However, at a threshold of $3,000 per incremental case detected, no screening is the preferred choice for all age groups: the general population, women who are 35 years and 40 years and older. This deterministic finding is also supported by the findings from the probabilistic analysis.

As apparent from the CEAC, CEAF, and EVPI curves in the results section, there is decision uncertainty to choose one strategy over another when the willingness to pay threshold. In the current scenario where the threshold is $3,000, there is no potential gain from the collection of perfect information of the parameters, and thus a further study of any sort is not justified for the general population or when women 35 years or older or 40 years older are screened. Thus, we can be certain while making the decision that no screening should be chosen for all age groups.

However, it is very important to be cautious while comparing the findings from different studies as there are variations in the values of the parameters used. The values of sensitivities, and FPRs, etc., differ and thus can produce different results as seen in one-way and two-way sensitivity analyses carried out in the study. Different studies use different cut-off values to categorize a woman at risk, which affects the sensitivity and FPR of a test. Similarly, the cost of the individual tests also might differ in different settings. For e.g., the cost of ultrasound for NT measurement was found to be less than half in Nepal compared to the Netherlands, and the

cost of DT was found to be higher (95). Moreover, the incidence of DS in the setting under study plays a very important role. Although similar for the general population, the incidence varies by the age group of women taken as subjects for study. This is highlighted by the result from this study where the incidence of DS for women 35 years and older and 40 years and older is considerably higher, which affects the CERs of the strategies and the conclusion that one strategy is more cost-effective than the other as shown by the sensitivity analysis. The willingness to pay threshold also is an important factor while comparing as different countries have different threshold values.

While most of our focus is on the cost-effectiveness of those strategies, we also have to keep in mind the safety, harm, and benefits of the tests, for instance, in terms of the number of fetal losses in the process. NIPT is the best in this context but costs considerably higher. Thus, CUB is the best intervention in terms of the lowest fetal losses and number of amniocentesis, which has been reported by various studies.

Although the cost per case detected value of NT and CUB are similar to other studies (sometimes even lower), the willingness to pay threshold of $3,000 per incremental case detected means that no screening should be the preferred choice from a healthcare perspective as of now for the general population, and even for women who are 35 years and older or 40 years and older.

However, if a scenario is introduced where the cost of NT is divided among the health payer which is the government in this case (30%) and individual (70%), then NT can be implemented as a screening strategy for women who are 40 years and older. Probabilistic analysis also shows that NT is 52% cost-effective at the current threshold level. In such a case, there is a value attached to the collection of further information about certain parameters, which might increase the certainty of the decision indicated by positive values of EVPI and partial EVPPI. Additional collection of information of input parameters which makes the most significant impact on the decision uncertainty are group of probabilities of spontaneous miscarriages and fetal loss d/t amniocentesis followed by test input values (sensitivity and FPRs) in the population. The cost of collection of additional information should not exceed the respective population EVPPI for the group of parameters.

The total annual budget and over several years also have to be kept in mind before making any decision as the budget of screening strategies can go very high and would not make any sense if it needed a huge fraction of the total budget for health. Although tests like CUB, TT, NIPT

have higher detection rates, lower FPRs and fewer number of fetal losses than other tests, the total budget of the intervention amounts to be a huge fraction of the total health budget of the country and thus would not make sense, no matter the perspective.

7.2 Limitations of the study

There are few limitations associated with the model and the study as a whole. The main limitation of the study is that the outcome is measured in terms of the number of cases detected and not QALYs or DALYs, which would have provided a better picture of the issue. The lack of data related to the cost of living with DS and lack of enough time were the reasons for that. Similarly, the inclusion of that dimension would have meant that I would have had to go to Nepal. This was not possible due to the current situation with COVID-19.

Although the costs were obtained from the major laboratories (cross verified from different sources and expert opinion), they are mostly from private laboratories where the prices can be on the higher side. This was because several of these tests were only provided through private laboratories.

Similarly, all the pregnant women in a fiscal year are assumed for the analysis. The acceptance rate of all the tests, be it screening or diagnostic test, is assumed to 100% due to lack of data whatsoever, which is far from reality and might exaggerate the cost-effectiveness of the strategies.

Another limitation of the study is that the input parameters like sensitivity and false-positive rates are taken from secondary sources of different settings. The values might differ for different settings. There are no studies done in Nepal concerning the sensitivities and false-positive rate, and thus there was no choice.

Finally, this is a simulated study of a hypothetical cohort of pregnant women. Different women have different characteristics which might affect the risk of having a fetus with DS. This has not been taken into consideration in this study due to the model.

