

Generating Evidence to Streamline the Clinical Pathway in Autism Spectrum Disorder Using Simulation Models: Cost-effectiveness Comparisons of Screening and Genetic Testing Strategies

by

Tracy Yuen

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Institute of Health Policy, Management and Evaluation
University of Toronto

© Copyright by Tracy Yuen 2017

Generating Evidence to Streamline the Clinical Pathway in Autism Spectrum Disorder Using Simulation Models: Cost-effectiveness Comparisons of Screening and Genetic Testing Strategies

Tracy Yuen

Doctor of Philosophy

Institute of Health Policy, Management and Evaluation
University of Toronto

2017

Abstract

The clinical pathway to diagnosis in autism spectrum disorder (ASD) is complex. The objective of this thesis was to estimate the health and monetary impact of changes in the provision of select ASD screening and diagnostic health services. First, a meta-analysis estimated that the Modified Checklist for Autism in Toddlers, a commonly used ASD screening tool, performed at a low-to-moderate accuracy among children with developmental delay (pooled sensitivity: 0.83, 95% credible interval [CrI] 0.75, 0.90; specificity: 0.51, 95% CrI: 0.41, 0.61; positive predictive values [PPV] in high-risk children: 0.55, 95% CrI: 0.45, 0.66; PPV in low-risk children: 0.07, 95% CrI: <0.01, 0.16) and its performance changed with patient characteristics. The second study evaluated the cost-effectiveness of universal or high-risk screening compared to standard care, surveillance monitoring, in ASD using discrete event simulation. Results demonstrated that universal screening would greatly burden the healthcare system by heightening demand for diagnostic services and increasing healthcare expenditure. High-risk screening, on the other

hand, could be a cost-effective strategy, yielding incremental cost-effectiveness ratios (ICERs) of \$1100-1900/child initiated treatment or diagnosed earlier. The last study compared the cost-effectiveness of genome (GS) or exome sequencing (ES) to chromosomal microarray (CMA) using a microsimulation model. The use of ES in children with syndromic features after a negative CMA could be cost-effective compared to CMA alone (ICERs \$5800-6000/child with pathogenic variant). If CMA was to be replaced by sequencing, GS would be the more cost-effective option. Findings from this thesis indicate that strategic resource allocation is crucial in ASD. Given the network of health, psychosocial and educational services required by individuals with ASD, changes in one component can have a large impact on the wait time, resource use and expenditure on downstream services that can affect children without ASD and extend beyond the healthcare system.

ACKNOWLEDGEMENTS

Sincere gratitude is due to my PhD supervisor, Dr. Wendy Ungar, and my committee members, Drs. Melissa Carter and Peter Szatmari. Not only for their invaluable support and guidance throughout the research process, but also for putting up with my flood of manuscript drafts in a short span of time. I would also like to thank all the clinicians, researchers and laboratory staff who shared their data and expertise which made the simulation models reflect clinical reality. Finally, I would like to acknowledge financial support from the Canadian Institutes of Health Research Autism Research Training Program, Doctoral Autism Scholars Award, Ontario Graduate Scholarship and RestraComp Hospital for Sick Children Foundation Scholarship.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	IV
TABLE OF CONTENTS	V
LIST OF TABLES	VIII
LIST OF FIGURES	IX
LIST OF APPENDICES	X
ABBREVIATIONS	XI
1. INTRODUCTION AND FRAMEWORK	1
1.1 Autism Spectrum Disorder	4
1.1.1 Delay in Diagnosis	5
1.1.2 Cost of ASD	6
1.2 Screening and Clinical Assessment	7
1.2.1 Surveillance	8
1.2.1.1 Nipissing District Developmental Screen	8
1.2.2 Screening	9
1.2.2.1 Modified Checklist for Autism in Toddlers	10
1.2.3 Diagnostic Assessment	11
1.2.3.1 Autism Diagnostic Observation Schedule	11
1.3 Clinical Genetic Testing	12
1.3.1 Fragile X Testing	16
1.3.2 Fluorescent <i>in Situ</i> Hybridization	16
1.3.3 Chromosomal Microarray	16
1.3.4 Clinical Genome and Exome Sequencing	17
1.3.5 Economic Evaluation of Genetic Testing	19
1.4 Study Rationale and Thesis Structure	20
2. META-ANALYSIS OF THE MODIFIED CHECKLIST FOR AUTISM IN TODDLERS	23
2.1 Preface	23
2.2 Modelling Accuracy Measures	23
2.3 Manuscript #1	24
2.3.1 Abstract	24
2.3.2 Introduction	26
2.3.3 Methods	27
2.3.3.1 Search Strategy	27
2.3.3.2 Selection Criteria	28
2.3.3.3 Data Extraction and Quality Appraisal	28
2.3.3.4 Statistical Analysis	29
2.3.3.5 Investigation of Heterogeneity	30
2.3.4 Results	30
2.3.4.1 Study Characteristics	30
2.3.4.2 Studies on Low-risk Children	31
2.3.4.3 Studies on High-risk Children	31
2.3.4.4 Sensitivity and Specificity	32

2.3.4.5 Positive Predictive Value	32
2.3.5 Discussion	33
2.3.5.1 Limitations	35
2.3.5.2 Conclusion	36
3. COST-EFFECTIVENESS OF UNIVERSAL OR HIGH-RISK SCREENING COMPARED TO SURVEILLANCE IN AUTISM SPECTRUM DISORDER	44
3.1 Preface	44
3.2. Manuscript #2	44
3.2.1 Abstract	44
3.2.2 Introduction	45
3.2.3 Methods	47
3.2.3.1 Study Design	47
3.2.3.2 Screening Strategies	47
3.2.3.3 Outcomes	48
3.2.3.4 Costing	49
3.2.3.5 Model Pathway	49
3.2.3.6 Statistical Analysis	50
3.2.4 Results	51
3.2.4.1 Cost-effectiveness Acceptability Curve	52
3.2.4.2 Sensitivity Analyses	53
3.2.5 Discussion	55
3.2.5.1 Limitations	57
3.2.6 Conclusion	58
4. COST-EFFECTIVENESS OF GENOME AND EXOME SEQUENCING IN CHILDREN WITH AUTISM SPECTRUM DISORDER	75
4.1 Preface	75
4.2 Manuscript #3	75
4.2.1 Abstract	75
4.2.2 Introduction	76
4.2.3 Methods	78
4.2.3.1 Genetic Testing Strategies	78
4.2.3.2 Outcome	78
4.2.3.3 Costing	79
4.2.3.4 Model Pathway	79
4.2.3.5 Statistical Analysis	80
4.2.4 Results	81
4.2.4.1 Alternate Parameter Distributions	81
4.2.4.2 One-way Sensitivity Analysis	82
4.2.5 Discussion	82
4.2.5.1 Limitations	84
4.2.5.2 Conclusion	85
5. DISCUSSION	99
5.1 Summary of Main Findings	99
5.2 Implications for ASD Clinical Care	101
5.3 Implications for the Healthcare System	103

5.4 Future Research	104
5.5 Conclusion	105
BIBLIOGRAPHY	107

LIST OF TABLES

Table 1.1 The prevalence and phenotypes of select genetic disorder and mutations associated with autism spectrum disorder. _____	14
Table 1.2 Recommendation for clinical genetic testing in autism spectrum disorder. _____	15
Table 2.1 Characteristics of the 13 studies included in the meta-analysis. _____	37
Table 2.2 M-CHAT results of the 13 studies included in the meta-analysis. _____	38
Table 2.3 Quality assessment of studies included in the meta-analysis. _____	39
Table 3.1 Static and time-varying attributes assigned to children in discrete event simulation model. _____	59
Table 3.2 Input parameter for accuracy of ASD diagnostic assessment and M-CHAT screening in discrete event simulation model. _____	60
Table 3.3 Cost items and resource use for discrete event simulation model. _____	61
Table 3.4 Input parameters used to estimate wait times for diagnostic assessment and ASD intervention in discrete event simulation model. _____	63
Table 3.5 Parameters and ranges included in the one-way deterministic sensitivity analysis. ____	64
Table 3.6 Mean and incremental outcomes and costs in the public payer perspective. _____	65
Table 3.7 Mean and incremental outcomes and costs in the societal perspective. _____	66
Table 3.8 Incremental cost-effectiveness ratios from one-way deterministic sensitivity analyses in the public payer perspective using one clinician type for all diagnostic assessment. ____	67
Table 3.9 Incremental cost-effectiveness ratios from one-way deterministic sensitivity analyses in the societal perspective using one clinician type for all diagnostic assessment. _____	68
Table 4.1 Mean and incremental costs and outcomes by test strategy. _____	86
Table 4.2 Incremental cost-effectiveness ratios from sensitivity analyses using alternate distributions. _____	87

LIST OF FIGURES

Figure 1.1 Diagram of theoretical framework. _____	4
Figure 2.1 PRISMA flow diagram of literature search. _____	40
Figure 2.2 Forest plots of sensitivity and specificity. _____	41
Figure 3.1 Schematic diagram of the clinical pathway in the discrete event simulation model. _____	69
Figure 3.2 Cost-effectiveness frontiers for the three ASD screening strategies in the discrete event simulation model in the public payer (top row) and societal (bottom row) perspectives. _____	70
Figure 3.3 Incremental costs and effects from the bootstrap simulation for high-risk screening and universal screening compared to surveillance in public payer (top row) and societal (bottom row) perspectives. _____	71
Figure 3.4 Cost-effectiveness acceptability curve for high-risk screening (top) and universal screening (bottom) compared to surveillance. _____	72
Figure 3.5 One-way deterministic sensitivity analysis for high-risk screening (left column) and universal screening (right column) compared to surveillance in the <i>public payer</i> perspective. Top row shows ICER per additional child diagnosed before 36 months, bottom row shows ICER per additional child initiated treatment before 48 months. _____	73
Figure 3.6 One-way deterministic sensitivity analysis for high-risk screening (left column) and universal screening (right column) compared to surveillance in the <i>societal</i> perspective. Top row shows ICER per additional child diagnosed before 36 months, bottom row shows ICER per additional child initiated treatment before 48 months. _____	74
Figure 4.1 Schematic diagram of the microsimulation model. _____	88
Figure 4.2 Cost-effectiveness acceptability curve of the three comparison strategies using chromosomal microarray as reference in the societal perspective. _____	89
Figure 4.3 One-way sensitivity analysis under the societal perspective. _____	90

LIST OF APPENDICES

Appendix 2.1 Questions on the QUADAS-2 modified for this study. _____	42
Appendix 2.2 Results of the meta-regression models. _____	43
Appendix 4.1 Input parameters used in the microsimulation model. _____	91
Appendix 4.2 Unit price and resource use of cost items in the microsimulation model. _____	93
Appendix 4.3 Values used in one-way sensitivity analysis and alternate parameter distributions for microsimulation model. _____	97
Appendix 4.4 Incremental costs and effects, from bootstrap simulation, of each alternate genetic testing strategy compared chromosomal microarray only in the societal perspective. _____	98

ABBREVIATIONS

ABA	Applied Behavior Analysis
AAP	American Academy of Pediatrics
ACMG	American College of Medical Genetics and Genomics
ADOS	Autism Diagnostic Observation Schedule
AOSI	Autism Observation Scale for Infant
ASD	Autism spectrum disorder
CCMG	Canadian College of Medical Geneticists
CEA	Cost-effectiveness analysis
CGES	Clinical genome and exome sequencing
CI	Confidence interval
CMA	Chromosomal microarray
CNV	Copy number variations
CrI	Credible interval
DD	Developmental disability
DES	Discrete event simulation
DSM	Diagnostic and Statistical Manual of Mental Disorder
EIBI	Early intensive behavioural intervention
ES	Exome sequencing
FISH	Fluorescent <i>in situ</i> hybridization
GS	Genome sequencing
ICD	International Classification of Diseases
ID	Intellectual disability
M-CHAT	Modified Checklist for Autism in Toddlers
NDDS	Nipissing District Developmental Screen
OHTAC	Ontario Health Technology Assessment Committee
PDD-NOS	Pervasive developmental disorder- not otherwise specified
PPV	Positive predictive value
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
SickKids	Hospital for Sick Children

1. INTRODUCTION AND FRAMEWORK

The presentation of autism spectrum disorder (ASD), with core symptoms of impaired social communication, repetitive behaviour and restricted interests, differs in pattern and severity across individuals (Lai, Lombardo, & Baron-Cohen, 2014). Individuals with ASD often have co-occurring medical conditions which requires additional clinical investigations in areas such as genetics, neuropsychiatry, endocrinology and gastroenterology (Lai et al., 2014). This is reflected in recent best practice guidelines in ASD and the list of recommended clinical investigations will likely increase as evidence on the etiology and clinical progression of ASD emerges (Anagnostou et al., 2014; Filipek et al., 2000; Johnson, Myers, & the Council on Children With Disabilities, 2007; Nachshen, Garcin, Moxness, Tremblay, Hutchinson, et al., 2008; Volkmar et al., 2014). Given limited resources and current long wait times for ASD services in Ontario (Auditor General of Ontario, 2015), evidence on how to streamline the clinical pathway in ASD through efficiency improvements is critical to ensure children are receiving the services they need in a timely manner.

One widely debated area is the best approach to identify children with ASD (Al-Qabandi, Gorter, & Rosenbaum, 2011; G. Dawson, 2016; Fein, 2016; Mandell & Mandy, 2015; Pierce, Courchesne, & Bacon, 2016; Powell, 2016; Robins et al., 2016; Silverstein & Radesky, 2016; Veenstra-VanderWeele & McGuire, 2016). The current guideline from the American Academy of Pediatrics (AAP) (Johnson et al., 2007) recommends universal screening (i.e. all children undergo screening at 18 and 24 months using a standardized ASD screening tool). But not all health professional associations or government bodies, such as the US Preventive Service Task Force or provincial health ministries across Canada (Nachshen, Garcin, Moxness, Tremblay, Hutchinson, et al., 2008; Siu et al., 2016), endorse universal screening. In Ontario, the use of a broad screening tool (i.e. Nipissing District Developmental Screen or Baby Rourke Records) is recommended at 18 months for all children and only those considered to be at high-risk undergo ASD screening (Nachshen, Garcin, Moxness, Tremblay, Hutchinson, et al., 2008; Williams, Clinton, & Canadian Pediatric Society, 2011). As initial symptoms of ASD can emerge in the first year of life, structured ASD screening could potentially identify these early risk markers that are unrecognized otherwise (Ozonoff, Heung, Byrd, Hansen, & Hertz-Picciotto, 2008;

Zwaigenbaum, Bauman, Fein, et al., 2015). In turn, universal screening is hypothesized to result in earlier diagnosis and treatment initiation, despite a lack of supporting evidence (Siu et al., 2016). Delay in ASD diagnosis could be attributed to health system inefficiencies rather than timing or method of ASD screening. Studies which demonstrated more children with ASD were detected using standardized tools compared to clinical observation have failed to illustrate how or if this translates to earlier age at diagnosis or treatment initiation (Al-Qabandi et al., 2011; Duby & Johnson, 2009).

Determining which of the recommended ASD screening tools should be used is a critical component of the potential effectiveness of universal screening. The Modified Checklist for Autism in Toddlers (M-CHAT) is one of the most cited and studied ASD screening tools that performs at moderate sensitivity and specificity in high-risk children, but its accuracy in the general population is less conclusive (Robins, Fein, Barton, & Green, 2001; Sunita & Bilszta, 2013; Yama, Freeman, Graves, Yuan, & Karen Campbell, 2012). If used on a population level, as in universal screening, low specificity and low ASD prevalence could lead to high proportion of children without ASD being referred for ASD diagnostic assessment. Given the long wait time for diagnostic assessment in Ontario (median 6 months; range 1 to 24 months) (Penner, 2016), the additional children from false positive ASD screening could further prolong wait time for children who require in-depth evaluation. As ASD diagnostic assessment is a lengthy process that could involve more than one clinician, increased demand could also lead to heightened healthcare expenditure and lost productivity.

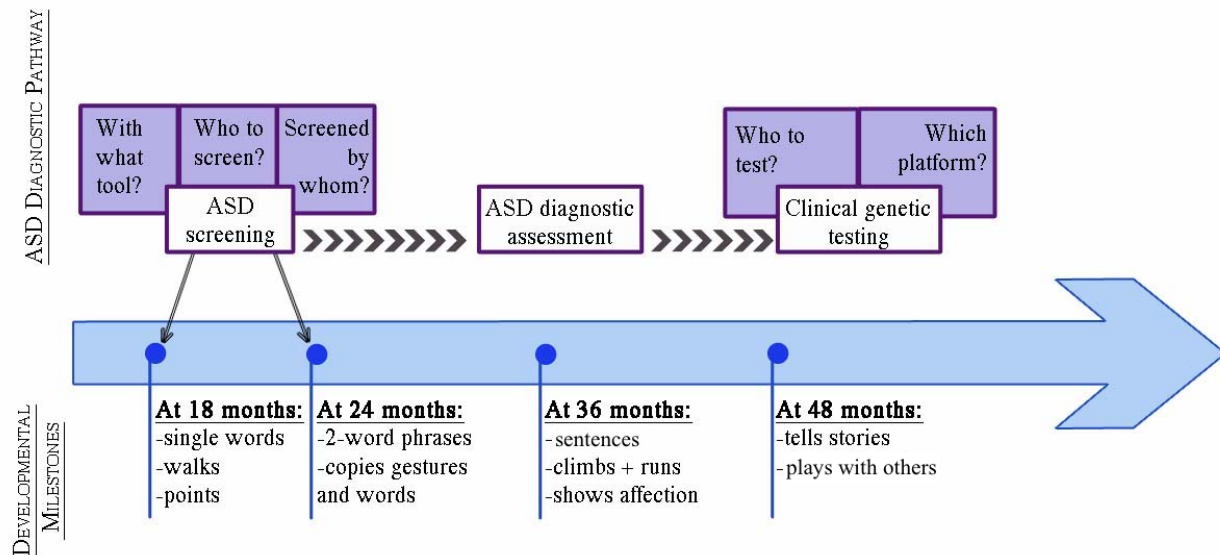
Similar concerns regarding uncertain health benefits and increase in healthcare expenditure also cloud the decision regarding whether newer genetic testing platforms should be used as part of ASD diagnostic assessment. Genetic tests that detect chromosomal abnormalities, such as karyotype or chromosomal microarray (CMA), are used after or during ASD diagnosis in order to delineate etiology, inform family planning and monitor for comorbid medical conditions (Anagnostou et al., 2014; Carter & Scherer, 2013; Schaefer & Mendelsohn, 2013). Some children are also tested earlier in the diagnostic process, before ASD is suspected, if they present with global developmental delays. Newer genetic tests which use next-generation sequencing technology, such as genome sequencing (GS) and exome sequencing (ES), can identify

additional rare variants in children with ASD. One of the drawbacks for the latter approaches is the high cost associated with DNA sequencing, bioinformatic analysis, clinical interpretation, and subsequent clinical assessments. Also, whether test results can inform treatment or diagnosis is not yet well understood since the evidence on the identified variants is still emerging. Currently, the use of clinical genome and exome sequencing (CGES) is reserved for children with ASD and syndromic features (i.e. children with clinical features in addition to those typically associated with idiopathic ASD) either as first-line testing or after a negative CMA. The use of CGES as first-line in these individuals could potentially eliminate genetic tests with lower resolution and eventually lead to more personalized intervention in the future when intervention based on genetic profile becomes available. Reduction of unnecessary testing could lead to lower healthcare costs and faster “diagnostic closure” for families.

The theoretical framework for this thesis is summarized in Figure 1.1. The overall goal is to generate evidence to help streamline the screening and diagnostic process in ASD through economic evaluations. The first objective is to summarize the sensitivity, specificity and positive predictive value (PPV) of the M-CHAT using a meta-analysis. Using this information, the second objective is to quantify the incremental costs and benefits of universal or high-risk screening compared to surveillance monitoring in ASD. The third objective is to determine which genetic test should be integrated into ASD care by comparing the cost-effectiveness of ES or GS to CMA in children recently diagnosed with ASD.

Information generated from this thesis can be used to inform decision makers and clinicians on efficiency of screening strategies and how to implement newer genetic testing platforms in clinical settings. More accurate recommendations can help reduce unnecessary clinical assessments, which can translate to reduced healthcare cost, health services wait time and psychological burden on families. Most importantly, a more streamlined clinical pathway in ASD could ensure children are receiving the care they need in a timely manner, which is critical for improving their functional outcome in the long run.

Figure 1.1 Diagram of theoretical framework.



1.1 Autism Spectrum Disorder

Autism spectrum disorder is a neurodevelopmental condition characterized by restrictive and repetitive behaviour or interests, impaired social interaction and delayed or atypical language development (Levy, Mandell, & Schultz, 2009; Newschaffer et al., 2007). Manifestations of the core symptoms vary across individuals in terms of severity and pattern, partly in relation to age and developmental attainment. While previous diagnostic criteria distinguish between autism, Asperger syndrome and pervasive developmental disorder- not otherwise specified (PDD-NOS), these subcategories were unified as ASD in the most recent version of the Diagnostic and Statistical Manual of Mental Disorder (DSM)-5 (American Psychiatric Association, 2013).

Autistic symptoms become more manageable with age in some cases, but most individuals require lifelong medical care and psychosocial support (Myers, Johnson, & the Council on Children with Disabilities, 2007). Given the multisystem nature of ASD symptoms and comorbid conditions, ASD treatment often involves a network of healthcare professionals. Interventions, such as early intensive behavioural intervention (EIBI), have been reported to be effective in improving intellectual abilities and adaptive skills in some individuals, such that developmental pathways revert to a more age-appropriate trajectory (Myers et al., 2007).

1.1.1 Delay in Diagnosis

Retrospective studies have reported that signs of ASD were evident in the early years among children later diagnosed with ASD and that their parents noted developmental concerns at around 18 months (Bryson, Rogers, & Fombonne, 2003). However, the median age at diagnosis ranged from 39.0 to 55.0 months across Canada between 2003 and 2010 (Ouellette-Kuntz et al., 2009). A Swedish study reported a gap of 20 to 60 months between first suspicion and formal diagnosis, and another study reported an average delay of 13 months between first clinical evaluation and formal ASD diagnosis (Sivberg, 2003; Wiggins, Baio, & Rice, 2006).

Children with ASD who receive intervention at a young age are reported to have improved intellectual quotient and adaptive behaviour compare to those who do not (Eldevik et al., 2009; Reichow & Wolery, 2009; Warren et al., 2011). Therefore, early diagnosis is one of the critical components to ensure children receive treatment soon after emergence of symptoms in order to potentially mitigate the impact of ASD. While one study (Horlin, Falkmer, Parsons, Albrecht, & Falkmer, 2014) did not find definitive evidence that increased medical cost was associated with delayed diagnosis, the findings did suggest that higher cost could be mediated by increased ASD symptoms in those diagnosed later on. Using a multivariate model to adjust for demographic characteristics, expenditure increased with additional ASD symptoms, AUD \$1400 (95% confidence interval [CI]: 860, 1900), but the difference between children with delayed and immediate diagnosis failed to reach statistical significance, AUD -\$1500 (95% CI: -6000, 3100).