It has been assumed that all the women who are diagnosed having a fetus with DS are assumed to choose to terminate. There might be women who decide to keep the baby. However, the choice between termination and keeping the baby does not have an impact on either the cost or the outcome and thus does not pose a limitation of the study.

7.3 Strength of the study

Despite the limitations, this study is the first cost-effectiveness analysis of prenatal screening strategies for DS in Nepal. This study will certainly raise the significance of prenatal screening of not just DS but other congenital anomalies as well and the need for cost-effective analysis before the policymakers introduce such programs. There are various NGOs who are working in the field of disability in Nepal and can use the findings of the study to plan their actions.

The total number of women in the analysis was taken from the real distribution of pregnant women in Nepal. Similarly, the number of women who are 35 years or older and 40 years or older are derived from the real data.

The assumption that the acceptance rate of all the tests is 100% can be seen as a strength as well. The acceptance rate of 100% refers to the best-case scenario. The results thus obtained are the maximum that can be achieved for any values of acceptance. Thus, if strategies are not cost-effective at a 100% acceptance rate, then chances are they will not be for any value of acceptance.

The study will also give a useful insight to the policy makers about the cost-effectiveness aspect of different health interventions and use it in the future as no such analysis is done right now before commencing a program in healthcare.

7.4 Further research

As apparent from the various assumptions made in the study, there is plenty of room for further research not just on the cost-effectiveness of the screening strategies but on various other parameters. Firstly, there is no incidence-related data of DS and other congenital anomalies that pertain exclusively to Nepal. I was obliged to use the data from studies conducted in India. This shows a huge need for a system for recording disability-related data. Similarly, the lack of primary sources of data related to sensitivities and FPRs of the screening tests from Nepal highlights the need to conduct studies measuring the effectiveness of such tests in a low-resource setting like Nepal.

There is also a need to explore the health-seeking behavior of people when it comes to situations like choosing; to take screening tests or a particular test over others, choosing safety over cost, etc. There is also a need for an assessment of the capacity of health facilities and health workers to conduct such screening tests. For instance, certified sonographers are needed for NT measurement, and they are very limited in number, which might be a barrier in the implementation of NT. Similarly, even when a certain strategy is found to be cost-effective,

there is a need to establish a proper system of referral for the women to go and access the services.

Different congenital anomalies like DS and NTDs can be detected using the same test. A study assessing the cost-effectiveness of such tests using the number of DS and NTDs as the outcome could present a different result.

Results from VOI analysis showed that there is some monetary gain attached to the collection of additional information on certain group of parameters to decreases the decision uncertainty, albeit for a specific scenario.

Lastly, although the strategies were found to be not cost-effective from a health care perspective, results might be different if the analyzed from a societal perspective taking into account the lifetime and all the indirect cost of DS lie productivity costs, medical treatment needed for persons with DS throughout their lifetime, etc. The strategies which are not cost-effective might turn out to be cost-effective, as discussed by some literature. Thus, a study from a societal perspective will give the best picture.

Chapter VIII: Conclusion

The choice of prenatal screening strategies for Nepal at the current willingness to pay threshold is straightforward, i.e., no screening regardless of the age group. No combination of the cost of the tests can change the result of cost-effectiveness of these strategies. NT can be introduced as a screening strategy if provided to women who are 40 years and older through a provision where the cost of NT is divided between the health system (30%) and individual (70%) as it will be cost-effective at the current WTP threshold. VOI analysis shows that the collection of further information is worthwhile for some parameters for the special scenario. However, this might result in an ethical dilemma as women who are 35 years and older are also at higher risk of having a fetus with DS but cannot be provided the test. Similarly, there is no room to assess the preference of people about the safety and effectiveness of other screening strategies as the health system of Nepal in the current scenario cannot afford to finance these tests. If a pregnant woman decides to go through prenatal screening on her own without financial help from the health care system, it would be a wise decision to choose NT (if cost is the priority) or CUB (if safety is the priority), but it would mean a catastrophic health care expenditure for both individual (if paid out-of-pocket) and health care system (if financed).

References

1. Update on overall prevalence of major birth defects--Atlanta, Georgia, 1978-2005. MMWR Morb Mortal Wkly Rep. 2008;57(1):1-5.
2. World Health A. Birth defects: report by the Secretariat. Geneva: World Health Organization; 2010.
3. Organization WH. Congenital anomalies 2020 [Available from: https://www-who-int.ezproxy.uio.no/news-room/fact-sheets/detail/congenital-anomalies.
4. Suzan LC. Birth defects epidemiology. European Journal of Medical Genetics. 2014;57(8):355-8.
5. Wu JL, Chen G, Song XM, Li CF, Zhang L, Liu L, et al. Spatiotemporal property analysis of birth defects in Wuxi, China. Biomed Environ Sci. 2008;21(5):432-7.
6. Xu XY, Yang JH, Ma XM, Liu AL, Liu K, He S, et al. [Neonatal complications and birth defects in infants conceived by in vitro fertilization]. Zhongguo Dang Dai Er Ke Za Zhi. 2015;17(4):350-5.
7. Lobo IZ, K. Birth defects: causes and statistics. Nature Education. 2008;1(1):18.
8. Christianson A, Howson CP, Modell B. March of Dimes: global report on birth defects, the hidden toll of dying and disabled children. Christianson, A and Howson, CP and Modell, B (2006) March of Dimes global report on birth defects: the hidden toll of dying and disabled children Research report March of Dimes Birth Defects Foundation, White Plains, USA. 2005.
9. Harris B, Bishop K, Kemeny H, Walker JS, Rhee E, Kuller J. Risk Factors for Birth Defects. Obstetrical & Gynecological Survey. 2017;72:123-35.
10. Oragnization WH. The global burden of disease. 2004.
11. Kazemi M, Salehi M, Kheirollahi M. Down Syndrome: Current Status, Challenges and Future Perspectives. Int J Mol Cell Med. 2016;5(3):125-33.
12. Wald N, & Leck, I. (Eds.). Antenatal and Neonatal Screening: Oxford University Press; 2000.
13. Mégarbané A, Ravel A, Mircher C, Sturtz F, Grattau Y, Rethoré MO, et al. The 50th anniversary of the discovery of trisomy 21: the past, present, and future of research and treatment of Down syndrome. Genet Med. 2009;11(9):611-6.
14. Salehi A, Ashford JW, Mufson EJ. The Link between Alzheimer's Disease and Down Syndrome. A Historical Perspective. Curr Alzheimer Res. 2016;13(1):2-6.
15. Hassold T, Sherman S. Down syndrome: genetic recombination and the origin of the extra chromosome 21. Clin Genet. 2000;57(2):95-100.
16. Oliver TR, Feingold E, Yu K, Cheung V, Tinker S, Yadav-Shah M, et al. New insights into human nondisjunction of chromosome 21 in oocytes. PLoS Genet. 2008;4(3):e1000033.
17. Flores-Ramírez F, Palacios-Guerrero C, García-Delgado C, Morales-Jiménez AB, Arias-Villegas CM, Cervantes A, et al. Cytogenetic profile in 1,921 cases of trisomy 21 syndrome. Arch Med Res. 2015;46(6):484-9.
18. Malini SS, Ramachandra NB. Possible risk factors for Down syndrome and sex chromosomal aneuploidy in Mysore, South India. Indian J Hum Genet. 2007;13(3):102-8.
19. Morris JK, Mutton DE, Alberman E. Revised estimates of the maternal age specific live birth prevalence of Down's syndrome. J Med Screen. 2002;9(1):2-6.
20. Morris JK, Wald NJ, Mutton DE, Alberman E. Comparison of models of maternal age-specific risk for Down syndrome live births. Prenat Diagn. 2003;23(3):252-8.
21. Yoon PW, Freeman SB, Sherman SL, Taft LF, Gu Y, Pettay D, et al. Advanced maternal age and the risk of Down syndrome characterized by the meiotic stage of chromosomal error: a population-based study. Am J Hum Genet. 1996;58(3):628-33.