Indicators of later age at diagnosis include being female, increased maternal age, comorbid medical conditions, being foreign-born and living in rural areas (Frenette et al., 2013; Valicenti-McDermott, Hottinger, Seijo, & Shulman, 2012). Conversely, greater symptom severity and higher socioeconomic status are associated with earlier diagnosis (Mandell, Novak, & Zubritsky, 2005). This could be partly attributed to past reports of lower compliance with the well-child visit schedule in parents with low socioeconomic status, poor physical and/or mental health, or no health insurance coverage (Chi et al., 2013; Jhanjee, Saxeena, Arora, & Gjerdingen, 2004). Parents have also reported inability to access appropriate healthcare services due to long wait time for ASD assessment and diagnosis (Auditor General of Ontario, 2013). Medical

professionals could also be hesitant to give a formal diagnosis if early clinical presentation is complex and many clinicians do not adhere to screening or diagnostic recommendations, both of which could further delay the diagnostic process (Daniels & Mandell, 2013; Zuckerman, Lindly, & Sinche, 2015).

1.1.2 Cost of ASD

The estimated cost of caring for a child with ASD varies across reports, depending on the age range of children, source of information and timing of the study. Estimates of annual cost per child based on private insurance claims were on the lower end (USD \$2575 (Croen, Najjar, Ray, Lotspeich, & Bernal, 2006)) and estimates from ASD-specific surveys were on the higher end (AUD \$34900 (Horlin et al., 2014) to €51877 (K. Järbrink, 2007)). Estimates based on Medicaid data (USD \$ 7198 (Peacock, Amendah, Ouyang, & Grosse, 2012)) or cross-sectional population-based surveys (USD \$6132 (Liptak, Stuart, & Auinger, 2006)) were in the middle of the range. The differences in estimates were largely driven by variation in the breadth of services covered by each method and the characteristics of the sample under study. The annual cost per child increased over time, with reports of a 3% increase over a 3-year period (USD \$22079 in 2000 vs. USD \$22772 in 2003 (Wang, Mandell, Lawer, Cidav, & Leslie, 2013)) to a 20% increase over a 4-year period (USD \$4965 in 2000 vs. USD \$5979 in 2004 (Leslie & Martin, 2007)). The cost also depended on age; at ages 3-5 years, the annual cost was an estimated USD \$8185, which increased to USD \$22079 at ages 17-20 years (Cidav, Lawer, Marcus, & Mandell, 2013). This trend was largely driven by increased consumption of long-term and inpatient care in adolescents and young adults, while the use of occupational or speech therapy and diagnostic assessment decreased with age. Increased cost was also reported for children with comorbid intellectual disability (USD \$19787 (Peacock et al., 2012)), epilepsy (USD \$11847 (Peacock et al., 2012)) or more severe forms of ASD (£11029 in autism vs. £8968 in ASD (Barrett et al., 2015)).

The differences between cost estimates based on private insurance and Medicaid claims could be an indication of differences in ASD services coverage (e.g. occupational, speech or ABA-based interventions, pharmaceutical treatments, etc.) or in the characteristics of the individuals subscribed to private vs. public insurance plans. This was highlighted by a study from Wang et

al. (2013), which reported the cost of care, from the health insurer perspective, for a child on Medicaid was almost 4 times higher than for someone with private insurance (USD \$22653 vs. \$5254). But children on private insurance were younger, more likely to be male and live in urban counties. Another limitation to studies based on private or Medicaid insurance claims is that they only consider cost items from the perspectives of the private insurer or the Ministry of Health, respectively. This criticism can also be applied to studies based on cross-sectional surveys as these questionnaires focus on a specific subset of health services (e.g. ambulatory, inpatient) and are unlikely to provide a comprehensive picture of the range of health, psychosocial and educational services used by individuals with ASD. As ASD manifests early in life and individuals often require lifelong support, the economic impact on family members and ASD services outside of healthcare such as special education, respite care or employment assistance programs could contribute to the total cost of ASD.

Some studies (Barrett et al., 2015; Krister Järbrink, Fombonne, & Knapp, 2003; PACT consortium et al., 2011) have designed surveys to estimate productivity loss in both caregivers and patients by valuing time spent on informal care or absence from work using the human capital approach. Other studies (Buescher, Cidav, Knapp, & Mandell, 2014; Cidav et al., 2013; Ganz, 2007) have measured productivity loss in caregivers or patients and the cost of non-medical services by aggregating from different sources or by estimating the expected level of service use based on functional outcomes. As published epidemiological data on functional outcomes in adults with ASD are scarce, some cost items were either left out or extrapolated from paediatric studies. Moreover, parental productivity loss is only a portion of the economic impact on family members; none of the studies included other spillover effects such as increased psychological stress or adverse health outcomes in caregivers (Cidav, Marcus, & Mandell, 2012; Montes & Halterman, 2008). Therefore, studies using a societal or family perspective have likely underestimated the true cost of ASD.

1.2 Screening and Clinical Assessment

Despite increasing knowledge about the genetic and neurophysiological basis of ASD, there is no biological investigation that can definitively diagnose the disorder. After first suspicion of possible autistic traits by a caregiver or a healthcare professional, the child is first evaluated for

developmental delay and then referred to undergo diagnostic assessment if suspicions warrant a full work up (Johnson et al., 2007). The process of diagnosis may entail multiple assessments by several specialists. A formal diagnosis of ASD is based on fulfilling standardized criteria, most commonly the DSM-5 (American Psychiatric Association, 2013) or International Classification of Disease-10 (World Health Organization, 1992). However, variations in symptom presentation and diagnostic criteria with imperfect accuracy and reliability can delay clinical confirmation on ASD diagnosis.

1.2.1 Surveillance

The purpose of surveillance is to continuously monitor a child's developmental trajectory such that any developmental abnormalities can be detected early on. Children considered to be at heightened risk (e.g. has a full sibling diagnosed with ASD, of pre-term birth or has specific genetic conditions (Grønberg, Schendel, & Parner, 2013; Limperopoulos et al., 2008; Ozonoff et al., 2011; Richards, Jones, Groves, Moss, & Oliver, 2015)) require additional attention and typically have a lower threshold for referral for further assessment. Many healthcare professional groups, including the Canadian Paediatric Society, recommend that parents document any observed developmental concerns and that physicians regularly probe for attainment of age-appropriate milestones (Johnson et al., 2007; Williams et al., 2011). In Ontario, a standardized approach using the Nipissing District Developmental Screen (NDDS) is recommended to detect developmental delays.

1.2.1.1 Nipissing District Developmental Screen

The NDDS is a general developmental screen, designed to evaluate a child's developmental achievements in sensory, cognition, motor, social-emotional and self-help domains. There are 13 age-specific versions of the NDDS, covering the key developmental periods between 1 month to 6 years, and each screen consists of a short checklist of age-appropriate skills. Completed by the caregiver or physician, failure to achieve two or more items on the checklist suggests the need for further testing. Validation studies reported a sensitivity of 50-83% and specificity of 63-96% compared to the Bayley Scales of Infant Development II or III (Cairney et al., 2016; Currie et al., 2012; Dahinten & Ford, 2004). The test-retest reliability over a 2 week delay was 60% and the agreement between 12- and 18-month NDDS screens was 65% (Cairney et al., 2016; Nagy,

Ryan, & Robinson, 2002). Despite moderate psychometric properties, only 16.5% family physicians in Ontario accessed (i.e. purchased or downloaded) the NDDS in 2010 (Limbos, Joyce, & Roberts, 2010).

1.2.2 Screening

The purpose of ASD screening is to distinguish whether a child has any symptoms consistent with ASD and, unlike surveillance, it is carried out at set intervals (Johnson et al., 2007). There are two levels of screening: “Level 1” is used in primary care settings to identify those at-risk for ASD from the general population and “level 2” is a more lengthy assessment administered to children considered to be at high-risk for developmental delay (Robins, 2008). Conducted by a physician or community-based healthcare provider, ASD screening can involve review of medical and family history, direct observation of child’s behaviour or administration of ASD screening tool (Anagnostou et al., 2014; Johnson et al., 2007). The use of a standardized screening tool is reported to be associated with increased referral to appropriate ASD assessment and services by 30% (Johnson et al., 2007). However, <50% of primary care physicians regularly implement standardized screening for ASD or developmental delay (Bethell, Reuland, Schor, Abrahms, & Halfon, 2011; Dosreis, Weiner, Johnson, & Newschaffer, 2006; Radecki, Sand-Loud, O’Connor, Sharp, & Olson, 2011).

Universal screening, a population-based approach where all children with and without developmental delay are screened, is recommended by the AAP at 18 and 24 months (Johnson et al., 2007). One study reported that children with autism who followed the well-child visit schedule were diagnosed 1.6 months earlier than children who did not (Daniels & Mandell, 2013). In a randomized controlled trial of 2103 children in urban primary care practices, the use of standardized screening tools (i.e. Ages and Stages Questionnaire II and M-CHAT) at 18- and 24-month well-child visits identified 2-4% more children with developmental delay compared to surveillance monitoring for delay in age-appropriate milestones (Guevara et al., 2013). However, the combination of imperfect accuracy of screening tools and low ASD prevalence could lead to unnecessary demand for subsequent diagnostic services (i.e. low PPV), increased psychological distress in parents or missed ASD diagnosis due to false negative screens (Al-Qabandi et al., 2011; Duby & Johnson, 2009).

One cross-sectional study (Pinto-Martin et al., 2008) compared universal screening with the M-CHAT to sequential screening with a generic screening tool and the M-CHAT; they reported discordance between the two screening strategies. Sixteen percent of the children who screened positive on the M-CHAT did not fail the general screen, and 14% who failed the general screen passed the M-CHAT. This indicates that the general screening tool and ASD-specific screening tool covered different developmental concerns and are not exchangeable (Pinto-Martin et al., 2008). Standardized screening tools with good psychometric properties include the Checklist for Autism in Toddlers (CHAT) (Baron-Cohen, Allen, & Gillberg, 1992), M-CHAT (Robins et al., 2001), and Infant-Toddler Checklist (ICT) (Wetherby & Prizant, 2002).

1.2.2.1 Modified Checklist for Autism in Toddlers

The M-CHAT is a 23-item questionnaire completed by caregivers (Robins et al., 2001). As the items are related to age-dependent milestones, the questionnaire is designed for children between 16 and 30 months. The threshold for a positive screen is either failing any three items or any two of the six critical items related to joint attention, social relatedness and communication (Robins et al., 2001).

The initial study used the screening tool during well-child visits at 15, 18, or 24 months; it identified 21 of 4797 children with ASD, only 4 of whom were flagged by healthcare providers prior to M-CHAT screening (Robins, 2008). Other validation studies on high-risk children indicate that the M-CHAT has moderate psychometric properties: sensitivity of 0.77–0.97, specificity of 0.38–0.99, and PPV of 0.06–0.92 (Anagnostou et al., 2014; Chlebowski, Robins, Barton, & Fein, 2013; Robins et al., 2001; Yama et al., 2012). However, one study in rural US reported poor internal consistency in caregivers with less than a high school education (Cronbach's alpha = 0.43) or of minority status (Cronbach's alpha = 0.53) (Scarpa et al., 2013). Among 1604 Canadian children considered at low-risk, the rate of false positive findings increased with age, highlighting the age-dependency of the screening instrument (Yama et al., 2012).

Although the M-CHAT is accessible at no cost for research and clinical purposes, its use requires staff training and possible update to the medical record system. One US study estimated the cost of implementing an M-CHAT screening program at a primary care clinic would be USD \$420.0 for training, USD \$22.8 per month for M-CHAT administration and reimbursement of USD \$38.8 per month from insurance payers (Gura, Champagne, & Blood-Siegfried, 2011).

1.2.3 Diagnostic Assessment

Children who screened positive are subsequently referred for diagnostic assessment. Assessment should be carried out by a multidisciplinary team, led by a clinician experienced in assessment of developmental disabilities for the particular age of the child (Zwaigenbaum et al., 2009). Several best practice guidelines (Anagnostou et al., 2014; Johnson et al., 2007; Nachshen, Garcin, Moxness, Tremblay, Hutchinson, et al., 2008; National Institute for Health and Care Excellence, 2011) recommend assessments to cover medical and developmental history, social and communication, cognition, adaptive functioning and physical examination. Direct testing using standardized measures is applicable for cognitive development using tools such as the Bayley Scales of Infant Development (Bayley, 2006) or the Mullen Scales of Early Learning (Mullen, 1995), adaptive skills using the Vineland Adaptive Behavior Scales II (Sparrow, Balla, & Cicchetti, 2005), and language using the Preschool Language Scale-IV (Zimmerman, Steiner, & Pond, 2002). Social interaction, communication and play skills should be assessed using structured observation with directed tasks and parental report for less frequent behaviours. The use of standardized diagnostic tools such as the Autism Diagnostic Observation Schedule (ADOS) (Lord et al., 2000) also recommended to better guide diagnosis.

1.2.3.1 Autism Diagnostic Observation Schedule

The ADOS is a semi-structured assessment tool that enables a clinician to observe a child's communication, social interaction, play and imagination by engaging them in a set of directed activities (Lord, Volkmar, DiLavore, & Risi, 1999). Each item is scored on a 4-point scale and diagnosis of ASD is based on combined individual scores in the communication and social domains. The ADOS is recommended for children with a non-verbal mental age of 15 months or more and its specificity is reported to be lower in younger toddlers (Gotham, Risi, Pickles, & Lord, 2007). The diagnostic algorithms were revised in the 2nd edition (Lord, Rutter, et al., 2012)

and a modified version is available for children younger than age 30 months (Lord, Luyster, Gotham, & Guthrie, 2012).

Validation studies reported moderate agreement between the ADOS and clinical diagnosis (Mazefsky & Oswald, 2006; Molloy, Murray, Akers, Mitchell, & Manning-Courtney, 2011). The ADOS did not perform as well in pre-school children or those with less typical symptoms (Le Couteur, Haden, Hammal, & McConachie, 2008). This suggests that the ADOS should supplement diagnostic assessment and not be used alone. Also, the interpretation of its results depends on the child's broader developmental and medical history.

Diagnosis and treatment recommendations are based on review of all test results and the clinical judgment of the multidisciplinary team or by a clinician with experience in ASD. Children with subtle symptoms in early years, especially those with typical language and intellectual development, can be difficult to diagnose before 18 months. Therefore, continuous surveillance and follow-up assessments are often necessary for children who do not receive an initial diagnosis of ASD.

As the knowledge base on the underlying biology of ASD increases, there are additional clinical tests that could be integrated into diagnostic assessment. These include neuroimaging, and testing in audiology, vision and genetics. However, most diagnostic guidelines are conservative and state that additional tests should only be carried out if the child presents specific symptoms, physical signs or a family history and the test results could guide diagnosis.

To date, there is no published economic evaluation comparing the costs and outcomes of alternative screening or diagnostic strategies in children.

1.3 Clinical Genetic Testing

Approximately 10-20% of children with ASD have detectable chromosome abnormalities (Carter & Scherer, 2013; Devlin & Scherer, 2012). Given strong familial aggregation in ASD (i.e. up to 20% in biological siblings, >90% in monozygotic twins), information on the genetic basis can help determine recurrence risk for siblings and potentially aid family planning (Carter & Scherer,

2013; Grønberg et al., 2013; Ozonoff et al., 2011; Schaefer & Mendelsohn, 2013; Tick, Bolton, Happé, Rutter, & Rijdsdijk, 2016). Identifying the underlying genetic disorders can also guide treatment and disease monitoring for known comorbid conditions.

Individuals with a genetic etiology for ASD may present with congenital abnormalities and/or dysmorphic features, as these features suggest insult in early morphogenesis (Miles et al., 2005). These individuals are often considered to have complex ASD; the male-to-female ratio is lower (3.5 M: 1F) and functional outcome is worse compared to those children without such features (Miles et al., 2005). The individuals with more complex clinical presentations typically have more *de novo* mutations or inherited rare mutations than those with essential ASD (Marshall et al., 2008). As the inheritance of rare genetic variants can be important in determining clinical significance of test results, first-degree family members, typically biological parents, may also undergo genetic testing (Scherer & Dawson, 2011).

Table 1.1 lists select genetic disorders and genes associated with ASD and Table 1.2 summarizes current clinical guidelines on recommended genetic testing in ASD. CMA is the recommended first-line test by all recently published guidelines (Carter & Scherer, 2013; CCMG Cytogenetics Committee, 2010; Miller et al., 2010; Schaefer & Mendelsohn, 2013) and Fragile X testing is recommended by some. Despite these recommendations, clinical genetic testing is selectively implemented. Studies based in the US reported approximately 30% of children with ASD underwent genetic testing and one study that reported 30 of 42 (71%) caregivers were not aware of its availability (Amiet, Couchon, Carr, Carayol, & Cohen, 2014; Chen, Xu, Huang, & Dhar, 2013; Vande Wydeven, Kwan, Hardan, & Bernstein, 2012). The low uptake could be due to limited health insurance coverage, as one study reported 71% of children with ASD in France, where there is universal healthcare, underwent testing (Amiet et al., 2014). There could be other factors that influence families' decision to undergo testing as another US study (Shea, Newschaffer, Xie, Myers, & Mandell, 2014) reported 64% of children with ASD enrolled in Medicaid received genetic counselling but only 7% underwent genetic testing.

Table 1.1 The prevalence and phenotypes of select genetic disorder and mutations associated with autism spectrum disorder.

Monogenic Disorders		
Syndrome (Gene)	Prevalence	Phenotype
Rett syndrome (MECP2)	0.8-1.3% for females ^{1,2}	ID, microcephaly/deceleration of head growth, language regression and stereotypic hand movement ^{3,4}
PTEN Mutation (PTEN)	4.7% in ASD with macrocephaly ²	Macrocephaly (>3SD) ^{3,4} , skin pigmentation ¹
Tuberous sclerosis (TSC1 and TSC2)	1.1-1.3 ⁴	Skin lesions or pigmentation; family history of seizure, skin lesion or ID ⁴
Fragile X (FMR1)	1-3% ³	Connective tissue disorder, hyperactivity, language delay, repetitive behaviour ⁴
Copy Number Variations		
Locus	Prevalence	Phenotype
15q11-13 duplication	1% ASD with normal karyotype ¹	Moderate-severe ID, epilepsy, hypotonia, impaired language ¹
16p11.2 duplication	1% ¹	Microcephaly ⁴
16p11.2 deletion		Macrocephaly, language delay ⁴ , ID ¹
22q11.2 duplication or deletion		Psychiatric disorder and congenital abnormality in deletion ¹
7q11.23 duplication	0.2% ¹	Short philtrum, thin lips, poor verbal skills ¹
1q21.1 duplication	0.2% ¹	Macrocephaly, frontal bossing, hypertelorism, ID, DD ¹
1q21.1 deletion		Microcephaly, mild ID, dysmorphic facial features, behavioural symptoms, eye abnormalities ¹
Selected Implicated genes		
Gene (Cytoband)	Prevalence	Phenotype
SHANK3 (22q13)	1.1% ² -1.5% ³	Early onset, social deficit ³
NRXN1 (2p16.3)	0.15% ² -0.4% ¹	Variable ¹
NLGN3, NLGN4 (Xq13.1, Xp22.32-31)	0.8% ²	No set morphology or phenotype; variable ASD severity, insidious or regressive onset ³

¹(Carter & Scherer, 2013). ²(Lintas & Persico, 2008). ³(Gurrieri, 2012). ⁴(Miles, 2011).

Table 1.2 Recommendation for clinical genetic testing in autism spectrum disorder.

	Fragile X	PTEN	MeCP2	Karyotyping	CMA	Clinical Features
American College of Medical Genetics and Genomics (2013) ¹	●	●	●		●	
Canadian College of Medical Geneticists (2010) ²					●	
International Standard Cytogenomics Array Consortium (2010) ³					●	
Carter and Scherer (2013)	●	●	●		●	MeCP2 in females with ID, PTEN for macrocephaly

¹(Schaefer & Mendelsohn, 2013). ²(CCMG Cytogenetics Committee, 2010). ³(Miller et al., 2010).

While conventional cytogenetic tests are used as diagnostic tools, newer genome-wide testing platforms can identify additional variants that could have unknown clinical significance or be predictive of health conditions not related to ASD. This raises an ethical concern in test administration when results do not inform treatment or diagnosis, and have uncertain meaning with respect to disease risks. The current Canadian guideline by the Canadian College of Medical Geneticists (CCMG) recommends only reporting variants that are related to ASD and have established benefits to patients and family members (Boycott et al., 2015). CCMG further states that bioinformatic analysis should be restricted to those possibly related to the suspected conditions in order to avoid identification of incidental findings, even if variants are medically actionable, because of the potential high cost of subsequent health services and psychological burden on families. However, the American College of Medical Genetics and Genomics (ACMG) created a list of genes for which variants considered to be likely pathogenic should be reported regardless of the age of the patient (Kalia et al., 2016). The ACMG further states that analysis of genetic information not associated with primary health condition should be left to the discretion of the testing laboratory (Green et al., 2013). The one thing that both guidelines agree on is that competent adults should have the option to refuse disclosure of incidental findings at

the time of DNA sample collection (American College of Medical Genetics and Genomics, 2014; Boycott et al., 2015).

1.3.1 Fragile X Testing

A majority of the ASD clinical guidelines (Johnson et al., 2007; Nachshen, Garcin, Moxness, Tremblay, Hutchinson, et al., 2008; National Institute for Health and Care Excellence, 2011) recommend genetic testing for Fragile X syndrome in children with intellectual disability and a family history of intellectual disability or Fragile X syndrome. Fragile X syndrome can be diagnosed using Southern blot analysis or polymerase chain reaction (Sherman, Pletcher, & Driscoll, 2005). The cost of Fragile X testing is \$325 and the turnaround time is approximately 4-6 weeks (Hospital for Sick Children, 2015).

1.3.2 Fluorescent *in Situ* Hybridization

Florescent *in situ* hybridization (FISH) is often used to confirm abnormalities found in genome-wide approaches. FISH can detect structural rearrangements, deletion or duplication of DNA segments in specific regions, depending on the specific set of probes used. The choice of which probe(s) to use is one of the limitations of FISH as it requires the ordering physician to know the specific chromosomal region(s) to be investigated. The cost of FISH ranges from \$550 to \$675 per sample, depending on the number of probes used (Tsiplova et al., 2017). The bioinformatic analysis and interpretation of results is quick and is typically within 1 month (The Centre for Applied Genomics, 2015).