22. Modi UJ NU, Aiyer S, Bharani S, Master DC, Shah T, et al. Study of malformations and Down syndrome in India (SOMDI): Baroda region. 1998.
23. Norton ME, Baer RJ, Wapner RJ, Kuppermann M, Jelliffe-Pawlowski LL, Currier RJ. Cell-free DNA vs sequential screening for the detection of fetal chromosomal abnormalities. Am J Obstet Gynecol. 2016;214(6):727.e1-6.
24. Nakata N, Wang Y, Bhatt S. Trends in prenatal screening and diagnostic testing among women referred for advanced maternal age. Prenat Diagn. 2010;30(3):198-206.
25. anomalies Esoc. Special report: Prenatal Screening Policies in Europe. 2010.
26. O'Leary P, Maxwell S, Murch A, Hendrie D. Prenatal screening for Down syndrome in Australia: Costs and benefits of current and novel screening strategies. Australian and New Zealand Journal of Obstetrics and Gynaecology. 2013;53(5):425-33.
27. Gadsbøll K, Petersen OB, Gatinois V, Strange H, Jacobsson B, Wapner R, et al. Current use of noninvasive prenatal testing in Europe, Australia and the USA: A graphical presentation. Acta Obstetricia et Gynecologica Scandinavica. 2020;99(6):722-30.
28. Tapon D. Prenatal Testing for Down Syndrome: Comparison of Screening Practices in the UK and USA. Journal of Genetic Counseling. 2010;19(2):112-30.
29. Palomaki GE, Knight GJ, Ashwood ER, Best RG, Haddow JE. Screening for down syndrome in the United States: results of surveys in 2011 and 2012. Arch Pathol Lab Med. 2013;137(7):921-6.
30. Lin S-Y, Hsieh C-J, Chen Y-L, Shaw SWS, Lin M-W, Chen P-C, et al. The impact of Down syndrome screening on Taiwanese Down syndrome births: a nationwide retrospective study and a screening result from a single medical centre. PLoS One. 2013;8(9):e75428-e.
31. Xu Y, Wei Y, Ming J, Li N, Xu N, Pong RW, et al. Cost-Effectiveness Analysis of Non-invasive Prenatal Testing for Down Syndrome in China. International Journal of Technology Assessment in Health Care. 2019;35(3):237-42.
32. Bank TW. GDP per capita- Nepal 2019 [Available from: https://data-worldbank-org.ezproxy.uio.no/indicator/NY.GDP.PCAP.CD?locations=NP.
33. Merchant A. The Borgen project [Internet]2018. [25th March,2021]. Available from: https://borgenproject.org/ten-facts-about-poverty-in-nepal/.
34. Programme UND. HUman Development Reports: United Nations Development Programme; 2020 [Available from: http://hdr.undp.org/en/countries/profiles/NPL.
35. Ghimire IBAS. National Budget 2020/21: An Analytical Review [Internet]: Nepal Economic Forum 2020. [cited 2021]. Available from: https://nepaleconomicforum.org/neftake/national-budget-2020-21-an-analytical-review/.
36. Global Health Expenditure Database [Internet]. World Health Organization. 2018. Available from: https://apps-who-int.ezproxy.uio.no/nha/database.
37. Nepal Go. National Health Policy 2074. In: Services DoH, editor. 2019.
38. Nepal Go. National Population and Housing Census 2011. In: Central Bureau of Statistics NPCS, Government of Nepal, editor. 2011.
39. Shrestha U, Bhattacharya S, Bhatta N, Jha C. Cytogenetic analysis of children with suspected genetic disorder. Kathmandu University medical journal (KUMJ). 2009;7:40-3.
40. project Dcp. Controlling birth defects: Reducing the hidden toll

of dying and disabled children in low-income countries. 2008.
41. World Health Organization. Regional Office for South-East A. Birth defects in South-East Asia: a public health challenge: situation analysis. New Delhi: WHO Regional Office for South-East Asia; 2013 2013.
42. Organization WH. Screening the genes. 2012 2012. Contract No.: 2021.
43. Khavjou OA, Anderson WL, Honeycutt AA, Bates LG, Razzaghi H, Hollis ND, et al. National Health Care Expenditures Associated With Disability. Med Care. 2020;58(9):826-32.

44. Mitra S, Findley PA, Sambamoorthi U. Health Care Expenditures of Living With a Disability: Total Expenditures, Out-of-Pocket Expenses, and Burden, 1996 to 2004. Archives of Physical Medicine and Rehabilitation. 2009;90(9):1532-40.
45. Altman BM, Cooper PF, Cunningham PJ. The Case of Disability in the Family: Impact on Health Care Utilization and Expenditures for Nondisabled Members. The Milbank Quarterly. 1999;77(1):39-75.
46. Mitra S, Palmer M, Kim H, Mont D, Groce N. Extra costs of living with a disability: A review and agenda for research. Disability and Health Journal. 2017;10(4):475-84.
47. Bodri D, Vernaeve V, Figueras F, Vidal R, Guillén JJ, Coll O. Oocyte donation in patients with Turner's syndrome: a successful technique but with an accompanying high risk of hypertensive disorders during pregnancy. Human Reproduction. 2006;21(3):829-32.
48. Sultan C, Biason-Lauber A, Philibert P. Mayer–Rokitansky–Kuster–Hauser syndrome: Recent clinical and genetic findings. Gynecological Endocrinology. 2009;25(1):8-11.
49. Cunningham SJ. Economic evaluation of healthcare – is it important to us? British Dental Journal. 2000;188(5):250-4.
50. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes: Oxford university press; 2015.
51. McGuire A, Henderson J, Mooney G. The economics of health care: an introductory text: Routledge & Kegan Paul; 1988.
52. Robinson R. Economic evaluation and health care. What does it mean? British Medical Journal. 1993;307(6905):670-3.
53. Goodacre S, McCabe C. An introduction to economic evaluation. Emergency Medicine Journal. 2002;19(3):198.
54. Consortium YHE. Cost-Effectiveness Analysis [online] 2016 [Available from: https://yhec.co.uk/glossary/cost-effectiveness-analysis/.
55. Palmer S, Byford S, Raftery J. Economics notes: types of economic evaluation. BMJ. 1999;318(7194):1349-.
56. Consortium YHE. Cost-Utility Analysis [online] 2016 [Available from: https://yhec.co.uk/glossary/cost-utility-analysis/.
57. Consortium YHE. Cost-Benefit Analysis [online] 2016 [Available from: https://yhec.co.uk/glossary/cost-benefit-analysis/.
58. Organization WH. Making choices in health: WHO guide to cost-effectiveness analysis 2003.
59. Ryder HF, McDonough C, Tosteson ANA, Lurie JD. Decision Analysis and Cost-effectiveness Analysis. Semin Spine Surg. 2009;21(4):216-22.
60. Andrew Briggs MS, and Karl Claxton. Decision Modelling for Health Economic Evaluation2006.
61. Aalabaf-Sabaghi M. Decision modelling for health economic evaluation. J Epidemiol Community Health. 2007;61(9):839-.
62. Haeussler K, den Hout Av, Baio G. A dynamic Bayesian Markov model for health economic evaluations of interventions in infectious disease. BMC Medical Research Methodology. 2018;18(1):82.
63. Consortium YHE. Markov Model [online] 2016 [Available from: https://yhec.co.uk/glossary/markov-model/.
64. Consortium YHE. Decision Tree [online] 2016 [Available from: https://yhec.co.uk/glossary/decision-tree/

65. Rautenberg T, Gerritsen A, Downes M. Health Economic Decision Tree Models of Diagnostics for Dummies: A Pictorial Primer. Diagnostics (Basel). 2020;10(3).

66. Consortium YHE. Incremental Cost-Effectiveness Ratio (ICER) 2016 [Available from: https://yhec.co.uk/glossary/incremental-cost-effectiveness-ratio-icer/.
67. Paulden M. Calculating and Interpreting ICERs and Net Benefit. PharmacoEconomics. 2020;38(8):785-807.
68. Consortium YHE. Net Monetary Benefit [online] 2016 [Available from: https://yhec.co.uk/glossary/net-monetary-benefit/.
69. Consortium YHE. Net Health Benefit [online] 2016 [Available from: http://www.yhec.co.uk/glossary/net-health-benefit/.
70. Alastair M. Gray PMC, Jane L. Wolstenholme, and Sarah Wordsworth. Applied Methods of Cost-effectiveness Analysis in Healthcare: Oxford University Press; 2010. 328 p.
71. Consortium YHE. Cost-Effectiveness Acceptability Curve (CEAC) 2016 [Available from: https://yhec.co.uk/glossary/cost-effectiveness-acceptability-curve-ceac/.
72. Consortium YHE. Cost-Effectiveness Acceptability Frontier (CEAF) [online] 2016 [Available from: https://yhec.co.uk/glossary/cost-effectiveness-acceptability-frontier-ceaf/.
73. Bolam FC, Grainger MJ, Mengersen KL, Stewart GB, Sutherland WJ, Runge MC, et al. Using the Value of Information to improve conservation decision making. Biol Rev Camb Philos Soc. 2019;94(2):629-47.
74. York Uo. Expected Value of Perfect Information (EVPI) [Available from: https://yhec.co.uk/glossary/expected-value-of-perfect-information-evpi/.
75. York Uo. Expected Value of Partially Perfect Information (EVPPI) [Available from: https://yhec.co.uk/glossary/expected-value-of-partially-perfect-information-evppi/.
76. Wilson ECF. A Practical Guide to Value of Information Analysis. PharmacoEconomics. 2015;33(2):105-21.
77. Strong M, Oakley JE, Brennan A. Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: a nonparametric regression approach. Med Decis Making. 2014;34(3):311-26.
78. Neumann PJ. Budget impact analyses get some respect. Value Health. 2007;10(5):324-5.
79. Mauskopf JA, Sullivan SD, Annemans L, Caro J, Mullins CD, Nuijten M, et al. Principles of good practice for budget impact analysis: report of the ISPOR Task Force on good research practices--budget impact analysis. Value Health. 2007;10(5):336-47.
80. Mauskopf J. Prevalence-based economic evaluation. Value Health. 1998;1(4):251-9.
81. Xu Y, Wei Y, Ming J, Li N, Xu N, Pong RW, et al. Cost-Effectiveness Analysis of Non-invasive Prenatal Testing for Down Syndrome in China. Int J Technol Assess Health Care. 2019;35(3):237-42.
82. Wald NJ, Kennard A, Hackshaw A, McGuire A. Antenatal screening for Down's syndrome. Health Technol Assess. 1998;2(1):i-iv, 1-112.
83. Shaffer ML, Watterberg KL. Joint distribution approaches to simultaneously quantifying benefit and risk. BMC Med Res Methodol. 2006;6:48.
84. Kurtovic-Kozaric A, Mehinovic L, Malesevic R, Mesanovic S, Jaros T, Stomornjak-Vukadin M, et al. Ten-year trends in prevalence of Down syndrome in a developing country: impact of the maternal age and prenatal screening. Eur J Obstet Gynecol Reprod Biol. 2016;206:79-83.
85. Doidge JC, Morris JK, Harron KL, Stevens S, Gilbert R. Prevalence of Down's Syndrome in England, 1998–2013: Comparison of linked surveillance data and electronic health records. International Journal of Population Data Science. 2020;5(1).
86. Research ICfBDSa. International Clearinghouse for Birth Defects Surveillance and Research Annual Report 2012. Rome, Italy. 2012.