1.3.3 Chromosomal Microarray

The ACMG (Manning, Hudgins, & Professional Practice and Guidelines Committee, 2010; Schaefer & Mendelsohn, 2013), CCMG (CCMG Cytogenetics Committee, 2010), International Standard Cytogenomics Array Consortium (Miller et al., 2010) and American Academy of Child and Adolescent Psychiatry (Volkmar et al., 2014) also recommend the use of chromosomal microarray (CMA), which has a diagnostic yield of 10-20% in ASD (Carter & Scherer, 2013). High-resolution CMA platforms can detect submicroscopic duplications or deletions and approximately 3 times more clinically relevant variants than karyotyping (CCMG Cytogenetics Committee, 2010; Shen et al., 2010). CMA has identified recurrent variants in several

chromosome regions (e.g. 2p16.3, 7q11.23, 16p11.2, 17p12 and 22q13) and they are associated with complete or partial loss of function of genes in the region (Ch'ng, Kwok, Rogic, & Pavlidis, 2015; Devlin & Scherer, 2012). For example, neurexin is a cell-surface receptor whose gene, NRXN1, lies at 2p16.3 and its deletion occurs in 0.4% individuals with ASD. However, some of these candidate genes are also associated with other neuropsychiatric disorders (e.g. schizophrenia, attention deficit hyperactive disorder) and mutations are also present in individuals without ASD (i.e. incomplete penetrance) (Devlin & Scherer, 2012; Warriar, Chee, Smith, Chakrabarti, & Baron-Cohen, 2015). Another limitation to CMA is that it cannot detect low-level mosaicism and balanced rearrangements which occur in approximately 1.3% of children with ASD (Xu, Zwaigenbaum, Szatmari, & Scherer, 2004).

The cost of CMA is \$744, accounting for the costs of equipment, testing of biological parents, validation testing and test result interpretation (Tsiplova et al., 2017). Moreover, the unit cost of CMA decreases as more samples are being tested simultaneously as multiple samples can be arrayed in parallel. The turn-around time for CMA is roughly 4-8 weeks (Ny Hoang, MSc, email communication, April 2016; Melissa Carter, MD, email communication, April 2017).

1.3.4 Clinical Genome and Exome Sequencing

CGES is increasingly being used in research and select clinical settings. CGES allows for more comprehensive and uniform coverage of the exome or genome, and preliminary studies in ASD have reported a higher detection rate of *de novo* or rare inherited mutations than previously reported by CMA studies (Jiang et al., 2013; Shen et al., 2010). Another advantage of CGES is that it does not require any prior knowledge of the underlying genetic abnormality (i.e. compared to single gene or gene panel tests) and might be more appropriate when clinical features suggest a single underlying genetic condition but a specific syndromic diagnosis cannot be made by clinical impression alone (Yu et al., 2013). CGES cannot detect structural rearrangements and GS has lower sensitivity for copy number variations (CNV) compared to CMA (Vermeesch et al., 2007; Yu et al., 2013). Evidence on phenotypes and comorbid conditions associated with variants detected by CGES is still emerging and the clinical significance of the identified variants can be uncertain. Currently, CGES is used selectively and is typically reserved for individuals with syndromic features (i.e. having one or more clinical features in addition to those typically

associated with idiopathic ASD, such as dysmorphic features, congenital abnormalities, and over- or under-growth) who have failed to receive a genetic diagnosis after exhausting all alternative tests (i.e. selective gene testing, CMA). Findings from one study comparing CGES to more targeted approaches support the conservative use of CGES because of its limitations (i.e. CGES cannot detect balanced rearrangements or some CNV). Also, using CGES is cost-effective as first-line only if it eliminates the need for further genetic testing (Shashi et al., 2014).

There are two types of CGES: ES and GS. By focusing on the protein coding regions that constitute 1% of the genome (i.e. exome), ES requires less bioinformatic analysis than GS. In ASD, however, additional mutations were identified with GS compared to ES, indicating potential causal contributions of regions that have unknown functional significance (Devlin & Scherer, 2012). One study comparing the two tests in ASD reported that GS could identify 9.2% of ASD-linked variants that were not detected by ES and GS has more coverage in chromosomal regions with high numbers of ASD susceptibility genes (Jiang et al., 2013). The diagnostic yield of ES in ASD was 4% (positive findings in 11 of 277 individuals) in one study (Yu et al., 2013) and 8% (8 of 95; 95% CI: 4%, 16%) in another (Tammimies et al., 2015). GS had a higher diagnostic yield; it identified pathogenic variants in 6 of 32 (19%) children with ASD in one study (Jiang et al., 2013) and in 36 of 85 (42%) children with ASD in another (Yuen et al., 2015). Overall, the reported diagnostic yields were lower in a study that excluded participants with previously diagnosed genetic disorders (Yu et al., 2013) and higher in studies that selected families based on specific criteria (e.g. complete DNA samples of both parents and siblings) (Jiang et al., 2013; Yuen et al., 2015). Differences in the set of genetic variants annotated as pathogenic for ASD further complicates direct comparison across studies and only some studies (Jiang et al., 2013; Tammimies et al., 2015; Yuen et al., 2015) used the pathogenic classification created by ACMG.

The laboratory testing costs of GS and ES have decreased over the years and National Institute of Health-Human Genome Research Institute estimated each megabase of DNA costs <US \$1.0 to sequence (Wetterstrand, 2015). Accounting for the costs of bioinformatic analysis, clinical interpretation of genetic variants, genetic testing of biological parents and validation testing of the identified variants, however, increases the cost to \$1978 for ES and \$3551 for GS (Tsiplova

et al., 2017). Moreover, delineating clinical significance of detected variants requires additional attention and turnaround time can take 20-40 weeks (Ny Hoang, MSc, email communication, 2016).

1.3.5 Economic Evaluation of Genetic Testing

Of the publications that compared the cost-effectiveness of clinical genetic testing in paediatrics, there were three studies that compared CMA to karyotyping in intellectual disability (ID) or developmental disability (DD) (Regier, Friedman, & Marra, 2010; Trakadis & Shevell, 2011; Wordsworth et al., 2007) and two studies that looked at the cost of CGES in the same conditions (Monroe et al., 2016; Tsiplova et al., 2017). Two were conducted in Canada and none considered costs from multiple payer perspectives or included downstream healthcare costs. Moreover, they did not account for differences in turnaround time of test results, the consequences of missing a genetic diagnosis, or wait times associated with accessing genetic testing and genetic counselling.

Trakadis and Shevell (Trakadis & Shevell, 2011) compared the cost-effectiveness of CMA to karyotyping in 114 children with DD who visited an academic paediatric neurology clinic between 2006 to 2009. The incremental cost of an additional diagnosis was CAD \$1379 for CMA compared to karyotyping. Also, CMA identified CNV in 9 children with normal karyotype and 37 FISH tests (i.e. the follow-up test for negative karyotype) would have been unnecessary if CMA was conducted as first-line. However, 21 (47%) children would have received the same clinical diagnosis without any genetic testing.

Another study compared CMA to karyotyping, but for children with ID in the UK (Wordsworth et al., 2007). The estimated cost of CMA was £442 and karyotyping was £117. The incremental cost per additional diagnosis was £2440, assuming the diagnostic yield of CMA was 15% higher than karyotyping and multi-telomere FISH testing. The authors also noted that the estimates were conservative given the costs of follow-up tests for CMA were excluded due to the lack of information.

Similar findings were reported in a study comparing CMA to karyotyping in children with ID, where the incremental cost-effectiveness ratio (ICER) was CAD \$2646 (95% CI: 1619, 5296) for each additional diagnosis (Regier et al., 2010). Based on the cost-effectiveness acceptability curve, CMA could be considered cost-effective compared to karyotyping if the willingness to pay for one additional diagnosis was CAD \$4550. Furthermore, the use of combined testing with karyotyping and CMA was more expensive and less effective compared to sequential testing (i.e. CMA only after negative karyotype). In another study by the same research group, the willingness to pay for CMA results in parents of children with DD was CAD \$1118 (95% CI: 498, 1788) (Regier, Friedman, Makela, Ryan, & Marra, 2009).

Monroe et al. (Monroe et al., 2016) estimated that if ES was offered to 17 individuals with ID as first-line diagnostic test, it would result in cost saving of USD \$4986-5699 in place of other subsequent genetic tests and USD \$2533 in place of subsequent metabolic testing.

Another study (Tsiplova et al., 2017) compared the costs and consequences of offering CMA, ES or GS to children with DD or ASD from the institutional payer perspective in Canada. During the 1st year of a 5-year program, the per-sample cost of CMA was CAD \$744 (95% CI: 714, 773), ES was CAD \$1655 (95% CI: 1611, 1699), GS using Illumina HiSeq 2500 was CAD \$5519 (95% CI: 5244, 5785) and GS using Illumina HiSeq X was CAD \$2851 (95% CI: 2750, 2956). The annual cost comprised of the cost of labour (e.g. DNA sample preparation and processing, clinical interpretation and report writing), bioinformatics (e.g. maintenance, computation, file storage), equipment (i.e. array or sequencing machine), supplies (e.g. reagents, shipping and handling of DNA samples) and follow-up testing (i.e. validation of positive tests in patient and follow-up tests in biological parents). Supplies made up the highest proportion of total cost across platforms at 40-74%. Compared to CMA, the cost of identifying an additional patient with any clinically-relevant genetic variant was CAD \$25458 for sequential testing with CMA and ES, CAD \$58959 for GS using HiSeq 2500 and \$26020 for GS using HiSeq X.

1.4 Study Rationale and Thesis Structure

As the evidence on the etiology and early risk markers of ASD emerges, clinical guidelines (Anagnostou et al., 2014; Johnson et al., 2007; Schaefer & Mendelsohn, 2013) are

recommending additional clinical services to identify early indicators. More ASD screening and diagnostic services further complicate the diagnostic pathway, potentially reducing the efficiency and effectiveness the healthcare system. In turn, tactful resource allocation strategies need to be identified such that children can access the services they need in a timely manner. The overall theme of this thesis focuses on estimating the impact of introducing two services within the ASD clinical pathway on the individual- and health system-level. This thesis consists of three manuscripts, followed by an in-depth discussion chapter.

Chapter 2 examines ASD screening. The objective was to conduct a meta-analysis to summarize the sensitivity, specificity and PPV of a commonly cited and used screening tool, the M-CHAT. A meta-regression was also carried out to estimate the extent to which the accuracy measures change in relation to age at screening, sex and study design. This information is necessary to determine the potential effectiveness of using the M-CHAT as a universal screening tool. Also, understanding how the accuracy measures varies by clinical characteristics would help clinicians better interpret an individual's M-CHAT score in relation to their clinical profile and decide if it is an appropriate screening tool to use.

The objective of Chapter 3 is to quantify the incremental costs and health outcomes of universal or high-risk screening compared to surveillance monitoring in ASD. All children were screened with the M-CHAT at 18 and 24 months in universal screening, and only children with a full sibling with ASD underwent M-CHAT screening at the same two time points in high-risk screening. The health outcomes, number of children accurately diagnosed with ASD prior to age 3 years and the number of children that initiated ASD treatment prior to age 4 years, were selected to be consistent with recommendations made to the Ontario Ministry of Children and Youth Services (Auditor General of Ontario, 2015).

Chapter 4 estimates the incremental costs and effects of using newer genetic testing approaches relative to the standard care in children diagnosed with ASD. Although GS and ES can identify additional pathogenic variants compared to CMA, they are much higher in cost and test results might not have immediate clinical benefit. In turn, the CEA compared the cost and number of

children with ASD diagnosed with rare pathogenic variants in CMA alone to three alternate testing strategies: ES for all children, GS for all children or CMA followed by ES in children who have syndromic features and had negative findings on CMA.

The discussion chapter, Chapter 5, summarizes the main findings of the three manuscripts and discusses their implications for the healthcare system and clinical care for children with ASD. Results from this thesis could help policy makers and clinicians determine whether implementation of universal screening and newer genetic testing services are worthwhile investments, and how to best integrate ASD screening and genomic sequencing in clinical care. From a methodological standpoint, studies in this thesis contributes to the limited body of work using simulation models in ASD and Bayesian meta-analysis using bivariate regression modelling.

2. META-ANALYSIS OF THE MODIFIED CHECKLIST FOR AUTISM IN TODDLERS

2.1 Preface

Among the growing list of recommended ASD screening tools, the M-CHAT is one that is most cited and it performs with moderate accuracy (Siu et al., 2016; Zwaigenbaum, Bauman, Fein, et al., 2015). Moreover, it has been validated in children with and without developmental concerns and translated into numerous languages (Canal-Bedia et al., 2011; Kara et al., 2014; Perera, Wijewardena, & Aluthwelage, 2009; Seif Eldin et al., 2008; Sunita & Bilszta, 2013). As it was designed to screen both low- and high-risk children aged 16 to 30 months (Robins, 2008), it could potentially be used in universal screening at 18 and 24 months. In order to determine if it would be appropriate for use on a population level, it is critical to understand how its psychometric properties change across subpopulations. The objective of this meta-analysis was to summarize the sensitivity, specificity and PPV of the M-CHAT in the published literature. Meta-regressions were carried out to quantify the extent to which the accuracy measures change in relation to the proportion of male, age at screening and study design.

2.2 Modelling Accuracy Measures

As the sensitivity and specificity in each study were correlated, these two measures were jointly modelled using a bivariate regression model to account for their covariance. The number of true positive, TP_i , and true negative, TN_i , were assumed to follow two binomial distributions:

$$TP_i \sim \text{Bin}(\text{sensitivity}_i, \text{positive}_i) \quad TN_i \sim \text{Bin}(\text{specificity}_i, \text{negative}_i)$$

where positive_i is the number of positive screens in the i th study and negative_i is the number of negative screens in the i th study. The study-specific values were then transformed on the logit scale and modelled as a bivariate normal distribution,

$$\text{logit} \begin{pmatrix} \text{sensitivity}_i \\ \text{specificity}_i \end{pmatrix} \sim N \begin{pmatrix} \mu_{\text{sensitivity}} & \sigma_{\text{sensitivity}}^2 & \rho \sigma_{\text{sensitivity}} \sigma_{\text{specificity}} \\ \mu_{\text{specificity}} & \rho \sigma_{\text{sensitivity}} \sigma_{\text{specificity}} & \sigma_{\text{specificity}}^2 \end{pmatrix}$$

where $\mu_{sensitivity}$ is the pooled sensitivity, $\mu_{specificity}$ is the pooled specificity, $\sigma_{sensitivity}^2$ is the variance of pooled sensitivity, $\sigma_{specificity}^2$ is the variance of pooled specificity, and $\rho\sigma_{sensitivity}\sigma_{specificity}$ is the covariance between pooled sensitivity and specificity.

PPV increases as the prevalence of the underlying condition increases. Therefore, study-specific PPV are recommended to be standardized prior to pooling (Bossuyt et al., 2013). The standardized PPV is calculated by applying a common prevalence rate to all study-specific sensitive and specificity,

$$PPV_i = \frac{sensitivity_i * prevalence}{sensitivity_i * specificity_i + (1 - specificity_i) * (1 - prevalence)}$$

where $sensitivity_i$ is the sensitivity of the i^{th} study, $specificity_i$ is the specificity of the i^{th} study and PPV_i is the standardized PPV of the i^{th} study. After standardization, the study-specific PPVs were modelled as a normal distribution,

$$PPV_i \sim N(\mu_{PPV}, \sigma_{PPV}^2)$$

where μ_{PPV} is the pooled PPV and is the σ_{PPV}^2 variance of the pooled PPV.

Heterogeneity across study-specific values was assessed using the I^2 -test (Higgins, 2003). To quantify the heterogeneity, the above two models (the bivariate model for sensitivity and specificity and the model for PPV) were extended into meta-regression models. The three covariates (proportion of males, age at screening and study design) was modeled one at a time due to the low number of studies. In turn, there were six meta-regression models in total.

The following is a manuscript submitted for publication, with minor modification on formatting such that it is cohesive with the rest of the thesis.

2.3 Manuscript #1

2.3.1 Abstract

Background: The Modified Checklist for Autism in Toddlers (M-CHAT) is one of the recommended autism spectrum disorder (ASD) screening tools. Validation studies reported

differences in psychometric properties across sample populations. This meta-analysis quantitatively summarized the sensitivity, specificity and positive predictive value (PPV) and estimated the extent to which the accuracy measures change in relation to age at screening, proportion of males and study design.

Method: Four electronic databases (Medline, PsychInfo, CINAHL and EMBASE) were searched to identify articles published between January 2001 and May 2016. Retrieved articles were evaluated for eligibility. Quality of the selected articles was assessed using the QUADAS-2. Study-specific sensitivity and specificity were pooled using a bivariate hierarchical regression model. PPVs were standardized using ASD prevalence reflective of a low-risk and a high-risk population, then pooled using a similar method. Meta-regressions were carried out with age at screening, study design and proportion of males as covariates.

Results: Thirteen studies were included. Pooled sensitivity was 0.83 (95% CrI: 0.75, 0.90), specificity was 0.51 (95% CrI: 0.41, 0.61), PPV was 0.55 (95% CrI: 0.45, 0.66) in high-risk children and 0.07 (95% CrI: <0.01, 0.16) in low-risk children. Sensitivity was higher for screening at 30 months compared to 24 months, but there were large variations in accuracy measures between studies based on design and the proportion of males.

Conclusion: Given its low specificity in high-risk children and lack of evidence on its psychometric properties in low-risk children, there is limited evidence supporting the use of the M-CHAT as a population screening tool. Clinicians should consider the child's age, sex and latent risk of ASD when deciding to use of the M-CHAT and interpreting its score.

Keywords: autism spectrum disorder; Modified Checklist for Autism in Toddlers; screening

Acronyms: ASD: autism spectrum disorder; CrI: credible interval; M-CHAT: Modified Checklist for Autism in Toddlers; PPV: positive predicative value.

2.3.2 Introduction

The Modified Checklist for Autism in Toddlers was designed to screen for autism spectrum disorder (ASD) in children between 16 and 30 months (Robins et al., 2001). Validation studies have reported moderate psychometric properties across various populations, potentially supporting its use as part of universal screening. The purpose of ASD screening using a standardized tool is to systematically identify early signs of ASD and the American Academy of Pediatrics (AAP) (Johnson et al., 2007) recommends that it should take place at 18 and 24 months for all children.

The original M-CHAT consists of 23 questions about behaviours that are potential early signs of ASD in very young children (Robins et al., 2001). Parents or caregivers complete the checklist based on their child's current skills and behaviours. Its brevity is one of its advantages, as is the fact that it does not require responses based on clinical observation by a trained clinician. A child is considered to be at-risk if he/she fails more than two of the six critical items (Critical6) or three of the 23 items (Total23) on the checklist. Among the 1293 children screened in the original study, 132 screened positive and 39 were later diagnosed with ASD. The M-CHAT Follow-up Interview, a 5-20 minutes structured phone interview used to confirm items endorsed by the parent, can be carried out in addition for children who screen positive prior to full developmental assessment. The follow-up interview was reported to improve positive predictive value (PPV) (Robins, 2008). A revised version, M-CHAT R/F, excluded 3 items from the original M-CHAT and was tested in a low-risk sample with PPV of 0.14 (Robins et al., 2014).

Since its publication in 2001, several studies have estimated the accuracy of the M-CHAT in children with or without developmental concerns screened at different ages. A “developmental concern” was typically defined as atypical development identified by a physician and leading to referral for diagnostic assessment or behavioural intervention. Previous studies (Siu et al., 2016; Sunita & Bilszta, 2013) that reviewed the validation studies identified methodological limitations such as high drop-out rates, reporting aggregated results from high- and low-risk children together and a lack of blinding between screening and assessment, which could have biased study findings. As the items on the checklist are related to age-specific development, psychometric properties also varied with age. One study reported that the proportion of positive

screens increased with age, particularly in children older than the intended age range of the M-CHAT at 33-48 months (Yama et al., 2012).

Given the heterogeneity in the presentation of ASD symptoms, variation in psychometric properties across study samples can be expected. Among children with ASD, observable variations in sensory, social communication, and repetitive behaviour have been reported by the second year of life, but the severity and type of deficits vary by age, cognitive development and co-occurring conditions (Elsabbagh & Johnson, 2010; Jones, Gliga, Bedford, Charman, & Johnson, 2014). Moreover, subtle delays present at around 12 months might not be ASD-specific and early ASD screening could falsely identify children with other developmental concerns (e.g. simple language delay) as having ASD (Zwaigenbaum, Bauman, Fein, et al., 2015).

Quantifying the direction and magnitude by which the accuracy of the M-CHAT differs according to key characteristics (e.g. age at screening, sex of the child) would help determine if the M-CHAT is appropriate for universal screening. Moreover, understanding how the psychometric properties change could help clinicians interpret an individual's M-CHAT score in relation to the patient's clinical profile.

The purpose of this study was to quantitatively summarize the accuracy (i.e. sensitivity, specificity and PPV) of the M-CHAT in children screened for ASD and to quantify the extent to which these measures of accuracy change in relation to the age at screening, gender distribution, study design and background risk for ASD.

2.3.3 Methods

2.3.3.1 Search Strategy

The literature search, data extraction and statistical analyses were based on the Cochrane guidelines for systematic reviews of diagnostic tests (Deeks, Bossuyt, & Gatsonis, 2013). A literature search was carried out using MESH headings and keywords associated with ASD (autistic disorder; autism spectrum disorder; autism; ASD; child development disorders, pervasive; Asperger's syndrome) and M-CHAT (M-CHAT, modified checklist for autism in

toddlers) using Medline, EMBASE, PsychInfo, and CINAHL in May 2016. After duplicated entries were removed, titles and abstracts of the retrieved articles were assessed by one reviewer (TY). Full texts of selected articles were reviewed based on the inclusion criteria. The reference lists of the included articles were scanned for additional articles.

2.3.3.2 Selection Criteria

Studies were included if they met all of these inclusion criteria: 1) screened children for ASD using the original M-CHAT, 2) reported the results of screening and clinical diagnosis of the children, 3) sample population was not selected based on any medical condition other than developmental delays (e.g. low birth weight, Down syndrome), and 4) published in English. As three American studies screened a small proportion of children using the Spanish version of the M-CHAT, the language criterion was broadened to include studies that screened >90% children using the English M-CHAT. The two studies that used the M-CHAT R/F were excluded as results could not be converted to match the original M-CHAT. Articles published between January 2001, year of publication of the M-CHAT, and May 2016 were screened. If more than one article used the same study population, only the article with the largest sample size was included. Figure 1 describes the number of articles retrieved and the reasons for exclusion as recommended by Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guideline (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009). Thirteen studies satisfied all inclusion criteria and were included in this meta-analysis.

2.3.3.3 Data Extraction and Quality Appraisal

For this meta-analysis, studies based on children with any developmental concerns were categorized as “high risk”, studies on children without identified developmental concerns were categorized as “low risk”, and studies that included both children with and without developmental concerns were categorized as “mixed”. A positive screen was defined as failing either of the scoring algorithms-- Critical6 (failing more than two of the six critical items) or Total23 (failing more than three items on the M-CHAT) (Robins et al., 2001). The reference standard for the screening test was a clinical diagnosis of ASD, as defined by each study.

Information on publication (author, year of publication, country of study), participants (age, sex, ASD diagnosis, high/low-risk), study procedure (method of recruitment, ASD diagnostic criteria, M-CHAT scoring algorithm) and results (number of positive and negative screens, true positive, true negative) were extracted from the included studies using a data collection form. As few studies carried out the recommended telephone follow-up phone interview (Robins et al., 2001), the M-CHAT results prior to the follow-up interview were used for the analysis. The corresponding authors were contacted if any information was missing from the publication.

The quality of the included articles was appraised independently by two researchers (TY, MP) using the QUADAS-2 (Whiting, 2011). It was modified to better evaluate biases in sample selection, implementation of the M-CHAT and diagnostic assessment of ASD studies (Appendix 2.1). Inter-rater agreement for the initial quality appraisal was quantified using Cohen's kappa. Disagreements in assessments were jointly reviewed until a consensus was reached between the two reviewers.

2.3.3.4 Statistical Analysis

The sensitivity and specificity were pooled using a bivariate hierarchical (random-effects) model under a Bayesian framework (Chu & Cole, 2006; Verde, 2010). A bivariate model was selected as it can account for the covariance between sensitivity and specificity in each study by jointly estimating the two values. For each study, the numbers of true positive and true negative screens were assumed to follow two binomial distributions. The study-specific values were then transformed on the logit scale and modelled as a bivariate normal distribution.

As the PPV varies depending on the prevalence of the underlying condition, study-specific PPV was standardized using ASD prevalence reflective of a low-risk (in general population, 1 in 68) and a high-risk (mean prevalence across included high-risk studies, 45.6%) population (Bossuyt et al., 2013; Centers for Disease Control and Prevention, 2016). For studies where sensitivity and specificity could not be calculated, the reported PPVs (i.e. non-standardized) on low-risk children were pooled with PPVs standardized using the low-risk ASD prevalence and reported PPVs on high-risk children were pooled with PPVs standardized using the high-risk prevalence. The study-specific PPVs were also pooled using Bayesian hierarchical models.

For the joint sensitivity and specificity model and the PPV regression models, the posterior distributions of the parameters were estimated using single-chain Gibbs sampling of 11,000 replications with the first 1000 iterations discarded. Non-informative prior distributions were used for all parameters, except a weakly informative scaled inverse Wishart distribution was used for the covariance matrix in the bivariate normal distribution. The estimated study-specific values, pooled estimates and their 95% credible intervals (CrI) were transformed back to the probability scale and summarized in forest plots. Convergence was assessed graphically and heterogeneity was quantified with I^2 -test.

2.3.3.5 Investigation of Heterogeneity

Meta-regressions were carried out to quantify the differences in sensitivity, specificity and PPV in relation to mean age at screening, proportion of males and study design (retrospective vs. prospective). A regression model was built for each of the three covariates paired with each outcome due to low number of studies. Non-informative normal prior distributions were used for all beta-coefficients in the models.

2.3.4 Results

2.3.4.1 Study Characteristics

Tables 2.1 and 2.2 summarized the characteristics and results of the included studies. All were carried out in North America except two (Charman et al., 2016; Koh et al., 2014). Five retrospective studies collected information using medical records (Cogan-Ferchalk, 2013; Goodwin, 2010; Koh et al., 2014; Ringwood, 2010; Salisbury, 2016). On average, children were screened at 21-41 months and four studies (Charman et al., 2016; Goodwin, 2010; Koh et al., 2014; Snow & Lecavalier, 2008) screened children above the M-CHAT recommended age range. ASD diagnostic assessment was carried out at 24-52 months. High-risk studies (n=12) had higher proportion of boys compared to girls, while low-risk studies had gender parity. The prevalence of ASD in studies based on children with developmental concerns (range 38-66%) was higher than that of the general population.

The quality ratings are summarized in Table 2.3. Inter-rater agreement of the initial assessments between the two reviewers was moderate-to-high ($\kappa=0.78$). Potential sources of bias in participant selection and implementation of M-CHAT and/or ASD diagnosis were identified in all studies. High risk of selection bias was introduced through volunteer bias from families, as indicated by moderate (50-60%) surveys response rates, or from potential selective referral by participating clinicians. The reported sociodemographic and functional characteristics of the sample populations in 70% of the studies can be considered reflective of children with ASD or developmental concerns typically presented at clinical settings, based on clinical opinion. In half of the studies, the interval between M-CHAT screening and clinical diagnosis was not reported or was greater than six months. Sixty percent of the studies did not include all participants in sensitivity and specificity calculations, with few providing the reasons for exclusion. Only two studies reported that the interpreter of the M-CHAT was blinded to the child's clinical diagnosis or that the M-CHAT was scored prior to clinical diagnosis. ASD diagnostic assessments were carried out by psychologists and, typically, one or more qualified professionals (i.e. psychometrist, developmental pediatrician, occupational therapist, speech language pathologist or special education teacher). One study did not provide details on diagnostic assessment (Kleinman et al., 2008).

2.3.4.2 Studies on Low-risk Children

One study (Robins, 2008) exclusively focused on low-risk children and two mixed studies (Kleinman et al., 2008; Robins et al., 2001) included children without developmental concerns. Only the PPVs were reported in all three studies, as children who screened negative on the M-CHAT did not undergo ASD diagnostic assessment. Thus, the sensitivity and specificity could not be calculated, and the estimated prevalences in Table 2.1 are likely underestimated for these studies. The PPV ranged from 0.06 – 0.11, but increased to 0.42 - 0.65 if the M-CHAT telephone follow-up interview was administered.

2.3.4.3 Studies on High-risk Children

Of the 12 studies that included children with developmental concerns, three (Charman et al., 2016; Fessenden, 2013; Goodwin, 2010) sampled from children referred for unspecified mental

health services, three (Cogan-Ferchalk, 2013; Kleinman et al., 2008; Robins et al., 2001) from children referred for early behavioural intervention and six (Eaves, 2006; Kleinman et al., 2008; Koh et al., 2014; Ringwood, 2010; Salisbury, 2016; Snow & Lecavalier, 2008) from children referred for diagnostic assessment for developmental conditions. Children were typically screened at 2-3 years old, with one study (Goodwin, 2010) reporting a mean age above 40 months. Clinical diagnosis of ASD was established on average within 6 months after screening in seven studies (Eaves, 2006; Goodwin, 2010; Kleinman et al., 2008; Robins, 2008; Robins et al., 2001; Snow & Lecavalier, 2008; Villalobos, 2011) and was not reported by three (Cogan-Ferchalk, 2013; Fessenden, 2013; Ringwood, 2010). Without the telephone interview, the sensitivity ranged from 0.64-0.96, specificity from 0.35-0.67 and PPV from 0.49-0.73.

2.3.4.4 Sensitivity and Specificity

The study-specific and pooled sensitivity and specificity are summarized in Figure 2.2. Based on the nine studies where sensitivity and specificity could be calculated, heterogeneity was high ($I^2 > 90$ for either measure) and the pooled sensitivity was 0.83 (95% CrI: 0.75, 0.90) and pooled specificity was 0.51 (95% CrI: 0.41, 0.61). Specificity was comparable across different ages at screening but sensitivity was higher at 30 months (sensitivity: 0.69, 95% CI: 0.19, 0.86; specificity: 0.46, 95% CI: 0.03, 0.64) compared to 24 months (sensitivity: 0.55, 95% CI: 0.02, 0.84; specificity: 0.45, 95% CI: <0.01, 0.71). No meaningful differences were observed in the sensitivity and specificity between retrospective and prospective studies or between low and high proportions of males (Appendix 2.2).

2.3.4.5 Positive Predictive Value

The PPVs from nine studies were standardized using the ASD prevalence in the general population (Centers for Disease Control and Prevention, 2016) and pooled with the reported PPVs on low-risk children from three studies (Kleinman et al., 2008; Robins, 2008; Robins et al., 2001). The pooled PPV in low-risk children was 0.07 (95% CrI: <0.01, 0.16). Large difference was estimated between the PPVs standardized using ASD prevalence in the general population (0.03, 95% CrI: <0.01, 0.20) and reported PPV in low-risk children that could not be standardized (0.24, 95% CrI: 0.08, 0.40). When the PPVs were standardized using the mean

ASD prevalence across the included high-risk studies, 45.6%, and combined with the reported PPV on high-risk children from three studies (Kleinman et al., 2008; Robins, 2008; Robins et al., 2001), the pooled PPV was 0.55 (95% CrI: 0.45, 0.66). Minimal difference was estimated between the standardized and non-standardized high-risk PPVs.

Given high variability across PPVs in low-risk children, only high-risk PPVs were analyzed using meta-regressions. The point estimate of the PPV in high-risk children increased as the proportion of males increased (75% male: 0.58, 95% CI: 0.52, 0.64; 100% male: 0.46, 95% CI: 0.14, 0.78) but the credible intervals overlapped. There were no clinically relevant differences in PPV by age at screening or study design (Appendix 2.2).

2.3.5 Discussion

This meta-analysis quantitatively summarized the accuracy of the M-CHAT and we conclude that it can identify ASD with low to moderate sensitivity and specificity among children with developmental concerns. Partially due to the low number of studies and high between-study heterogeneity, the estimated values were imprecise and the credible intervals were wide. Meta-regression results suggest that the M-CHAT can better identify children with ASD among older children (at 30 compared to 24 months). Although screening at 18 months is recommended by the AAP (Johnson et al., 2007) and within the intended age range of the M-CHAT, the accuracy measures could not be precisely predicted for this age in this meta-analysis as few studies screened 18 month-old children. Standardized PPVs were sensitive to changes in ASD prevalence and the pooled values were only slightly higher than the pre-test odds in both low- and high-risk populations. Findings from the meta-regressions, however, should be interpreted with caution as adjusting for confounding was not feasible due to low number of studies. While age and gender distribution, on average, did not differ by study design, there could be unaccounted factors which contributed to identified differences in accuracy measures.

Given the moderate specificity of the M-CHAT and low prevalence of ASD in the general population, many children referred for subsequent follow-up would be due to a false positive M-CHAT screen and not require in-depth assessment. In turn, this would further delay access for diagnostic assessment for children who would benefit from additional follow-up. Similarly, the

moderate sensitivity suggests that a proportion of children who would benefit from further assessment would not be able to access it because they were falsely identified as screen negative. Both scenarios could potentially delay access to ASD interventions that could revert the child's developmental trajectory to a more age-appropriate level.

Methodological issues were identified in all studies, with high risk of biases in participant selection and interpretation of M-CHAT or clinical diagnosis. In particular, bias from selective referral by participating clinicians, M-CHAT scoring not blinded to clinical diagnosis and a lack of detail on diagnostic assessment were consistently noted. As clinicians could be more likely to diagnose children with ASD after a positive M-CHAT screen, reported accuracy measures could be overestimated due to the lack of blinding. The inconsistent and often long time intervals between screening and diagnosis were also a source of concern; given rapid development that occurs in toddlers, longer intervals could bias study findings in either direction. The lack of complete follow-up in studies with low-risk children also prevented estimation of the pooled sensitivity and specificity in low-risk populations. Difference between reported PPVs in low-risk children and PPVs standardized using the ASD prevalence in the general population suggests that the children in low-risk studies might not be truly low-risk or representative of the general population. Evidence on the accuracy of M-CHAT or other ASD screening tools in the general population is one of the components needed to determine the clinical utility of universal ASD screening, along with timely access to diagnostic assessment and intervention.

Although the M-CHAT is intended to be used in both low- and high-risk children (Robins, 2008; Robins et al., 2001), study findings indicate that it performs with low-to-moderate accuracy in high-risk children and there was a lack of evidence supporting its use in low-risk children. Furthermore, the low pooled specificity of the M-CHAT suggests that it would not be appropriate as a universal screening tool on its own. If all screen-positive children were referred for ASD diagnostic assessment, the children with false positive screens would needlessly undergo lengthy assessment and decrease the efficiency of the diagnostic pathway (i.e. lengthen the wait time for children who truly require further assessment and increase healthcare expenditures). Although the pooled sensitivity in high-risk children is in the moderate range, the

consequences of a false negative screen could be severe if it delays ASD diagnosis and timely access to appropriate interventions.

The M-CHAT R/F (Robins et al., 2014) reports greater PPV, but it also requires a structured phone interview following a positive screen. Implementation of screening with a standardized tool is already low in clinical settings (Bethell et al., 2011; Radecki et al., 2011). The additional follow-up interview would likely further reduce uptake or compliance with protocol, which would decrease accuracy of screening. However, routine developmental surveillance (i.e. careful monitoring of any signs of developmental concerns over time by a general paediatrician or practitioner) is a critical and recommended (Nachshen, Garcin, Moxness, Tremblay, & Hutchison, 2008; National Institute for Health and Care Excellence, 2011) component of routine check-ups. The use of a standardized ASD tool (e.g. the M-CHAT) might be more appropriate as a second-line screen carried out by primary care physicians for children with developmental concerns to better guide the course of follow-up assessment.

2.3.5.1 Limitations

While this meta-analysis was carried out in accordance with the Cochrane guideline (Deeks et al., 2013), there are limitations. First, the diagnostic criteria for ASD have changed from Diagnostic and Statistical Manual of Mental Disorder (DSM) IV to DSM-5 (American Psychiatric Association, 2013). Validation studies reported that children with severe ASD remain on the spectrum but some children with milder symptoms or those without repetitive behaviour do not (Huerta, Bishop, Duncan, Hus, & Lord, 2012; McPartland, Reichow, & Volkmar, 2012). In turn, the reported PPV and sensitivity could be lower using the DSM-5 criteria as children with milder symptoms who screened positive may be less likely to be ultimately diagnosed with ASD. Although there is a general consensus on the criteria for ASD (e.g. in DSM), the reference standard of clinical diagnosis was not uniformly applied across all studies due to differences in the how diagnostic assessment was carried out and the type of clinicians involved. In turn, this could contribute to heterogeneity between studies. This, however, unlikely biased study findings as the methods of assessment used in most studies could be considered reflective of clinical practise and variations in how the “gold standard” is applied could be considered inherent component of psychiatry.

Another limitation is that only studies that primarily used the English version of the M-CHAT were included. Although the M-CHAT has been validated in other languages and countries (Canal-Bedia et al., 2011; Kamio et al., 2014; Kara et al., 2014), the heterogeneity across these studies was too great to be pooled in a meta-analysis. Comparison of its performance across ethnocultural groups is necessary to ensure that it can correctly identify children with ASD across the general population. Since there were only two validation studies using the M-CHAT R/F at the time of this meta-analysis, the accuracy of the revised version, though often used clinically, could not be summarized. Lastly, only one reviewer screened the retrieved articles for inclusion.

2.3.5.2 Conclusion

This meta-analysis provided quantitative evidence that the M-CHAT performs with low-to-moderate sensitivity and specificity in children with developmental concerns. Although the M-CHAT was designed to be used in both low- and high-risk children, there was a lack of evidence supporting its use in the former group. Identified heterogeneity in accuracy measures emphasize that clinicians should account for the age and sex of the child when interpreting the M-CHAT scores. The existence of developmental concerns should also be considered when deciding to use the M-CHAT, as the PPV was much higher in high-risk compared to low-risk children. A low pooled specificity in high-risk children suggests that if it was used on a population level, high proportion of children without ASD would be referred for diagnostic assessment due to the low specificity of the M-CHAT, which would decrease the efficiency of the diagnostic pathway. Rather, standardized screening tools, such as the M-CHAT, may be used in children with developmental concerns by primary care physicians to better guide subsequent referral.

Key Points

- the M-CHAT performs with low-to-moderate accuracy in identifying ASD among children with developmental concerns
- clinicians should consider a child's age, sex and presenting developmental concerns when deciding to use the M-CHAT and interpreting its results

Table 2.1 Characteristics of the 13 studies included in the meta-analysis.

Study Characteristics				Study Participants				ASD diagnosis	
Author	Year	Location	Study Design	High or low risk	N	No. Male (%)	Age at Screening (months) mean (std)	Diagnostic Criteria	ASD Prevalence
Charman	2015	UK	Prospective	High	120	100 (83%)	35.3 (8.3)	ICD-10	45.80%
Cogan-Ferchalk	2013	US	Retrospective	High	222	179 (81%)	Median: 2 years	Based on education classification	49.10%
Eaves	2006	Canada	Prospective	High	82	68 (81%) ^b	37.2 (6.4) ^b	DSM-IV	65.90%
Fessenden	2013	US	Prospective	High	80	56 (70%)	26.8	ADOS module 1	47.50%
Goodwin	2010	Canada	Retrospective	High	148	115 (78%)	40.8 (14.3)	Not reported	57.40%
Koh	2014	Singapore	Retrospective	High	580	435 (75%)	35 (8)	DSM-IV TR	34.10%
Ringwood	2010	US	Retrospective	High	303	241 (80%)	27.6 (5.2)	DSM-IV	38.00%
Salisbury	2016	US	Retrospective	High	479	379 (77%)	30.8 (4.1)	Not specified	60.80%
Snow	2008	US	Prospective	High	56	44 (79%)	34.9 (8.7)	DSM-IV	65.90%
Villalobos	2011	US	Prospective	High	142	Not reported	Range: 14-31	DSM-IV TR	37.50%
Robin	2008	US	Prospective	Low	4797	2384 (50%)	20.9 (3.1)	DSM-IV	34.40%
Kleinman	2008	US	Prospective	Mixed	3793	2003 (53%)	21.0 (3.4)	DSM-IV	3.60%
Robin	2001	US	Prospective	Mixed	1293	693 (54%)	Range: 18-30	DSM-IV	3.00%

Values in table reflect study participants used to calculate accuracy measures in each paper, unless otherwise stated. The values may differ from sample characteristics reported in the original studies as some did not include all children in the final analysis. Additional information was provided by corresponding authors.

^a High-risk was defined as studies with children with any identified developmental concerns, low-risk refers to studies with children without identified developmental concerns, mixed refers to studies with both high-risk and low-risk children.

^b Of the entire sample (i.e. with children excluded from calculation of accuracy measures).

ADOS: Autism Diagnostic Observation Schedule; ASD: autism spectrum disorder; DSM: Diagnostic and Statistical Manual of Mental Disorder; ICD: International Classification of Diseases.

Table 2.2 M-CHAT results of the 13 studies included in the meta-analysis.

Author	Year	Sensitivity	Specificity	Positive Predictive Value
Charman	2015	0.80	0.40	0.53
Cogan-Ferchalk	2013	0.64	0.60	0.61
Eaves	2006	Critical6: 0.77 Total23: 0.92	Critical6: 0.43 Total23: 0.27	Critical6: 0.70 Total23: 0.69
Fessenden	2013	0.89	0.62	0.68
Goodwin	2010	0.88	0.51	0.71
Koh	2014	0.79	0.67	0.55
Ringwood	2010	0.96	0.35	0.53
Salisbury	2016	0.74	0.57	0.73
Snow	2008	Critical6: 0.70 Total23: 0.88	Critical6: 0.38 Total23: 0.38	Critical6: 0.79 Total23: 0.83
Villalobos	2011	not calculated	not calculated	0.06
Robin	2008	not calculated	not calculated	0.06
Kleinman	2008	not calculated	not calculated	0.36
Robin	2001	not calculated	not calculated	0.30

Values may differ from the original papers due to differences in reporting (e.g. disaggregation by scoring method, age group, etc.).

Additional information was provided by corresponding authors.

Critical6: positive screen defined as failing more than two of the six critical items. Total23: positive screen defined as failing more than three items of the 23 items. Both scoring algorithms were used for all other studies (i.e. a positive screen defined as failing more than two of the critical six items and/or more than three of any 23 items).

Table 2.3 Quality assessment of studies included in the meta-analysis.

Author (Year)	<u>Risk of Bias</u>				<u>Concerns of Applicability</u>		
	Participant Selection	Index Test	Reference Standard	Flow and Timing	Participant Selection	Index Test	Reference Standard
Charman (2015)	●	●	●	●	●	●	●
Cogan-Ferchalk (2013)	●	●	●	●	●	●	●
Eaves (2006)	●	●	●	●	●	●	●
Fessenden (2013)	●	●	●	●	●	●	●
Goodwin (2010)	●	●	●	●	●	●	●
Kleinman (2008)	●	●	●	●	●	●	●
Koh (2014)	●	●	●	●	●	●	●
Ringwood (2010)	●	●	●	●	●	●	●
Robin (2001)	●	●	●	●	●	●	●
Robin (2008)	●	●	●	●	●	●	●
Salisbury (2016)	●	●	●	●	●	●	●
Snow (2008)	●	●	●	●	●	●	●
Villalobo (2011)	●	●	●	●	●	●	●

●: low concern. ●: high concern.

Figure 0.1 PRISMA flow diagram of literature search.