87. Lanzoni M, Kinsner-Ovaskainen, A., Morris, J. and Martin, S. EUROCAT - Surveillance of congenital anomalies in Europe: epidemiology of Down syndrome 1990 - 2014. 2019. Report No.: 978-92-76-00574-2.
88. Mai CT, Isenburg JL, Canfield MA, Meyer RE, Correa A, Alverson CJ, et al. National population-based estimates for major birth defects, 2010-2014. Birth Defects Res. 2019;111(18):1420-35.
89. Li C, Shi L, Huang J, Qian X, Chen Y. Factors associated with utilization of maternal serum screening for Down syndrome in mainland China: a cross-sectional study. BMC Health Services Research. 2016;16(1):8.
90. Coppedè F. Risk factors for Down syndrome. Archives of Toxicology. 2016;90(12):2917-29.
91. Cuckle H, Maymon R. Development of prenatal screening--A historical overview. Semin Perinatol. 2016;40(1):12-22.
92. Comas C, Torrents M, Muñoz A, Antolín E, Figueras F, Echevarría M. Measurement of nuchal translucency as a single strategy in trisomy 21 screening: should we use any other marker? Obstet Gynecol. 2002;100(4):648-54.
93. Pandya PP, Snijders RJ, Johnson SP, de Lourdes Brizot M, Nicolaids KH. Screening for fetal trisomies by maternal age and fetal nuchal translucency thickness at 10 to 14 weeks of gestation. BJOG: An International Journal of Obstetrics & Gynaecology. 1995;102(12):957-62.
94. Michailidis GD, Spencer K, Economides DL. The use of nuchal translucency measurement and second trimester biochemical markers in screening for Down's Syndrome. BJOG: An International Journal of Obstetrics & Gynaecology. 2001;108(10):1047-52.
95. Hoogendoorn M, Evers SM, Schielen PC, van Genugten ML, de Wit GA, Ament AJ. Costs and effects of prenatal screening methods for Down syndrome and neural tube defects. Community Genet. 2008;11(6):359-67.
96. Hwa HL, Yen MF, Lin CL, Ko TM, Hsieh FJ, Chen TH. Cost-effectiveness analysis of triple test in second-trimester maternal serum screening for Down's syndrome: an experience from Taiwan with decreasing birth rate but increasing population of old pregnant women. J Eval Clin Pract. 2008;14(2):191-7.
97. Hwa H-L, Yen M-F, Lin C-L, Ko T, Hsieh F-J, Chen T. Cost-effectiveness analysis of triple test in second-trimester maternal serum screening for Down's syndrome: An experience from Taiwan with decreasing birth rate but increasing population of old pregnant women. Journal of evaluation in clinical practice. 2008;14:191-7.
98. Cusick W, Buchanan P, Hallahan T, Krantz D, Larsen J, Macri J. Combined first-trimester versus second-trimester serum screening for Down syndrome: A cost analysis. American journal of obstetrics and gynecology. 2003;188:745-51.
99. Ökem ZG, Örgül G, Kasnakoglu BT, Çakar M, Beksaç MS. Economic analysis of prenatal screening strategies for Down syndrome in singleton pregnancies in Turkey. European Journal of Obstetrics and Gynecology and Reproductive Biology. 2017;219:40-4.
100. Li B, Sahota D, Lao T, Xu J, Hu S, Zhang L, et al. Applicability of first-trimester combined screening for fetal trisomy 21 in a resource-limited setting in mainland China. BJOG: An International Journal of Obstetrics & Gynaecology. 2016;123(S3):23-9.
101. Christiansen M, Olesen Larsen S. An increase in cost-effectiveness of first trimester maternal screening programmes for fetal chromosome anomalies is obtained by contingent testing. Prenat Diagn. 2002;22(6):482-6.
102. Vintzileos AM, Ananth CV, Smulian JC, Day-Salvatore DL, Beazoglou T, Knuppel RA. Cost-benefit analysis of prenatal diagnosis for Down syndrome using the British or the American approach. Obstet Gynecol. 2000;95(4):577-83.