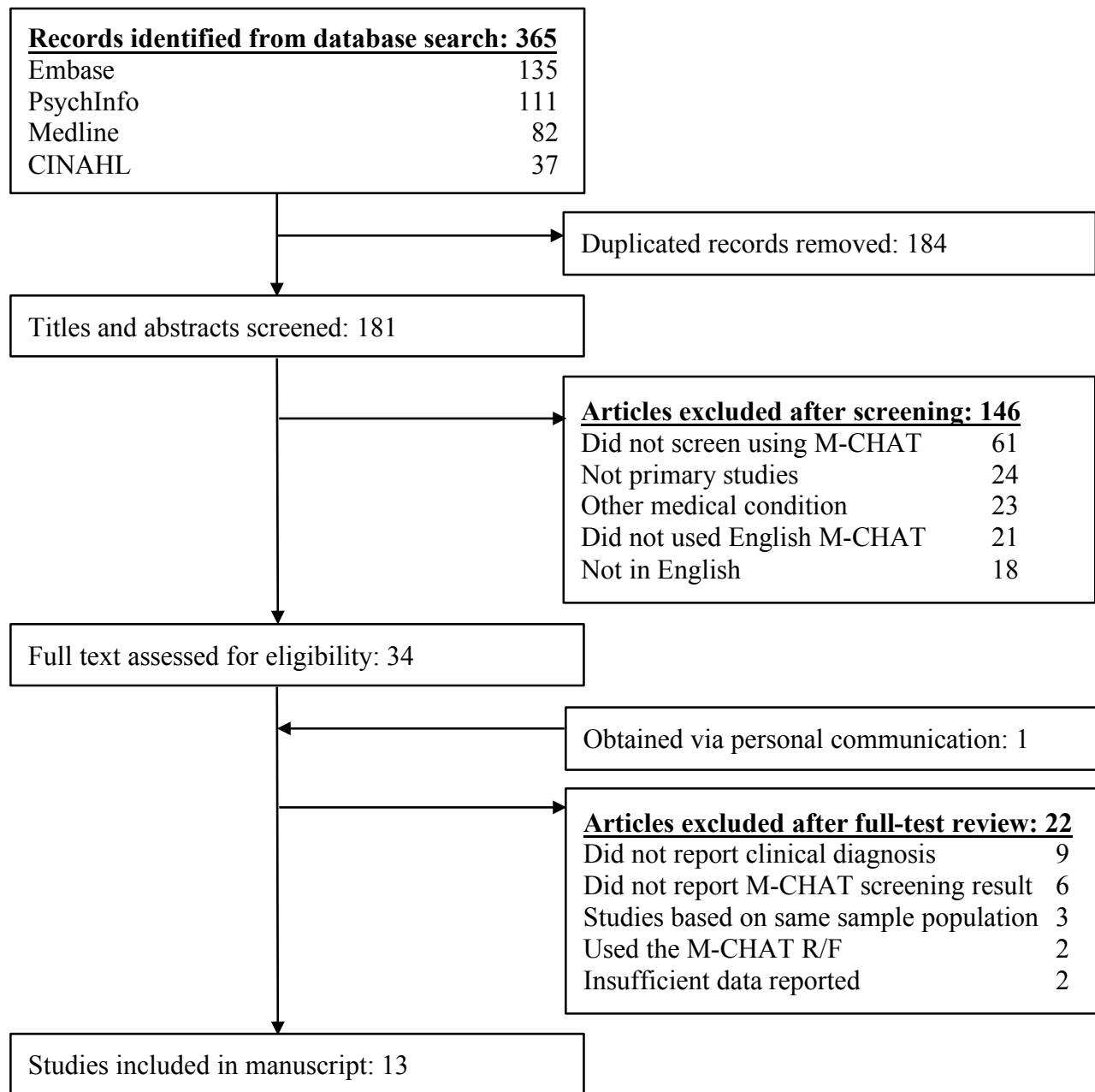
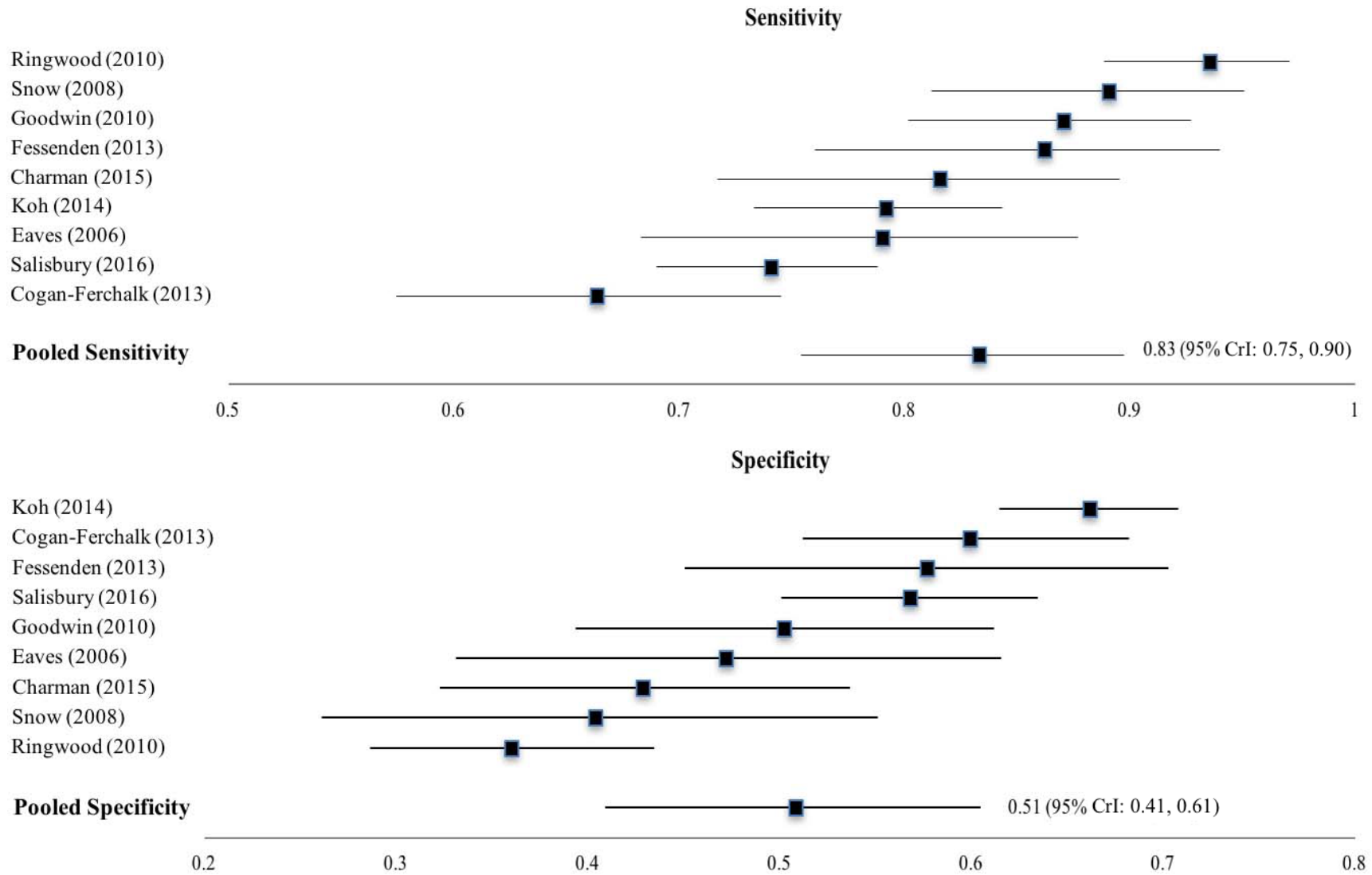


Figure 0.2 Forest plots of sensitivity and specificity.



Appendix 0.1 Questions on the QUADAS-2 modified for this study.

Part I: Participant Selection

1. How were participants recruited?
2. What was the source of the participants?
3. Were exclusion and inclusion criteria reported?

Part II: Index Test

1. What was the index test?
2. Was the administrator blinded to the child's clinical diagnosis?
3. Was the results interpreted without knowledge of the child's clinical diagnosis?
4. Was the scoring method pre-specified?

Part III: Reference Standard

1. Was the ASD diagnostic criteria reported?
2. Was a standardized tool used (e.g. ADOS, ADI-R)?
3. Was there a clinical assessment by a psychologist, psychiatrist and/or developmental pediatrician?
4. Were children likely to be correctly classified as ASD?
5. Were all participants diagnosed using the same method?
6. Was the reference standard implemented and interpreted without knowledge of the index test result?

Part IV: Flow and Timing

1. Was there an appropriate time between M-CHAT screening and clinical diagnosis?
2. Were the number of participant at each stage reported?
3. Were the reasons for non-participant at each stage given?
4. Were all patients included in the analysis?

Appendix 0.2 Results of the meta-regression models.

	Sensitivity	Specificity	Positive Predictive Value
Age			
At 24 months	0.55 (0.02, 0.84)	0.45 (<0.01, 0.71)	0.56 (0.48, 0.64)
At 30 months	0.69 (0.19, 0.86)	0.46 (0.03, 0.64)	0.55 (0.50, 0.61)
Male			
75% male	0.82 (0.73, 0.89)	0.53 (0.43, 0.63)	0.58 (0.52, 0.64)
100% male	0.84 (0.73, 0.92)	0.51 (0.36, 0.66)	0.46 (0.14, 0.78)
Study Design			
Retrospective	0.81 (0.70, 0.92)	0.55 (0.41, 0.67)	0.54 (0.37, 0.71)
Prospective	0.81 (0.67, 0.91)	0.48 (0.33, 0.63)	0.52 (0.37, 0.66)

Separate meta-regression models were constructed for each covariate. Models for sensitivity and specificity included the nine studies for which these measures could be calculated. Meta-regressions using PPV as the outcome included the twelve studies with high-risk children in their study sample, except one (Villalobos, 2011) was excluded from the meta-regression for gender distribution and two (Robins et al., 2001; Villalobos, 2011) from the meta-regression for age due to lack of appropriate data.

3. COST-EFFECTIVENESS OF UNIVERSAL OR HIGH-RISK SCREENING COMPARED TO SURVEILLANCE IN AUTISM SPECTRUM DISORDER

3.1 Preface

Using the pooled accuracy information from Chapter 2, this CEA compared the incremental costs and effects of universal or high-risk screening to surveillance monitoring in ASD.

The CEA was carried out using a DES model. Originated from operational research, DES is an event-based modelling technique that allows for tracking of individuals along a network of pathways. DES can estimate patient flow (e.g. wait time, resource use) in response to changes in system parameters (e.g. clinician availability, patients' demand for services). In turn, DES is able to evaluate the hypothesis that more vigorous screening could lead to earlier diagnosis and treatment initiation. The application of DES is particularly appropriate in ASD as it models entities (i.e. individual children) interacting with elements in a pathway (i.e. screening, diagnostics assessment) based on the entities' attributes (i.e. clinical characteristics, developmental trajectory). Attributes can be constant, time-varying or change in response to events, the combination of which would capture the complexity of clinical and developmental changes in young children.

The rest of this chapter is the manuscript which will be submitted for publication.

3.2. Manuscript #2

3.2.1 Abstract

Importance: The American Academy of Pediatric recommends that all children undergo screening for autism spectrum disorder (ASD) at 18 and 24 months. There is no direct evidence that active screening results in improved outcomes compared to surveillance monitoring.

Objective: To estimate the cost-effectiveness of universal or high-risk screening for ASD at 18 and 24 months compared to surveillance monitoring.

Design: A discrete event simulation model replicated the clinical pathway in ASD from birth to age 6 years to estimate the costs, in public payer and societal perspectives, and outcomes under three screening strategies. Costs and outcomes were discounted at 3%. Model parameters were

estimated from published literature, a large prospective cohort study and data from Statistics Canada.

Setting: Children in Ontario, Canada.

Participant: A cohort of children born within one year in Ontario, Canada. Each child was assigned clinical and developmental characteristics which influenced their pathway of care.

Main outcome and measures: Incremental cost per child correctly diagnosed with ASD before 36 months, and per child correctly diagnosed and started ASD treatment before 48 months.

Results: For high-risk screening, incremental cost was \$1196/child diagnosed before 36 months and \$1856/child initiated treatment before 48 months in the societal perspective. The incremental costs for the same outcomes were \$71213/child and \$138676/child for universal screening. Results were sensitive to changes in wait times for diagnostic assessment or ASD intervention and in the accuracy of the screening tool.

Conclusion and relevance: Universal screening at 18 and 24 months would greatly burden the healthcare system by increasing the demand of ASD diagnostic services and healthcare expenditure. The number of children referred for diagnostic assessment was 6-fold higher compared to surveillance, which would further delay access for children who require in-depth behavioural or psychiatric evaluation, ASD-related or not. Limiting standardized ASD screening to children considered at heightened risk of ASD could be a cost-effective strategy. As the goal is to promote early access to ASD intervention, reducing or eliminating wait times for ASD intervention could have greater impact compared to more vigorous screening.

3.2.2 Introduction

The effectiveness of universal screening to detect early signs of autism spectrum disorder (ASD) has been widely debated (Al-Qabandi et al., 2011; G. Dawson, 2016; Fein, 2016; Mandell & Mandy, 2015; Pierce et al., 2016; Powell, 2016; Robins et al., 2016; Silverstein & Radesky, 2016; Veenstra-VanderWeele & McGuire, 2016). The American Academy of Pediatrics (AAP) recommends that all children be screened with an ASD-specific tool, such as the Modified Checklist for Autism in Toddlers (M-CHAT) (Robins et al., 2001), at 18 and 24 months, a period when rapid neurodevelopment occurs (Johnson et al., 2007). This is in contrast to a current clinical practise often referred to as “surveillance”, where clinicians continuously monitoring children for signs of abnormalities over the course of development (Filipek et al., 2000;

Nachshen, Garcin, Moxness, Tremblay, Hutchinson, et al., 2008; Volkmar et al., 2014). Structured screening early in life could potentially identify initial presentations of ASD, such as atypical social and communication development that occurs as early as 12 months (Ozonoff et al., 2008; Zwaigenbaum, Bauman, Fein, et al., 2015). Moreover, ASD could be accurately diagnosed in some children prior to age 2 years (Chawarska, Klin, Paul, & Volkmar, 2007; Guthrie, Swineford, Nottke, & Wetherby, 2013) and improvement in adaptive behaviour, social skills and IQ have been reported in children with ASD who received behavioural and/or developmental intervention prior to age 3 years (Zwaigenbaum, Bauman, Choueiri, et al., 2015). Although there is strong theoretical rationale for the AAP recommendation, there is no direct evidence that universal screening can lead to earlier access to treatment or improved ASD outcomes over time (Siu et al., 2016).

Given the moderate accuracy of most published ASD screening tools and the low prevalence of the condition in the general population (Centers for Disease Control and Prevention, 2016; Johnson et al., 2007; Siu et al., 2016; Zwaigenbaum, Bauman, Fein, et al., 2015), one consequence of universal screening is that many children without ASD would be unnecessarily referred for additional investigations. The wait time for diagnostic assessment is already high in Ontario (median 26 weeks) (Penner, 2016) and the additional children from false positive ASD screening would further delay access for children who require in-depth evaluation. As ASD diagnostic assessment is a lengthy process that can involve multiple clinicians, unnecessary assessment can greatly increase healthcare expenditures and productivity lost.

A potentially more efficient strategy could be active screening targeted towards subgroups known to be at heightened risk or to be under-diagnosed for ASD. For example, children with a first-degree family member diagnosed with ASD are considered to be at heightened risk for ASD given high familial recurrence rate (Grønberg et al., 2013; Ozonoff et al., 2011), along with babies of preterm birth or have specific genetic conditions (Limperopoulos et al., 2008; Richards et al., 2015). Since they typically require closer monitoring, they might benefit from ASD-specific screening to detect early risk markers that are often unrecognized.

The objective of this study was to compare the incremental cost-effectiveness of universal or high-risk screening to surveillance in ASD from the provincial government and societal perspectives.

3.2.3 Methods

3.2.3.1 Study Design

A discrete event simulation model was used to predict the costs and consequences of a heterogeneous cohort of children born within one year in Ontario, Canada in each of the three ASD screening scenarios. The model was calibrated to reflect the epidemiology and clinical manifestation of ASD in children across Ontario. Each hypothetical child was randomly assigned a developmental trajectory and a set of clinical characteristics, summarized in Table 3.1, that influenced their pathway of care. The developmental trajectory was described by two time-varying attributes: attainment of age-specific developmental milestones as defined by the Centers of Disease Control and Prevention (CDC) (Centers for Disease Control and Prevention, 2015), and presence of severely low adaptive functioning as measured on the Vineland Adaptive Behavior Scales (Sparrow et al., 2005). Clinical characteristics included static attributes such as sex of the child, latent ASD diagnosis and whether the child was considered high-risk. High- or recurrent risk children were defined as children with one or more full sibling diagnosed with ASD. The probability of being classified as high-risk was based on the joint distribution of latent ASD diagnosis, recurrent risk and the probability of having an older sibling. The values were determined by random sampling from distributions estimated from published literature, information from Statistics Canada or a prospective cohort study, the Infant Sibling Study (Zwaigenbaum et al., 2012). The same cohort of children was then replicated and each cohort underwent one of the three screening strategies. The time horizon of the model was from birth to age 6 years. All costs and outcomes beyond one year were discounted at 3%.

3.2.3.2 Screening Strategies

Three screening strategies were compared. The reference approach was the current clinical practice in North America, surveillance monitoring for potential signs of developmental delay for all children, hereon referred to as surveillance. The two comparator screening strategies were

1) screening using the Modified Checklist for Autism in Toddlers (M-CHAT) (Robins et al., 2001) for all children, hereon referred to as universal screening, and 2) M-CHAT screening for high risk children and surveillance for low-risk children, hereon referred to as high-risk screening. The different screening approaches occurred at the 18- and 24-month well-child visits only. All children, regardless of strategy, underwent surveillance monitoring for developmental delay at 36-, 48- and 60-month well-child visits. The frequency of screening followed the well-child visit schedule recommended by the AAP (American Academy of Pediatrics, 2015).

A child was screened positive by surveillance if they did not attain the age-appropriate developmental milestones or if they had severely low adaptive functioning. A positive screen on the M-CHAT was based on the joint probability of having an observed ASD status and the accuracy of the M-CHAT (Table 3.2). The M-CHAT was selected as the screening tool for this model as it is a commonly used clinical tool that has been validated in multiple study populations and performs with moderate accuracy (Siu et al., 2016). Although there is a revised version of the M-CHAT (M-CHAT R/F) (Robins et al., 2014), there is no published information on its specificity and it has not been widely validated.

3.2.3.3 Outcomes

The two outcomes of this study were 1) the number of children correctly diagnosed with ASD before 36 months and 2) the number of children with a correct ASD diagnosis and initiated ASD intervention prior to 48 months. These two outcomes were selected to be consistent with recommendations made to the Ontario Ministry of Youth and Child Services regarding ASD service provision (Auditor General of Ontario, 2015). ASD intervention included government funded generic applied behavior analysis (ABA)-based therapy and early intensive behavioural intervention (EIBI). This model did not consider private ASD services because there is no systematic documentation on the proportion of families who seek private care and on how using private services alters the wait time and consumption of publicly funded ASD services.

3.2.3.4 Costing

Cost items were considered from the societal and public payer (i.e. provincial government) perspectives. Details on the cost items, valuation and sources of information are summarized in Table 3.3. The cost of each clinical visit for screening and ASD diagnosis was based on the physician fee schedules or was obtained via personal communication with professional associations. The duration and cost of ASD diagnostic assessment varied by the type of clinician(s) involved. This model assumed that ASD assessment was conducted by a psychologist, psychiatrist, developmental pediatrician or a multidisciplinary team (consisting of a developmental pediatrician, psychological, speech language pathologist and occupational therapist) and the probability of each clinician type is based on a recent national survey (Penner, 2016). For the societal perspective, parental time lost from accompanying their child to all screening and diagnostic assessments were valued using the sex-specific hourly wage (Statistics Canada, 2015b). This model assumed that the primary caregiver was female and only she accompanied the child to all medical visits. All cost items were expressed in 2016 Canadian dollars.

3.2.3.5 Model Pathway

Figure 3.1 describes the pathway of care for the three screening strategies. Each hypothetical child entered the model at the Ontario birth rate and the model generated children for one year. Each child was assigned a set of attributes as described above and waited for the first screening assessment which took place at 18 months.

For all three strategies, a child screened negative would wait for the next screening assessment based the well-child visit schedule. Children who screened positive were referred for ASD diagnostic assessment. The wait time of which was estimated from the current wait time for ASD diagnostic assessment in Canada (Table 3.4) and the number of children referred for diagnosis in the model. It was modelled as a log-normal distribution to reflect the skewness of current wait time distribution. The accuracy of ASD diagnostic assessment (Table 3.2) was based on published literature (Huerta et al., 2012; McPartland et al., 2012) and the model assumed that the probabilities did not vary by clinician type. Children diagnosed with ASD, both true positive and

false positive, were referred for ASD intervention. Children who were not diagnosed returned to screening based on the well-child visit schedule.

Children diagnosed with ASD were referred to EIBI if they were severely impaired based on their adaptive functioning status and to generic ABA-based intervention otherwise. For children with false positive ASD diagnosis, they remained on the wait list and potentially underwent treatment until their developmental trajectory or adaptive functioning improved to the age-appropriate range. After this, they returned to scheduled well-child visits. This is assumed to be reflective of clinical reality as clinicians cannot know the true latent ASD diagnosis of a child and are unlikely remove an ASD diagnosis unless the child shows significant improvement.

The wait time for generic ABA-based therapy and EIBI were estimated from the current wait times in Ontario, the number of available spots and the length of each type of intervention (Table 3.4). A child exited the model if they initiated either form of intervention or reached age 6 years. The discrete event simulation model was built using MatLab 2017a (MATLAB, 2017).

3.2.3.6 Statistical Analysis

The predicted outcomes and costs for the two alternate screening strategies were compared to standard care, surveillance, using incremental analysis and were summarized as incremental cost-effectiveness ratios (ICERs). The ICERs were expressed as the incremental cost per additional child correctly diagnosed with ASD by 36 months and the incremental cost per additional child correctly diagnosed and initiated ASD intervention by 48 months. The mean costs and outcomes of the three strategies were also compared graphically using a cost-effectiveness frontier.

Uncertainties in the ICERs were assessed using a non-parametric method. Bootstrapped sampling was used to simulate 1000 cohorts. For each simulated cohort, the incremental costs and outcomes for each comparator relative to surveillance were estimated and plotted on a cost-effectiveness plane. A cost-effectiveness acceptability curve (CEAC) was also constructed by plotting the proportion of ICERs that were below the willingness-to-pay threshold, using a series of thresholds from \$0-300000.

One-way deterministic sensitivity analyses were used to quantify uncertainties in model parameters. The parameters assessed were selected based on their impact on ASD epidemiology and influence on the pathway of care. The parameters (Table 3.5) included factors that influenced the demand for ASD services (i.e. prevalence of ASD, recurrent risk of ASD, accuracy of the M-CHAT), the cost of diagnostic assessment (i.e. type of clinician(s) administering the ASD diagnostic assessment), the efficiency of the diagnostic pathway (i.e. wait times for ASD diagnostic assessment or ASD intervention) and discount rate.

3.2.4 Results

A cohort of 139789 children, 2065 (1.5%) of whom had a latent ASD diagnosis, was generated and put through the model. The proportion of children with ASD correctly diagnosed before 36 months was 10% for surveillance, 15% for high-risk screening and 31% for universal screening. The mean costs and outcomes for the three screening strategies are summarized in Table 3.6 for the public payer perspective and in Table 3.7 for the societal perspective. Compared to surveillance, both outcomes were 2 times higher in high-risk screening and 3 times higher in universal screening.

In the public payer perspective, the mean costs per 1000 children were comparable between surveillance (\$292702±163973) and high-risk screening (\$293538±166356), but was almost 2 times higher for universal screening (\$491413±237071). The cost-effectiveness frontiers (Figure 3.2, top row) do not suggest any strategy was dominated for either outcome. Compared to surveillance, the ICER for high-risk screening was \$1101 per additional child diagnosed before age 3 years and \$1709 per additional child initiated treatment before age 4 years. Using the same reference strategy, the ICER for universal screening was \$65002 children per additional child diagnosed before 36 months and \$126576 per additional child initiated treatment before 48 months.

The same pattern in costs was observed in the societal perspective and the cost-effectiveness frontiers (Figure 3.2, bottom row) also do not suggest any strategy was dominated. Compared to surveillance, the ICER for high-risk screening was \$1196 per additional child diagnosed before

age 3 years and \$1856 per additional child that initiated treatment before age 4 years. Using the same reference strategy, the ICER for universal screening was \$71213 children per additional child diagnosed before 36 months and \$138676 per additional child that initiated treatment before 48 months.

The bootstrap simulation (Figure 3.3) shows high uncertainty in all ICERs. For high-risk screening compared to surveillance, the ICERs were divided across the four quadrants of the cost-effectiveness plane. This indicates that neither strategy was definitively more costly or more effective. The incremental effect per 1000 children was 0.75 (95% CI: -1.92, 3.43) for number of children diagnosed before 36 months and 0.49 (95% CI: -1.88, 2.86) for initiated treatment before 48 months. The incremental cost per 1000 children was \$802 (95% CI: -6405, 8009) in public payer and \$872 (95% CI: -7165, 8909) in societal perspective. For universal screening compared to surveillance, 3% and 8% of the iterations were less effective for diagnosis and treatment, respectively, but all iterations were more expensive. The incremental effect was 3.14 (95% CI: -1.01, 7.29) for diagnosed before age 3 years and 1.64 (95% CI: -1.49, 4.77) for initiated treatment before age 4 years. The incremental cost per 1000 children was \$198566 (95% CI: 182783, 214349) in public payer and \$215528 (95% CI: 200053, 235003) in societal perspective.