103. Gekas J, Gagné G, Bujold E, Douillard D, Forest JC, Reinharz D, et al. Comparison of different strategies in prenatal screening for Down's syndrome: cost effectiveness analysis of computer simulation. BMJ. 2009;338:b138.
104. Gilbert RE, Augood C, Gupta R, Ades AE, Logan S, Sculpher M, et al. Screening for Down's syndrome: effects, safety, and cost effectiveness of first and second trimester strategies. BMJ. 2001;323(7310):423-5.
105. Norton ME, Jacobsson B, Swamy GK, Laurent LC, Ranzini AC, Brar H, et al. Cell-free DNA analysis for noninvasive examination of trisomy. N Engl J Med. 2015;372(17):1589-97.
106. Mackie FL, Hemming K, Allen S, Morris RK, Kilby MD. The accuracy of cell-free fetal DNA-based non-invasive prenatal testing in singleton pregnancies: a systematic review and bivariate meta-analysis. Bjog. 2017;124(1):32-46.
107. Gil MM, Quezada MS, Revello R, Akolekar R, Nicolaides KH. Analysis of cell-free DNA in maternal blood in screening for fetal aneuploidies: updated meta-analysis. Ultrasound Obstet Gynecol. 2015;45(3):249-66.
108. Okun N, Teitelbaum M, Huang T, Dewa CS, Hoch JS. The price of performance: a cost and performance analysis of the implementation of cell-free fetal DNA testing for Down syndrome in Ontario, Canada. Prenatal Diagnosis. 2014;34(4):350-6.
109. Evans MI, Sonek JD, Hallahan TW, Krantz DA. Cell-free fetal DNA screening in the USA: a cost analysis of screening strategies. Ultrasound in Obstetrics & Gynecology. 2015;45(1):74-83.
110. Walker BS, Nelson RE, Jackson BR, Grenache DG, Ashwood ER, Schmidt RL. A Cost-Effectiveness Analysis of First Trimester Non-Invasive Prenatal Screening for Fetal Trisomies in the United States. PLoS One. 2015;10(7):e0131402.
111. Ayres AC, Whitty JA, Ellwood DA. A cost-effectiveness analysis comparing different strategies to implement noninvasive prenatal testing into a Down syndrome screening program. Australian and New Zealand Journal of Obstetrics and Gynaecology. 2014;54(5):412-7.
112. Neyt M, Hulstaert F, Gyselaers W. Introducing the non-invasive prenatal test for trisomy 21 in Belgium: a cost-consequences analysis. BMJ Open. 2014;4(11):e005922-e.
113. Rautenberg T, Gerritsen A, Downes M. Health Economic Decision Tree Models of Diagnostics for Dummies: A Pictorial Primer. Diagnostics. 2020;10(3):158.
114. Nepal Go. The Right to Safe Motherhood and Reproductive

Health Act, 2075 (2018). In: Comission NL, editor.: Nepal Law Comission; 2018. p. 13.
115. Hutubessy R, Chisholm D, Edejer TT. Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. Cost Eff Resour Alloc. 2003;1(1):8.
116. Leech AA, Kim DD, Cohen JT, Neumann PJ. Use and Misuse of Cost-Effectiveness Analysis Thresholds in Low- and Middle-Income Countries: Trends in Cost-per-DALY Studies. Value Health. 2018;21(7):759-61.
117. Torgerson DJ, Spencer A. Marginal costs and benefits. BMJ. 1996;312(7022):35-6.
118. Nepal Go. The Act Relating to Rights of Persons with Disabilities, 2074 (2017). Nepal Law Commission; 2017. p. 34.
119. Caughey AB, Kuppermann M, Norton ME, Washington AE. Nuchal translucency and first trimester biochemical markers for down syndrome screening: a cost-effectiveness analysis. Am J Obstet Gynecol. 2002;187(5):1239-45.
120. Kott B, Dubinsky TJ. Cost-effectiveness model for first-trimester versus second-trimester ultrasound screening for Down syndrome. J Am Coll Radiol. 2004;1(6):415-21.
121. Shengmou L, Min C, Chenhong W, Shengli L, Jiansheng X, Hui Y, et al. [Effects, safety and cost-benefit analysis of Down syndrome screening in first trimester]. Zhonghua Fu Chan Ke Za Zhi. 2014;49(5):325-30.