3.2.4.1 Cost-effectiveness Acceptability Curve

Figure 3.4 (top) shows the CEAC for high-risk screening compared to surveillance. Examining a willingness-to-pay threshold of \$0 reveals that approximately 40% of the simulated cohorts had lower cost and greater effectiveness (dominance), in both public payer and societal perspectives, for high-risk screening than for surveillance. Starting at the threshold of \$18000, the proportion of ICERs below the threshold plateaued at 70% for diagnosed before 36 months and at 65% for initiated treatment before 48 months. The lower proportion of iterations considered cost-effective for treatment initiation compared to diagnosis was due to a smaller incremental effect for this outcome.

The CEAC for universal screening compared to surveillance (Figure 3.4, bottom) shows much wider range of acceptability. Half of the iterations were below the threshold of \$60000 for

diagnosed before 36 months and of \$120000 for initiated treatment before 48 months. The proportion of ICERs considered cost-effective began to level at a threshold of \$300000, with 91% below the threshold for diagnosed before 36 months and 78% for initiated treatment before 48 months. The difference in the proportion of iterations below the threshold between the two outcomes is due to differences in incremental effectiveness.

3.2.4.2 Sensitivity Analyses

Figures 3.5 and 3.6 show the results of the one-way deterministic sensitivity analyses under the public payer and the societal perspectives, respectively. The ICERs for high-risk screening compared to surveillance using either outcome were most sensitive to changes in the accuracy of the M-CHAT. If the sensitivity and specificity of the M-CHAT were both 50% (i.e. same as chance), the ICERs increased by 48-52% for diagnosed before 36 months and by 118-124% for initiated treatment before 48 months. On the other hand, if sensitivity and specificity were higher at 80%, the ICERs decreased by 56% for diagnosed before 36 months and by 58% for initiated treatment before 48 months. The ICERs for universal screening compared to surveillance were not sensitive to a decrease in accuracy of the M-CHAT for either outcome (increase of 7% if sensitivity and specificity were at 50%), but they dropped by approximately 30% if sensitivity and specificity increased to 80%.

The incremental costs per additional child that initiated treatment before 48 months were sensitive to changes in wait times for ASD diagnostic assessment, generic ABA-based therapy and EIBI.

If the wait time for EIBI was eliminated, the number of children that initiated treatment before 48 months increased by 80-100% in all three strategies. Compared to surveillance, the incremental cost per additional child that initiated treatment before age 48 months decreased by 51% for high-risk screening and 83% for universal screening when EIBI wait time was eliminated.

Eliminating the wait time for generic ABA-based therapy did not have an impact on the ICERs for high-risk screening but decreased the ICERs for universal screening by 17%. On the other

hand, if wait time increased by 30% from the base case scenario, the ICERs for initiated treatment before 48 months increased by 54% for high-risk screening and by 87% for universal screening.

While the ICERs for both comparator strategies were sensitive to changes in diagnostic wait time, the impact was more pronounced for universal screening because more children are referred for diagnostic assessment. If the mean wait time shifts towards the high-end of the distribution in the base case scenario, the ICER for universal screening compared to surveillance increased by 34-38% for diagnosed by 36 months and by 130-136% for treatment initiation by 48 months. The ICERs for high-risk screening compared to surveillance increased by 34-40% for either outcome.

Increased ASD prevalence resulted in lower ICERs for both comparator strategies. This could be attributed to a decrease in the proportion of children without a latent ASD diagnosis among those referred for diagnostic assessment as prevalence increased. As the prevalence increased to 6/1000 boys and 3/1000 girls, the ICER per additional child initiated treatment before 48 months decreased by 12% for high-risk screening and 5% for universal screening.

Type of clinician(s) who administered the ASD diagnostic assessment was also influential (Tables 3.8 and 3.9). If all diagnostic assessments were carried out by a developmental paediatrician or by a psychologist, the ICERs per additional child diagnosed before 36 months decreased by more than half for either alternate screening strategy compared to reference case scenario, which reflected the current mixture of clinicians that carry out diagnostic assessment. On the other hand, the ICERs were much higher if all assessments were carried out by a multidisciplinary team. As diagnostic assessment by a multidisciplinary team also took longer to complete, the difference in ICERs by clinician type was most pronounced under the societal perspective when parental time lost from accompanying their child to clinical visits was valued. Compared to the reference case scenario, the ICERs for universal screening relative to surveillance increased by 70-90% when a multidisciplinary team administered all assessments.

3.2.5 Discussion

The hypothesized effectiveness of universal screening using an ASD-specific tool for all children at 18 and 24 months has been largely based on evidence from cross-sectional studies or clinical experience (Dawson, 2016; Fein, 2016; Pierce et al., 2016; Powell, 2016; Robins et al., 2016; Silverstein & Radesky, 2016; Veenstra-VanderWeele & McGuire, 2016). This simulation study adds to the debate from an economic perspective and demonstrates that universal screening would greatly burden the healthcare system by heightening healthcare expenditures and demand for ASD diagnostic services. Although universal screening was 3 times more effective in correctly identifying children with ASD before age 3 years compared to surveillance (4.43±63.53 vs. 1.38±35.44 children with ASD per 1000 children), the number of children referred for diagnostic assessment was 6-fold higher. Increased referral would not only prolong wait time for diagnostic assessment, but also increase consumption of downstream intervention by children who do not have ASD or might not benefit from it. As most children were on wait lists beyond the time horizon of the model (i.e. at age 6 years) the estimated wait times were right-censored. Limiting active ASD screening to children considered at heightened risk for ASD (e.g. children with one or more full sibling diagnosed with ASD) could be a more cost-effective option at ICER of \$1100-1900 per additional child diagnosed before 36 months or initiated treatment before 48 months. Given the high cost of universal screening and inconclusive evidence on the efficacy of interventions for all individuals with ASD, it does not fulfill the criteria for population-based screening by Wilson and Jungner (Fletcher-Watson, McConnell, Manola, & McConachie, 2014; Oono, Honey, & McConachie, 2013; Reichow, Barton, Boyd, & Hume, 2012; Wilson & Jungner, 1968).

The cost-effectiveness of ASD screening compared to surveillance was dependent on the accuracy of the screening tool. Due to low specificity of the M-CHAT, a large proportion of children referred for ASD diagnostic assessment did not have ASD nor did they require additional behavioural assessment or treatment. Using an ASD screening tool that performs at 80% sensitivity and specificity significantly decreased the ICERs for either strategy, but the ICERs for universal screening remained high (\$43000-47000/child diagnosed by 36 months, \$87000-95000/child initiated treatment by 48 months).

The impact of high false positive referral due to low specificity of the M-CHAT was particularly noticeable for universal screening and the ICERs were highly sensitive to changes in diagnostic wait time. As wait time increased with additional children referred, the ICER per child initiating treatment before 48 months more than doubled when the wait time increased by 30% from base case assumption. Despite the high cost of universal screening, the bootstrap simulations (Figure 3.3) indicated that it was not always more effective compared to surveillance. Although there is no information on how much the government or society is willing pay for an additional child diagnosed or initiating treatment earlier, the CEACs (Figure 3.4) indicate that universal screening compared to surveillance was not cost-effective 10-30% of the time even at a willingness-to-pay threshold of \$300000 per additional child.

A targeted screening approach might be optimal, especially if ASD prevalence is higher. The ICERs for high-risk screening compared to surveillance decreased to below \$1700 as prevalence increased. However, the CEACs indicated that high-risk screening did not always lead to more children diagnosed or initiating treatment earlier compared to surveillance. The proportion of ICERs considered cost-effective plateaued starting at a willingness-to-pay threshold of \$18000 and 30-35% of the iterations cannot be considered cost-effective. This study defined high-risk children as those with one or more full sibling diagnosed with ASD, but another criterion to define the target subpopulation that are known to be at heightened risk for ASD or are under-diagnosed might be appropriate.

Due to the long wait time for EIBI in Ontario, few children initiated EIBI prior to 48 months. Eliminating wait time for EIBI resulted in an 80-100% increase in the number of children initiating treatment before 48 months in all three screening strategies, with the highest percentage increase in surveillance. Moreover, the number of children initiating treatment by 48 months from surveillance without wait time for generic ABA-based therapy was similar to the number from high-risk screening with the current wait time. A recent Ontario study estimated lifetime savings of CAD \$267000 and gains of 2.52 disability-free life years per individual if EIBI wait time was eliminated (Piccininni, Bisnaire, & Penner, 2017). Considering that the hypothesized benefits of universal screening are mediated through earlier access to ASD treatment, the present

study supports resource allocation towards reducing or eliminating wait time for ASD intervention rather than more vigorous screening.

The type of clinician who carried out ASD diagnostic assessment also had a large impact on the incremental costs. While this model assumed the referral was random, it is likely based on the child's presenting symptoms and availability of clinicians in the area. The purpose of ASD diagnostic assessment is not only to correctly classify the child's developmental concerns, but also to identify the strengths and weaknesses in a child's development and behaviour such that they are referred to an appropriate ASD intervention. Although diagnostic assessment by psychologists or developmental paediatricians only were the least expensive options, the assessment might not be able to identify all of the potential needs that ASD services should address, especially considering the wide spectrum of concerns presenting in children with ASD.

3.2.5.1 Limitations

Due to the complexity of the pathway to ASD diagnosis and determining developmental trajectories in children, the model made several assumptions which might limit the generalizability of the study findings. This model accounted for the large heterogeneity in the clinical presentation of ASD symptoms by modelling the children's developmental trajectories using a probabilistic approach and by allowing the trajectories to vary by sex, time and latent risk of ASD. However, the trajectories were described by two variables, attainment of age-appropriate developmental milestones and the presence of severely low adaptive functioning, which might not capture all clinical symptoms typically monitored over time in surveillance. Therefore, the number of children screened positive, and in turn costs and outcomes, could be underestimated for surveillance, which could have biased the ICERs in either direction.

The time horizon of this model was limited to birth to age 6 years. Children have additional avenues to receive an ASD diagnosis (e.g. through the education system) and this could confound measurement of the costs and effects of the screening strategies. Moreover, the benefits of early diagnosis and intervention initiation might not be apparent until later in the child's life. The high cost ASD over the lifetime is primarily driven by productivity lost and the cost of support services in adult years (Ganz, 2007; Motiwala, Gupta, Lilly, Ungar, & Coyte,

2006). If earlier treatment could revert the child's development to a more age-appropriate trajectory, thus improving daily functioning and independence, costs incurred over their lifetime would be greatly reduced. In turn, ICERs for more vigorous screening strategies would be lower. However, there is no published evidence that on the long-term treatment outcomes in children identified through screening.

Lastly, this model did not consider private ASD interventions because the amount of out-of-pocket expenditure could not be estimated without reliable data on resource use. Study findings indicate that reduction in wait times would have a large impact on the number of children that initiated treatment before 48 months. If a large proportion of parents of children with ASD are willing to pay out-of-pocket for ASD intervention, thus bypassing the long wait time for publicly funded interventions, the effectiveness of all three strategies would be underestimated. In addition, while cost to the government (i.e. public payer) would be decreased when parents pay out-of-pocket for intervention, the higher costs of private services would result in higher societal costs overall. Since reduction in wait time had a greater impact on the effectiveness of surveillance compared to the two alternate screening strategies, incremental effectiveness would decrease and the ICERs would be higher if private intervention was considered.

3.2.6 Conclusion

This study demonstrated that screening all children using an ASD-specific tool, such as the M-CHAT, would greatly burden the healthcare system by increasing healthcare expenditures and demand for ASD diagnostic services. Compared to surveillance, universal screening cannot be considered cost-effective or efficient, even if the screening tool performed at high sensitivity and specificity. A tailored screening approach targeting children at heightened risk for ASD could be a cost-effective option, especially if ASD prevalence is higher. Reduction in wait times for publicly funded ASD interventions, such as generic ABA-based therapy and EIBI, had the highest impact on the number of children starting treatment earlier. If the ultimate goal for children with ASD to receive treatment earlier, resources should be allocated toward reducing wait times for ASD diagnostic assessment and intervention rather than more vigorous ASD screening.

Table 3.1 Static and time-varying attributes assigned to children in discrete event simulation model.

Attributes	Distributions ¹ (mean, sd)			Source
Static				
Sex (Male)	0.51, 0.05			(Statistics Canada, 2015a)
Latent ASD	M: 0.02, 0.005 F: 0.01, 0.002			(Ouellette-Kuntz et al., 2014)(Centers for Disease Control and Prevention, 2016)
Recurrent risk	M: 0.26, 0.02 F: 0.09, 0.02			(Ozonoff et al., 2011)
Time-varying				
	18 months	24 months	36, 48, 60 months	
Severely low Vineland composite score	M+HR: 0.11, 0.04	M+HR: 0.08, 0.03	M+HR: 0.24, 0.26	(Sparrow et al., 2005; Zwaigenbaum et al., 2012)
	M+LR: 0.02, 0.01	M+LR: 0.02, 0.01	M+LR: 0.02, 0.01	
	F+HR: 0.08, 0.04	F+HR: 0.10, 0.04	F+HR: 0.26, 0.07	
	F+LR: 0.02, 0.01	F+LR: 0.02, 0.01	F+LR: 0.02, 0.01	
Delayed developmental milestone	M+HR: 0.08, 0.02	M+HR: 0.20, 0.04		(Hagan, Shaw, & Duncan, 2008; Zwaigenbaum et al., 2012)
	M+LR: 0.02, 0.01	M+LR: 0.05, 0.02		
	F+HR: 0.06, 0.04	F+HR: 0.18, 0.04		
	F+LR: 0.02, 0.01	F+LR: 0.05, 0.02		
Observed ASD status	M+HR: 0.18, 0.04	M+HR: 0.53, 0.06	same as latent ASD	(Zwaigenbaum et al., 2012)
	M+LR: 0.01, 0.01	M+LR: 0.01, 0.01	same as latent ASD	
	F+HR: 0.15, 0.06	F+HR: 0.55, 0.09	same as latent ASD	
	F+LR: 0.01, 0.01	F+LR: 0.01, 0.09	same as latent ASD	

¹Distributions were modelled as beta distributions in the discrete event simulation model.

ASD: autism spectrum disorder; F: female; HR: high-risk; LR: low-risk; M: male; sd: standard deviation.

Recurrent risk was defined as children with one or more full sibling diagnosed with ASD

Table 3.2 Input parameter for accuracy of ASD diagnostic assessment and M-CHAT screening in discrete event simulation model.

	Sensitivity	Specificity	Source
M-CHAT screening	N(0.8, 0.04)	N(0.5, 0.05)	(Yuen et al., 2017, <i>publication submitted for review</i>)
Diagnostic assessment	N(0.8, 0.05)	N(0.8, 0.007)	(Huerta et al., 2012; McPartland et al., 2012)

$N(\mu, \sigma)$ represents the Normal distribution where μ is the mean and σ is the standard deviation.

M-CHAT: Modified Checklist for Autism in Toddlers.

Table 3.3 Cost items and resource use for discrete event simulation model.

Cost Items	Distribution of Unit Cost (\$)	Source	Resource Use (Number of units)	Source
Screening				
Enhanced well-baby visit	N(51.8, 4.9) ¹	(Ministry of Health and Long-Term Care, 2015; Régie de l'assurance maladie du Québec, 2015)	At 18 months: 1	(Williams et al., 2011)
Paediatric assessment	N(52.2, 10.0) ²	(Ministry of Health and Long-Term Care, 2015; Régie de l'assurance maladie du Québec, 2015)	At 24/36/48/60 months: 1	(American Academy of Pediatrics, 2015)
Diagnostic assessment				
Psychiatrist	N(427.5,16.6) ³	(Ministry of Health and Long-Term Care, 2015; Régie de l'assurance maladie du Québec, 2015)	19%: 1	(Penner, 2016)
Psychologist	N(225.0,3.9)	(British Columbia Psychological Association, 2014; Ontario Psychological Association, 2015)	24%: 1	(Penner, 2016)
Paediatrician	Initial: N(138.5,14.0) ⁴ Follow-up: N(73.1,9.1)	(Ministry of Health and Long-Term Care, 2015; Régie de l'assurance maladie du Québec, 2015)	26%: Of which, 20%: 1 80%: 2	(Penner, 2016)
Multi-disciplinary team			31%: 1	(Penner, 2016)
Paediatrician	Initial: N(138.5,14.0) Follow-up: N(73.1,9.1)	(Ministry of Health and Long-Term Care, 2015; Régie de l'assurance maladie du Québec, 2015)	20%: 1 80%: 2	(Penner, 2016)
Psychologist	N(225.0,3.9)	(British Columbia Psychological Association, 2014; Ontario Psychological Association, 2015)	1	(Penner, 2016)
Speech Language therapist	N(190.0,9.7)	(Ontario Association of Speech-Language Pathologists and Audiologists, 2016)	1	(Penner, 2016)

Occupational therapist	N(60.0,3.1)	(Ontario Society of Occupational Therapist, email communication, Jan 2016)	1	(Penner, 2016)
Hourly wage for women	N(21.6,0.8)	(Statistics Canada, 2015b)	0.5-1 units for screening 1-2.5 units for diagnostic assessment	(Penner, 2016; Weiss, Whelan, McMorris, Carroll, & Canadian Autism Spectrum Disorders Alliance, 2014)

N(μ, σ) represents normal distribution with mean μ and standard deviation σ .¹ fee code: A262, 09127, ² fee codes: A002, 09127, ³ fee codes A667, 08935, ⁴ fee codes: A265, K123, 09165, 15164.

Table 3.4 Input parameters used to estimate wait times for diagnostic assessment and ASD intervention in discrete event simulation model.

	Value	Source
ASD diagnostic assessment wait time (minutes)	LogN(12.8, 0.15)	(Penner, 2016)
Generic ABA-based therapy		
Wait time (minutes)	N(564,480, 40,320)	(Auditor General of Ontario, 2015)
Length of intervention (minutes)	N(262,800, 20,160)	(Auditor General of Ontario, 2015)
Spots available	9400	(Auditor General of Ontario, 2015)
EIBI		
Wait time (minutes)	N(11,179,360, 40,320)	(Auditor General of Ontario, 2015)
Length of intervention (minutes)	N(1,314,000, 262,800)	(Auditor General of Ontario, 2015)
Spots available	1400	(Auditor General of Ontario, 2015)

$N(\mu, \sigma)$ represents the Normal distribution where μ is the mean and σ is the standard deviation.

$\text{LogN}(\mu, \sigma)$ represents the Log-normal distribution where μ is the location parameter and σ is the scale parameter.

ABA: applied behavior analysis; ASD: autism spectrum disorder; EIBI: early intensive behavioural intervention.

Table 3.5 Parameters and ranges included in the one-way deterministic sensitivity analysis.

Parameters	Ranges	Source
Recurrent risk of ASD	Male: 0.19-0.35 Female: 0.057-0.14	(Ozonoff et al., 2011)
Prevalence of ASD	Male: 22.9/1000- 24.3/1000 Female: 4.9/1000 -5.6/ 1000	(Centers for Disease Control and Prevention, 2016)
Wait times		
Diagnostic assessment	4-24 weeks	(Penner, 2016)
Generic ABA-based therapy	0-75 weeks	(Auditor General of Ontario, 2013, 2015)
EIBI	0-36 months	(Auditor General of Ontario, 2013, 2015)
Accuracy of M-CHAT	Sensitivity: 0.5-0.8 Specificity: 0.5-0.8	
Clinician for ASD diagnostic assessment		
Psychiatrists	100%	
Psychologist	100%	
Developmental paediatrician	100%	
Multi-disciplinary team	100%	
Discount rate	0-5%	

ABA: applied behavior analysis; ASD: autism spectrum disorder; EIBI: early intensive behavioural intervention; M-CHAT: Modified Checklist for Autism in Toddlers.

Table 3.6 Mean and incremental outcomes and costs in the public payer perspective.

	Surveillance	High-risk Screening	Universal Screening
Outcome (per 1000 children)			
Number of children diagnosed before 36 months			
Mean±sd	1.375±35.444	2.134±44.143	4.432±63.532
Incremental	ref	0.759	3.057
Number of children initiated treatment before 48 months			
Mean±sd	0.782±26.344	1.271±33.583	2.352±45.65
Incremental	ref	0.489	1.570
Cost - public payer (per 1000 children)			
Mean±sd	292702±163973	293538±166356	491413±237071
Incremental	ref	836	198712
Screening (mean; % total cost)	232354 (79%)	232059 (79%)	198809 (40%)
Diagnostic assessment (mean; % total cost)	60348 (21%)	61479 (21%)	292604 (60%)
ICER			
\$/child diagnosed before 36 months	ref	1101	65002
\$/child initiated treatment before 48 months	ref	1709	126576

ICER: incremental cost-effectiveness ratio; sd: standard deviation.

Table 3.7 Mean and incremental outcomes and costs in the societal perspective.

	Surveillance	High-risk Screening	Universal Screening
Outcome (per 1000 children)			
Number of children diagnosed before 36 months			
Mean±sd	1.375±35.444	2.134±44.143	4.432±63.532
Incremental	ref	0.759	3.057
Number of children initiated treatment before 48 months			
Mean±sd	0.782±26.344	1.271±33.583	2.352±45.65
Incremental	ref	0.489	1.570
Cost - societal (per 1000 children)			
Mean±sd	358067±181461	358975±184073	575768±263302
Incremental	ref	908	217701
Screening (mean; % total cost)	291140 (81%)	290789 (81%)	250592 (44%)
Diagnostic assessment (mean; % total cost)	66927 (19%)	68186 (19%)	325176 (56%)
ICER			
\$/child diagnosed before 36 months	ref	1196	71213
\$/child initiated treatment before 48 months	ref	1856	138676

ICER: incremental cost-effectiveness ratio; sd: standard deviation.

Table 3.8 Incremental cost-effectiveness ratios from one-way deterministic sensitivity analyses in the public payer perspective using one clinician type for all diagnostic assessment.

Clinician Type	High-risk Screening		Universal Screening	
	Diagnosed by 36 months	Initiated treatment by 48 months	Diagnosed by 36 months	Initiated treatment by 48 months
Psychiatrists	1348	2074	69550	140953
Psychologist	520	800	31614	64070
Developmental paediatrician	476	732	26351	53403
Multidisciplinary team	1864	2867	115327	233726

Table 3.9 Incremental cost-effectiveness ratios from one-way deterministic sensitivity analyses in the societal perspective using one clinician type for all diagnostic assessment.

Clinician Type	High-risk Screening		Universal Screening	
	Diagnosed by 36 months	Initiated treatment by 48 months	Diagnosed by 36 months	Initiated treatment by 48 months
Psychiatrists	1402	2156	73432	148820
Psychologist	529	814	33470	67831
Developmental paediatrician	530	815	29814	60421
Multidisciplinary team	2075	3192	129736	262928

Figure 0.1 Schematic diagram of the clinical pathway in the discrete event simulation model.

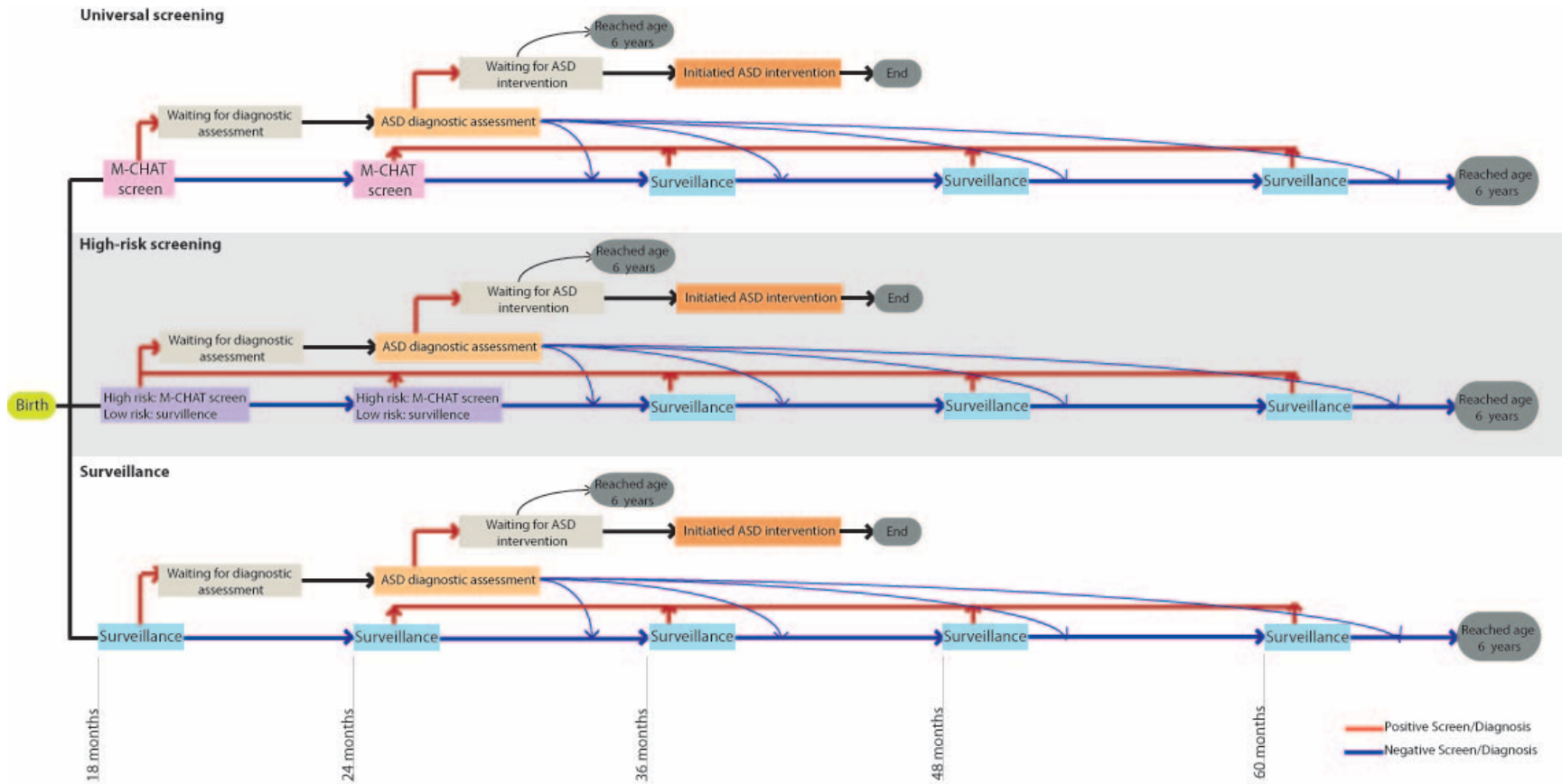


Figure 0.2 Cost-effectiveness frontiers for the three ASD screening strategies in the discrete event simulation model in the public payer (top row) and societal (bottom row) perspectives.

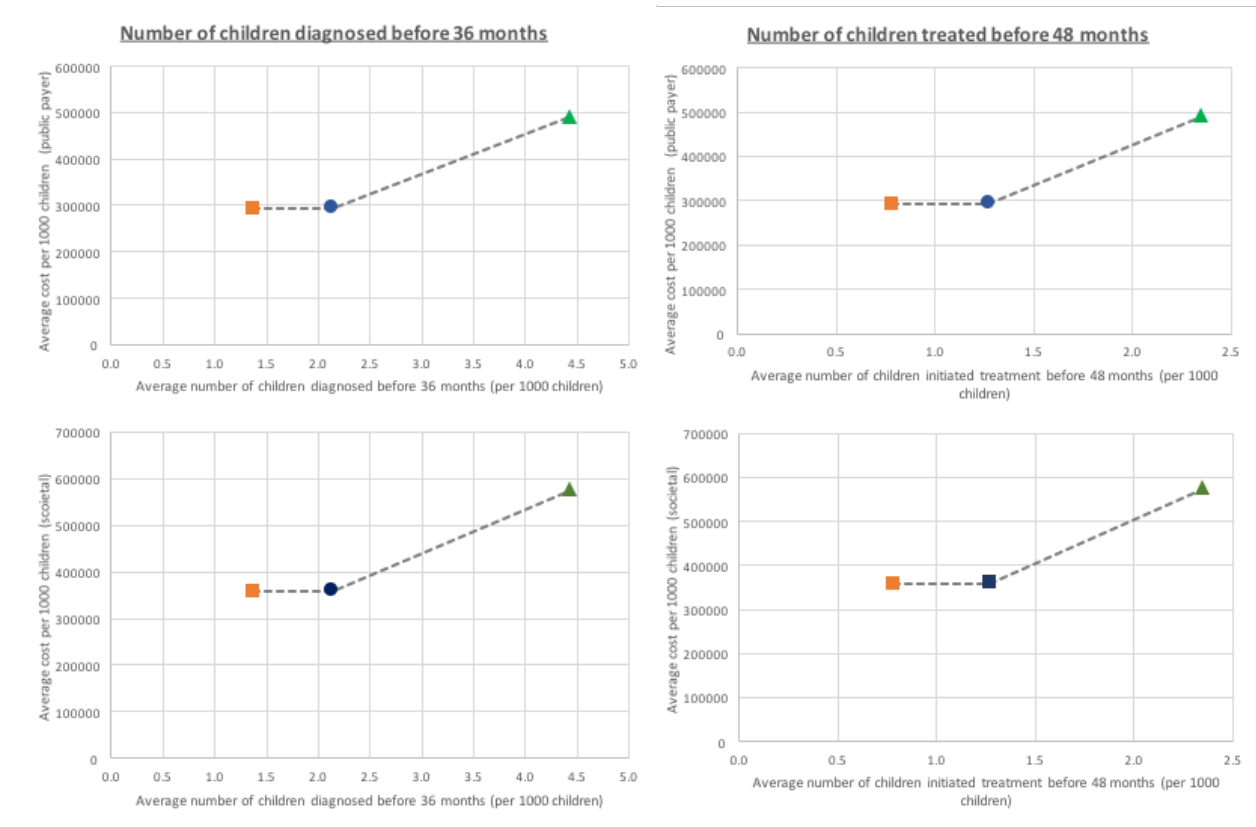


Figure 0.3 Incremental costs and effects from the bootstrap simulation for high-risk screening and universal screening compared to surveillance in public payer (top row) and societal (bottom row) perspectives.

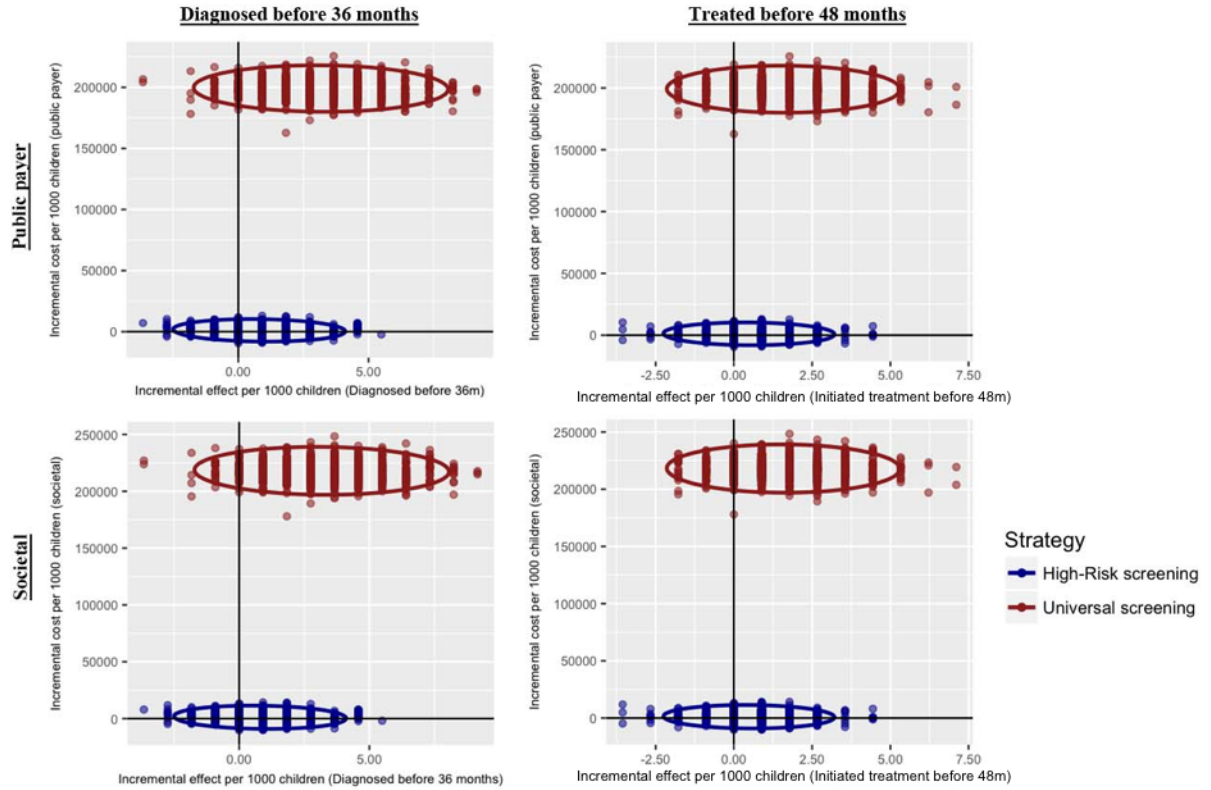


Figure 0.4 Cost-effectiveness acceptability curve for high-risk screening (top) and universal screening (bottom) compared to surveillance.

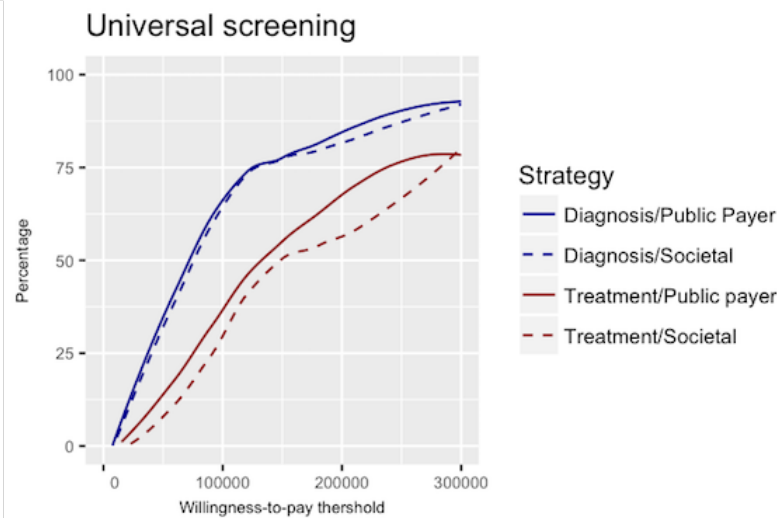
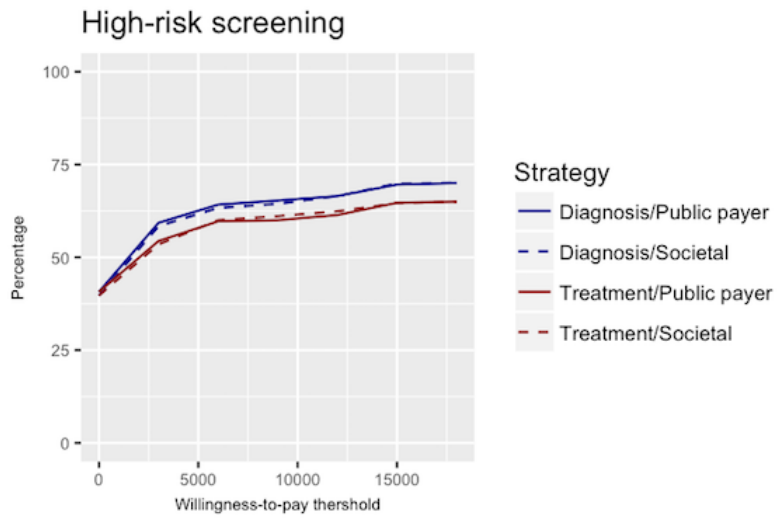


Figure 0.5 One-way deterministic sensitivity analysis for high-risk screening (left column) and universal screening (right column) compared to surveillance in the *public payer* perspective. Top row shows ICER per additional child diagnosed before 36 months, bottom row shows ICER per additional child initiated treatment before 48 months.

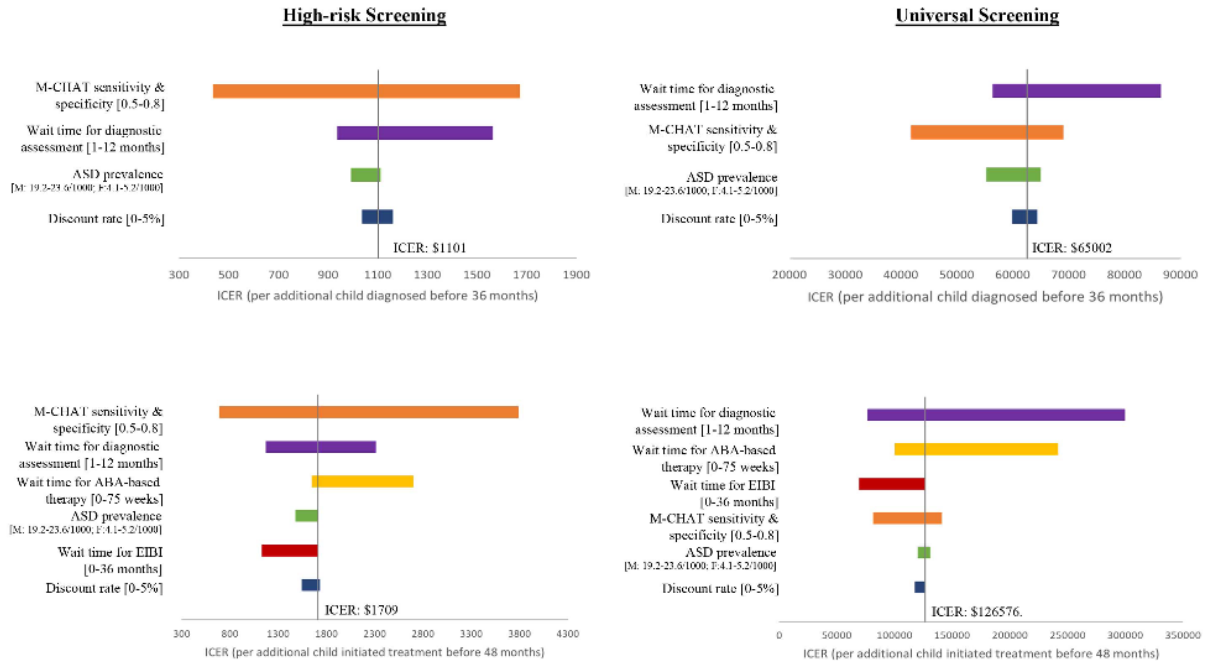
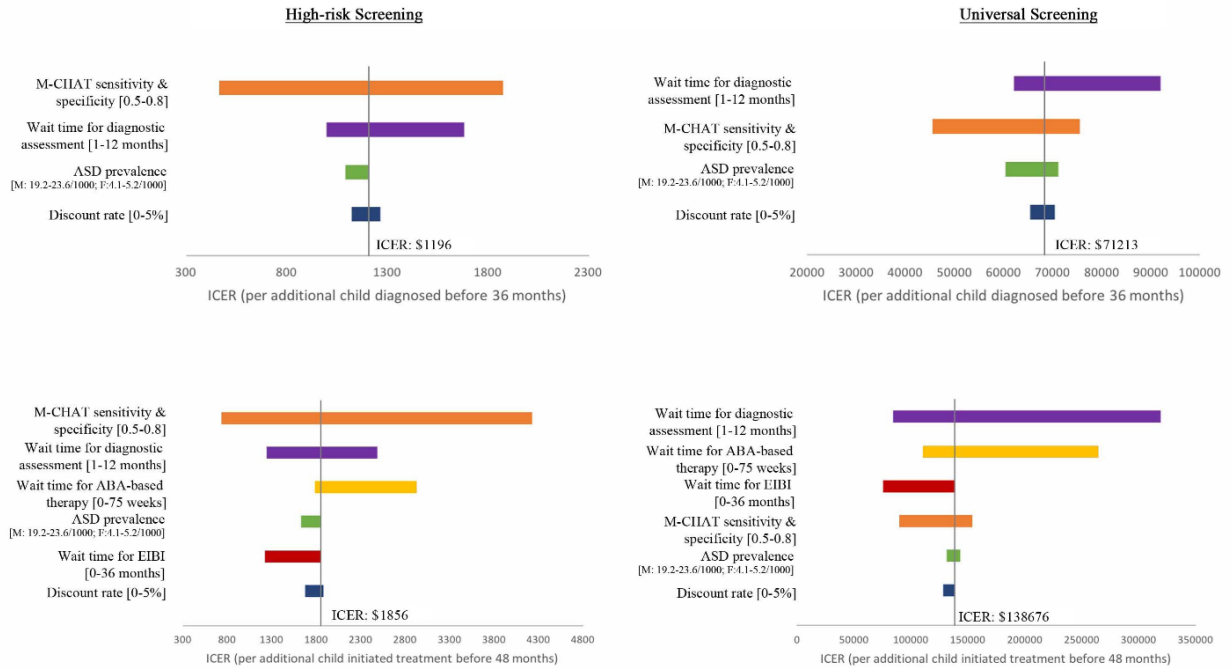


Figure 0.6 One-way deterministic sensitivity analysis for high-risk screening (left column) and universal screening (right column) compared to surveillance in the *societal* perspective. Top row shows ICER per additional child diagnosed before 36 months, bottom row shows ICER per additional child initiated treatment before 48 months.



4. COST-EFFECTIVENESS OF GENOME AND EXOME SEQUENCING IN CHILDREN WITH AUTISM SPECTRUM DISORDER

4.1 Preface

Genetic testing using CMA is one of the newer additions to the long list of recommended clinical investigations for children with ASD (Anagnostou et al., 2014). Newer genetic testing platforms, GS and ES, are being used in select tertiary settings to identify pathogenic variants that are potentially associated with ASD, in particular when CMA is non-diagnostic. However, the interpretation of these test results is often uncertain and their costs are much higher compared to CMA. This chapter compares different genomic sequencing strategies to CMA in order to determine how or if these newer platforms should have broader implementation across clinical settings.

The following manuscript has been submitted for publication and was reformatted to match the style of the thesis.

4.2 Manuscript #3

4.2.1 Abstract

Purpose: Genome (GS) and exome sequencing (ES) could potentially identify pathogenic variants with greater sensitivity than chromosomal microarray (CMA) in autism spectrum disorder (ASD), but are costlier and result interpretation can be uncertain. Study objective was to compare the costs and outcomes of four genetic testing strategies in children with ASD.

Methods: A microsimulation model estimated the outcomes and costs (in societal and public payer perspectives) of four genetic testing strategies: CMA for all, CMA for all followed by ES for those with negative CMA and syndromic features (CMA+ES), ES or GS for all.

Results: Compared to CMA, the incremental cost-effectiveness ratio (ICER) per additional child identified with rare pathogenic variants within 18 months of ASD diagnosis was \$5997.8 for CMA+ES, \$13504.2 for ES and \$10784.5 for GS in the societal perspective. ICERs were sensitive to changes in ES or GS diagnostic yields, wait times for test results or pre-test genetic counselling, but were robust to changes in the ES or GS costs.

Conclusion: Strategic integration of ES into ASD care could be a cost-effective strategy. Laboratory and genetic services need to be scaled up prior to clinical implementation to ensure timely access.

Keywords: autism spectrum disorder; genome sequencing; exome sequencing; cost-effectiveness analysis; health services

4.2.2 Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental condition that occurs in roughly 1 in 68 children in North America (Centers for Disease Control and Prevention, 2016). The disorder may run in families and twin studies reported heritability estimates greater than 90% (Tick et al., 2016). Candidate genes have also been identified, but often, the same variants can be found in individuals without ASD and with other neuropsychiatric disorders (Devlin & Scherer, 2012). Despite a high degree of genetic heterogeneity in ASD, attempting to determine the underlying genetic etiology is a component of recommended clinical care, to better inform family planning and identify comorbid medical conditions (Carter & Scherer, 2013; Schaefer & Mendelsohn, 2013).

Among individuals with ASD, approximately 10% have a Mendelian genetic condition, 5% have a cytogenetically visible chromosomal rearrangement and up to 10% have one or more rare submicroscopic copy number variants (CNVs) (Carter & Scherer, 2013; Devlin & Scherer, 2012). While the proportion of *de novo* CNVs is higher in families with at least one individual diagnosed with ASD compared to families with none, only 20-35% of rare CNVs are *de novo* (Devlin & Scherer, 2012; Iossifov et al., 2014). This suggests testing of biological parents, along with the individual with ASD, is often necessary to interpret genetic test results. Additional rare pathogenic variants are reported in syndromic individuals (i.e. with congenital anomalies in addition to clinical features typically associated with idiopathic ASD) (Marshall et al., 2008; Tammimies et al., 2015). However, individuals with syndromic features and no rare CNVs may require more thorough examination into their genome.

The current clinical guidelines by the American College of Medical Genetics and Genomics (ACMG), the International Standard Cytogenomic Array Consortium (ISCA) and the American Academy of Child and Adolescent Psychiatry recommend the use of chromosomal microarray (CMA) in all individuals diagnosed with ASD (Anagnostou et al., 2014; Miller et al., 2010; Schaefer & Mendelsohn, 2013; Volkmar et al., 2014). CMA can detect submicroscopic CNVs and has a diagnostic yield of 10-20% in individual with ASD (Carter & Scherer, 2013). Other genetic testing may be indicated depending on the specific clinical features and may include Fragile X testing, *MeCP2* testing for females and *PTEN* testing for individuals with absolute macrocephaly (Carter & Scherer, 2013; Miller et al., 2010; Schaefer & Mendelsohn, 2013).