Appendix

Parameters (35 years or above)				
Parameters	**Value**	**Distribution**	**Alpha**	**Beta**
Irg (in %)	0.32	Beta	67.68	143.82
Total population	231308			
Test input values				
SensNT	0.700	Beta	234.5	100.5
FPRNT	0.024	Beta	97.576	3968.091
T+NT	0.026	Beta	97.35752	3623.805
SensDT	0.620	Beta	64.12105	39.3
FPRDT	0.050	Beta	94.95	1804.05
T+DT	0.052	Beta	94.76578	1733.842
SensCUB	0.818	Beta	194.0309	43.17069
FPRCUB	0.021	Beta	97.879	4563.026
T+CUB	0.024	Beta	97.62141	4047.591
SensTT	0.651	Beta	34.249	18.36083
FPRTT	0.047	Beta	95.253	1931.407
T+TT	0.049	Beta	95.05779	1847.561
SensQT	0.760	Beta	23.24	7.338947
FPRQT	0.090	Beta	90.91	919.2011
T+QT	0.092	Beta	90.69346	893.5644
SensNIPT	0.990	Beta	54.83627	0.553902
FPRNIPT	0.001	Beta	3.995	3991.005
T+NIPT	0.004	Beta	99.57936	23810.18
SensAmn	0.993	Beta	38.32259	0.270149
FPRAmn	0.001	Beta	99.8586	71227.71
PPVNT	0.086	Beta	91.35274	975.6473
PPVDT	0.038	Beta	96.13337	2414.963
PPVCUB	0.111	Beta	88.77397	709.9204
PPVTT	0.043	Beta	95.70016	2152.225
PPVQT	0.026	Beta	97.33426	3590.482
PPVNIPT	0.761	Beta	23.17326	7.291385
T+AmNT	0.086	Beta	91.28398	966.5014
T+AmDT	0.039	Beta	96.02445	2343.507
T+AmCUB	0.112	Beta	88.72686	706.2084
T+AmTT	0.044	Beta	95.59488	2096.191
T+AmQT	0.028	Beta	14602.29	1.84E-06
T+AmNIPT	0.756	Beta	274.8419	0.00067
T+CUBNIPT	0.111	Beta	88.79645	711.7037
T+CUBNIPTAmn	0.985	Beta	0.509514	0.007731

Parameters (40 years and above)				
Parameters	**Value**	**Distribution**	**Alpha**	**Beta**
Irg (in %)	1.5	Beta	67.6800	143.8200
Total population	139948			
Test input parameters				
T+NT	0.0341	Beta	96.5519	2731.5635
SensDT	0.6200	Beta	64.1211	39.3000
FPRDT	0.0500	Beta	94.9500	1804.0500
T+DT	0.0586	Beta	94.0865	1512.8555
SensCUB	0.8180	Beta	194.0309	43.1707
FPRCUB	0.0210	Beta	97.8790	4563.0258
T+CUB	0.0330	Beta	96.6715	2836.7694
SensTT	0.6510	Beta	34.2490	18.3608
FPRTT	0.0470	Beta	95.2530	1931.4066
T+TT	0.0561	Beta	94.3379	1588.4651
SensQT	0.7600	Beta	23.2400	7.3389
FPRQT	0.0900	Beta	90.9100	919.2011
T+QT	0.1001	Beta	89.8950	808.6053
SensNIPT	0.9900	Beta	54.8363	0.5539
FPRNIPT	0.0010	Beta	3.9950	3991.0050
T+NIPT	0.0158	Beta	98.4007	6115.7241
SensAmn	0.9930	Beta	38.3226	0.2701
FPRAmn	0.0014	Beta	99.8586	71227.7128
PPVNT	0.3076	Beta	68.9367	155.2061
PPVDT	0.1588	Beta	83.9573	444.6126
PPVCUB	0.3723	Beta	62.3951	105.1868
PPVTT	0.1742	Beta	82.4070	390.6842
PPVQT	0.1139	Beta	88.4918	688.1398
PPVNIPT	0.9378	Beta	5.2826	0.3504
T+AmNT	0.3064	Beta	69.0563	156.3426
T+AmDT	0.1589	Beta	83.9507	444.3587
T+AmCUB	0.3706	Beta	62.5696	106.2643
T+AmTT	0.1741	Beta	82.4134	390.8861
T+AmQT	0.1144	Beta	228888.6757	0.0000
T+AmNIPT	0.9313	Beta	116.7401	0.0005
T+CUBNIPT	0.3692	Beta	62.7077	107.1259
T+CUBNIPTAmn	0.9913	Beta	-	-