Genome sequencing (GS) and exome sequencing (ES) are increasingly being used in research and clinical settings to identify genetic variants that would be missed by platforms with lower resolution. In one study (Tammimies et al., 2015) where individuals with ASD underwent both ES and CMA, eight of 95 participants received an ASD-related molecular diagnosis on ES while only two were identified by CMA. There are no published articles that directly compared the diagnostic yield of GS to CMA in ASD, but two studies (Jiang et al., 2013; Yuen et al., 2015) reported GS detected pathogenic variants in 19% and 42% individuals with ASD, respectively.

Despite higher diagnostic yields, high costs and uncertain clinical relevance of test results have limited the use of ES and GS clinically for ASD. Also, policies for coverage by provincial health care plans in Canada are in the process of development. GS is not yet available clinically. For ASD, ES may be offered for individuals with syndromic features who received a negative CMA. The use of ES or GS as first-line tests in these individuals could potentially eliminate the need for CMA and reduce healthcare costs by shortening the time to molecular diagnosis and discovering additional medically actionable variants. To date, there is no published economic evaluation on genetic testing in ASD (Ziegler, Rudolph-Rothfeld, & Vonthein, 2017).

The objective of this study was to compare the cost-effectiveness of using GS or ES to CMA in children with ASD. A cost-effectiveness analysis predicted the cost, in the public payer (i.e. provincial government) and societal perspectives, and outcomes of four genetic testing strategies.

4.2.3 Methods

A microsimulation model was constructed to predict the costs and consequences in a cohort of children diagnosed with ASD in Ontario, Canada under four genetic testing strategies. Each simulated child in the cohort of 1000 was assigned a set of characteristics (Appendix 4.1) selected because of their influence on the pathway of care. The time horizon was two years starting from the time of ASD diagnosis and model cycle length was one week. Costs and outcomes that surpassed one year were discounted at 3% (Sanders et al., 2016) to reflect the present value of costs and benefits incurred in the future.

4.2.3.1 Genetic Testing Strategies

Four strategies were compared. The reference approach was based on the current guideline in Ontario (Anagnostou et al., 2014), which is CMA as a first-line test for all children diagnosed with ASD. The second comparator was CMA followed by ES for children with syndromic features in addition to ASD with negative results on CMA, referred to as CMA+ES. Syndromic features were defined as clinical features not typically found in idiopathic ASD; these included microcephaly, macrocephaly, and congenital anomalies. The two other comparators were using ES or GS as first-line genetic testing for all children with ASD, instead of CMA. The specific testing platforms modelled were GeneChip System 3000Dx (Affymetric, U.S.A.) for CMA, HiSeq 2500 System (Illumina, U.S.A.) for ES and HiSeq X (Illumina, U.S.A.) for GS.

4.2.3.2 Outcome

Study outcome was the number of children diagnosed with rare pathogenic genetic variants within 18 months of ASD diagnosis. For this study, a rare pathogenic variant could be primary (i.e. known to be associated with ASD and/or other clinical features) or secondary (i.e. medically actionable genetic variants that are recommended to be reported by the ACMG (Green et al., 2013; Kalia et al., 2016)). The outcome accounted for differences in wait times for test results and in the length of post-test genetic counselling required between genetic testing strategies. ES and GS test results require longer times to process the samples, interpret the results, and communicate results to patients, since specific expertise of a genetics professional is required.

4.2.3.3 Costing

Cost items (Appendix 4.2) were considered from the public payer and societal perspectives and were generated using a probabilistic approach. The costs of each genetic test was the total of each component cost based on a microcosting study (Tsiplova et al., 2017). Validation tests were carried out for all positive and a proportion of negative ES and GS tests. Follow-up testing using Sanger sequencing for the parents' DNA and fluorescence in situ hybridization (FISH) for both parents' and child's DNA were carried out for all children with validated positive findings. As Fragile X syndrome cannot be detected by any of the comparators, Fragile X testing was included as a fixed cost for individuals with intellectual disability (Carter & Scherer, 2013; Filipek et al., 2000; Johnson et al., 2007; Schaefer & Mendelsohn, 2013; Vermeesch et al., 2007) and positive test results were not counted in the outcome for all four strategies.

Direct healthcare costs also included genetic counselling before and after genetic testing. The length of the sessions and the type of clinician delivering the results varied by genetic testing platform and type of test results. Under the societal perspective, parental time lost from accompanying their child to medical visits was estimated from the total duration of all medical visits and valued using the human capital approach using sex-specific hourly wages in adults (Statistics Canada, 2015b). As both biological parents undergo genetic testing with the child (i.e. trio testing) to delineate inheritance pattern of identified variants, the model assumed one male and one female adult accompanied each child to all assessments. All cost items are expressed in 2016 Canadian dollars (median 2016 conversion CAD\$ 1.00 = US\$ 0.77).

4.2.3.4 Model Pathway

Figure 4.1 is a schematic diagram of the health states in the microsimulation model built using TreeAge 2015 (TreeAge Software, 2015). First, the child was referred to pre-test genetic counselling which consisted of one or two consultation sessions with a medical geneticist and then with a genetic counsellor. The model assumed that at the end of the pre-test counselling session venipuncture for the child and both parents were carried out and samples were delivered to the laboratories for genetic testing. Each child received either a positive primary finding only, a positive secondary finding only, positive primary and secondary findings or a negative finding.

The families received genetic test results for the child and parents (if follow-up testing for parents occurred) during the post-test counselling session. It was assumed that families received negative test results through a telephone call by a medical geneticist. Children with positive primary and/or secondary findings underwent in-person post-test genetic counselling sessions, with longer sessions for secondary findings. The child then exited the model after post-test counselling was completed. Wait times, length and frequency of counselling sessions and diagnostic yields were randomly generated from distributions estimated from published literature or in consultation with clinicians (Appendix 4.1). Diagnostic yield for GS was modified in consultation with geneticists because it has yet to be validated for clinical use and the number of genetic variants reported in clinical settings is lower compared to research settings due to more stringent criteria. The wait time for test results (Appendix 4.1) varied by testing platform and type of results; it included time needed for processing the samples, interpreting test results, validation testing, follow-up testing in child and parents (if needed), and writing up the laboratory report.

4.2.3.5 Statistical Analysis

Predicted outcomes and costs of the three alternate genetic testing strategies were compared to the reference strategy, CMA, using incremental analysis and summarized as incremental cost-effectiveness ratios (ICERs). The ICER is the difference in cost divided by the difference in outcome between two strategies and represents the incremental cost per unit gain of health outcome. For this study, the ICER was expressed as the incremental cost per additional child identified with any rare pathogenic variants within 18 months of ASD clinical diagnosis.

Uncertainties in the ICERs were assessed using bootstrap sampling to mimic 1000 study replications. ICERs were calculated for each hypothetical study replication and were plotted on a cost-effectiveness plane. A cost-effectiveness acceptability curve (CEAC) was also constructed by plotting the proportion of ICERs that were below the willingness-to-pay (WTP) threshold, using a series of thresholds ranging from \$0 to \$20000/child.

The influence of changes in input parameters on outcomes was quantified using one-way sensitivity analysis. As GS and ES are emerging technologies in ASD and their clinical utilities

are uncertain, the analysis was repeated using two additional sets of parameter distributions that reflected the “best case” and “worst case” scenarios. The “best case” distributions had shorter wait times for test results and higher diagnostic yields. Best and worst case distributions were centered on the upper and lower 95% confidence interval (CI) of the distributions in the original model (e.g. wait time distribution in best case scenario centered on lower bound of the original CI). The alternate distributions and the ranges for one-way sensitivity analysis are listed in Appendix 4.3.

4.2.4 Results

Table 4.1 summarizes the mean and incremental costs and outcomes for the four strategies. The mean costs, in either perspective, and outcomes were approximately 2 times higher in ES and 3 times higher in GS compared to CMA. Using CMA as the reference strategy, ES alone was the least cost-effective option of the three comparators while CMA+ES had the lowest ICER.

All ICERs from bootstrap simulation comparing the three comparators to CMA alone were in the northwest quadrant of the cost-effectiveness plane (Appendix 4.4), indicating that the comparators were higher in cost and also identified more children with pathogenic variants within 18 months. The CEAC (Figure 4.2) revealed that all simulated ICERs could be considered cost-effective at thresholds of \$10000 for CMA+ES and \$15000 for GS. Consistent with previous results, ES compared to CMA was the least cost-effective option and 1% of the simulated ICERs remained above the WTP threshold of \$20000/child.

4.2.4.1 Alternate Parameter Distributions

In both “best case” and “worst case” scenarios (Table 4.2), ES was the least cost-effective strategy of the three comparators and had the highest ICER. Higher diagnostic yields and shorter wait times in the “best case” scenario also led to increases in validation testing, follow-up testing for parents, and more and longer post-testing counselling sessions, which increased costs. The effectiveness of GS was highly sensitive to the wait time for test results and <1% of children, compared to 61.2% in the “best case” and 28.5% in the reference case scenario, received their

test results within 18 months of ASD diagnosis under the “worst case” scenario. On the other hand, GS had the lowest ICER in the “best case” scenario.

4.2.4.2 One-way Sensitivity Analysis

The tornado diagram summarizes the results of one-way sensitivity analysis under the societal perspective (Figure 4.3). All ICERs were sensitive to changes in wait time for pre-test genetic counselling. The effectiveness of CMA+ES began to decrease when pre-test wait time exceeded 53 weeks because fewer children received their ES test results within 18 months. When pre-test wait time reached 58 weeks, the addition of ES for syndromic patients with a negative CMA resulted in an effectiveness of serial testing that was equivalent to CMA alone, but CMA+ES had higher costs (incremental cost \$221.9; incremental effect: 0.00). Similarly, the effectiveness of ES and GS was lower than CMA such that CMA dominated ES when pre-test wait time was 55 weeks (incremental cost \$891.3; incremental effect -0.02) and dominated GS when wait time was 29 weeks (incremental cost \$2113.8; incremental effect: -0.03).

The results were also sensitive to diagnostic yields of ES or GS. When the diagnostic yield of ES changed from 5% below base value to 5% above, the ICER for CMA+ES dropped slightly (by 5%) but the decrease in the ICER for ES alone was more pronounced at 20%. The same percentage change from the base value of GS diagnostic yield resulted in a 14% reduction in the ICER for GS. Consistent with the results from the “worst case” scenario, the effectiveness of GS was highly sensitive to the wait time for GS test results; the ICER doubled when wait time for primary or negative GS results increased from the base value by 3 weeks (i.e. from 48 to 51 weeks). All ICERs were robust to changes in the cost of sequencing equipment or supplies and in the discount rate (Figure 4.3). Findings were consistent between societal and public payer perspectives (data available upon request).

4.2.5 Discussion

Findings from this study indicate that using ES or GS for all children diagnosed with ASD was costly relative to CMA. Compared to CMA, the cost per additional child diagnosed with rare pathogenic variants within 18 months of ASD diagnosis was approximately \$14000 for ES and

\$11000 for GS. However, the addition of ES for children with a negative CMA and syndromic features could be a cost-effective option at ICER \$6000/child.

The limited influence changes in the costs of GS or ES had on the ICERs, especially compared to diagnostic yields of ES or GS, emphasized that the focus of “value for money” decision-making should be on the impact of genetic testing as well as on costs. Although the continuing decrease in laboratory prices for ES and GS make both sequencing techniques increasingly attractive, their clinical utility needs to be better established prior to widespread clinical use in ASD. To date, there are few published studies that compared the clinical utility of GS or ES to CMA in ASD and the reported diagnostic yield could be inflated due to differences in genetic variant annotation between research and clinical populations. The three studies which used ES or GS in ASD (Jiang et al., 2013; Tammimies et al., 2015; Yuen et al., 2015) classified variants as ASD-specific based on a list of known ASD-risk genes, the clinical significance of which may differ when interpreted in the context of the individual’s clinical presentation and cascade testing of parents and other family members. As the list of possible ASD-associated variants continues to grow as new evidence emerges, diagnostic yield will likely increase in the future. In turn, the higher demand for genetic counsellors and medical geneticists to interpret and communicate test results may prolong wait time and decrease clinical efficiency. Similarly, the clinical utility of test results is also dependent on the accessibility and availability of the appropriate follow-up diagnostic assessment, preventive intervention and treatment for comorbid medical conditions associated with the detected pathogenic variants.

Streamlining the pipeline for clinical sequencing and automating annotation of pathogenic variants are also critical given the ICERs were highly sensitive to changes in wait time. The use of ES as a diagnostic test is currently available in select tertiary hospitals in Canada and GS is not offered clinically as yet. In turn, the pipeline is being refined and the turnaround time for test results will likely shorten with increased familiarity with and automation of the new sequencing techniques. Delays due to the availability of genetic counsellors and medical geneticists to review genetic findings and deliver test results will likely remain a critical issue as the demand for genetics services increases with genomic sequencing being offered to more patient groups. The time interval used for the outcome in this study was lenient to account for the novelty of GS

and ES, but a desirable turnaround time for testing and results reporting should be less than 18 months. The estimated mean times from referral to completion of all post-test genetic counselling were 10.8 and 17.3 months for ES and GS, respectively, which might be considered excessively long for some families. As ES and GS results could be used to inform family planning and immediate risk of comorbid conditions, delay in receiving test results could reduce the perceived utility of test results. Therefore, successful implementation of GS and ES requires strengthening of clinical genetics and laboratory services such that they are accessible in a timely manner.

To date, there is no clinical guideline on the use of GS or ES in children with ASD. ES is currently used in specific tertiary settings and often limited to children with a clinical indication of an underlying Mendelian (i.e. single-gene) condition. This study demonstrates that the use of ES after a negative CMA in these individuals could be considered cost-effective compared to CMA alone if the payer is willing to spend \$10000 to identify each additional child with a pathogenic variant. A recent study in intellectual disability (Monroe et al., 2016) reported potential cost savings by using ES as first-line test in place of other genetic and metabolic investigations. Our findings indicate that GS would be the more cost-effective approach compared to ES, if CMA was to be replaced by sequencing. However, policies on provincial health coverage for ES or GS are still being developed in Canada.

4.2.5.1 Limitations

Although the genetic test results do not directly influence ASD treatment, the diagnosis of comorbid conditions can change the course of non-ASD related healthcare. Inclusion of follow-up assessment and treatment of comorbid conditions was beyond the scope of the current study, but their inclusion would likely increase the clinical utility of ES and GS. However, how patient management, health services use and health outcomes change after receiving genetic test results is still under study. A dynamic lifetime model is needed to estimate the potential impact of genetic testing and timing of intervention on long-term health outcomes for identified childhood-onset and adult-onset conditions and on ASD over time. Another limitation was that the model assumed that pattern of referral for additional testing was consistent between clinicians. The characteristics defined as syndromic in this study might not be considered syndromic by all

clinicians, which could lead to differential referral for ES in CMA+ES and bias the estimation of both costs and health outcome in either direction.

This study defined the outcome as the number of children identified with pathogenic variants within 18 months of ASD diagnosis, but might have excluded some potential benefits of genetic testing. For example, genetic test results can be used to guide family planning or provide an explanation for a child's ASD diagnosis. Inclusion of such benefits in the study outcome would have lowered the ICERs for GS and ES, but there is no systematic documentation on the personal utility of genetic test results.

4.2.5.2 Conclusion

Study results indicate that ES and GS are costly compared to CMA with ICERs up to \$10784.5 for GS and \$13504.2 for ES. Strategic implementation of ES in ASD, however, could be a cost-effective option. While the continuous drop in prices for GS and ES make the new technologies attractive options, whether and how these test results will influence clinical care needs to be better established. Moreover, clinical genetic and laboratory services need to be strengthened in order to handle anticipated increase in demand and ensure equitable access. Despite the promise of higher diagnostic yield of GS and ES compared to CMA, the long wait time for genetic services and high costs of follow-up testing mitigate the potential benefit of these new technologies. In order to increase the cost-effectiveness of GS and ES, the focus should shift from solely on improving diagnostic yield and lowering costs to consider the increasing need for genetic services and the utility, both personal and clinical, of test results.

Table 4.1 Mean and incremental costs and outcomes by test strategy.

Strategy	Cost (Public Payer)		Cost (Societal)		Effect		ICER	
	Mean (sd)	Incremental	Mean (sd)	Incremental	Mean (sd)	Incremental	Public Payer	Societal
CMA	996.9 (284.3)	Ref	1079.9 (287.8)	Ref	0.090 (0.285)	ref	ref	ref
CMA+ES	1211.8 (554.2)	214.9	1301.8 (559.9)	221.9	0.127 (0.332)	0.037	5808.9	5997.8
ES	1878.1 (155.5)	881.2	1971.2 (170.1)	891.3	0.156 (0.360)	0.066	13352.9	13504.2
GS	3089.9 (241.8)	2093.0	3193.7 (260.3)	2113.8	0.285 (0.442)	0.196	10678.8	10784.5

CMA: chromosomal microarray; ES: exome sequencing; ICER: incremental cost-effectiveness ratio. GS: genome sequencing; sd: standard deviation.

Table 4.2 Incremental cost-effectiveness ratios from sensitivity analyses using alternate distributions.

Strategy	“Best case”		“Worst case”	
	Public Payer	Societal	Public Payer	Societal
CMA	ref	ref	ref	ref
CMA+ES	\$5757.4	\$5930.7	\$14125.0	\$14506.3
ES	\$6393.5	\$6492.7	\$59500.0	\$60013.3
GS	\$4325.1	\$4384.4	dominated	dominated

CMA: chromosomal microarray; ES: exome sequencing; ICER: incremental cost-effectiveness ratio. GS: genome sequencing

Figure 0.1 Schematic diagram of the microsimulation model.

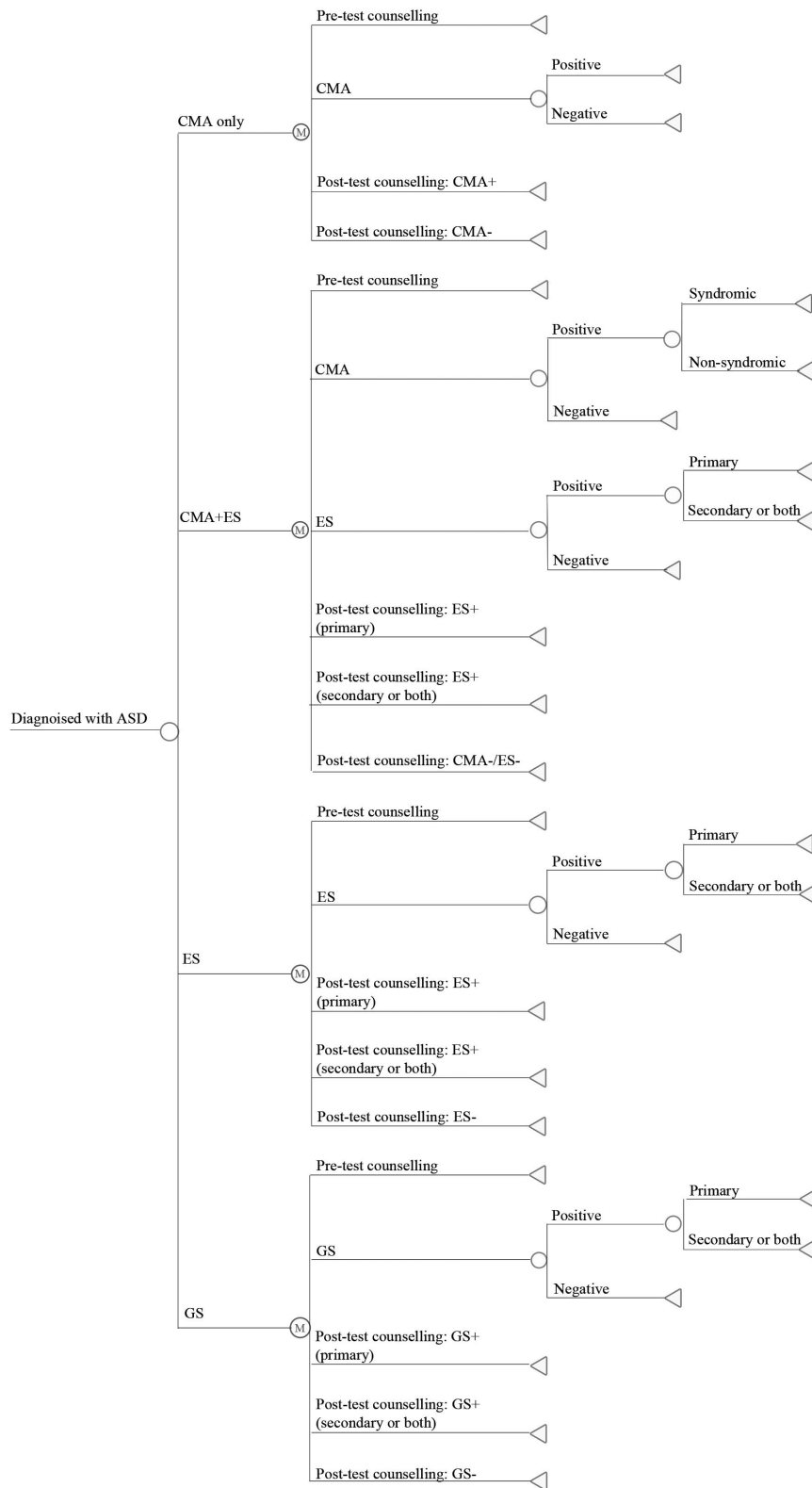


Figure 0.2 Cost-effectiveness acceptability curve of the three comparison strategies using chromosomal microarray as reference in the societal perspective.

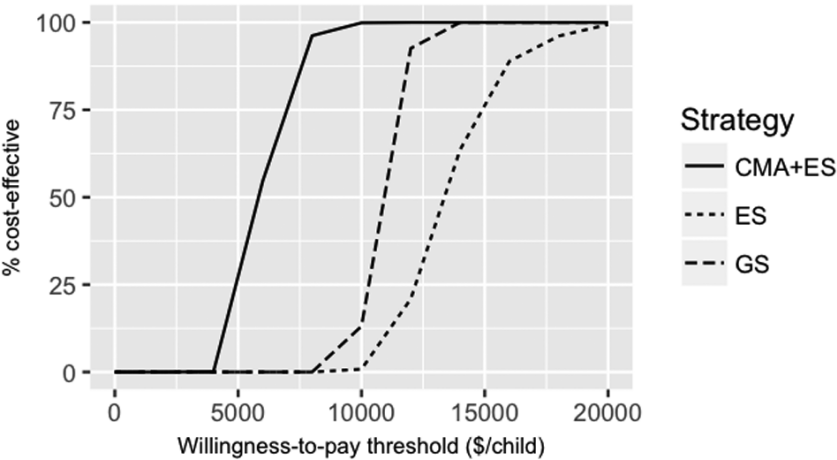
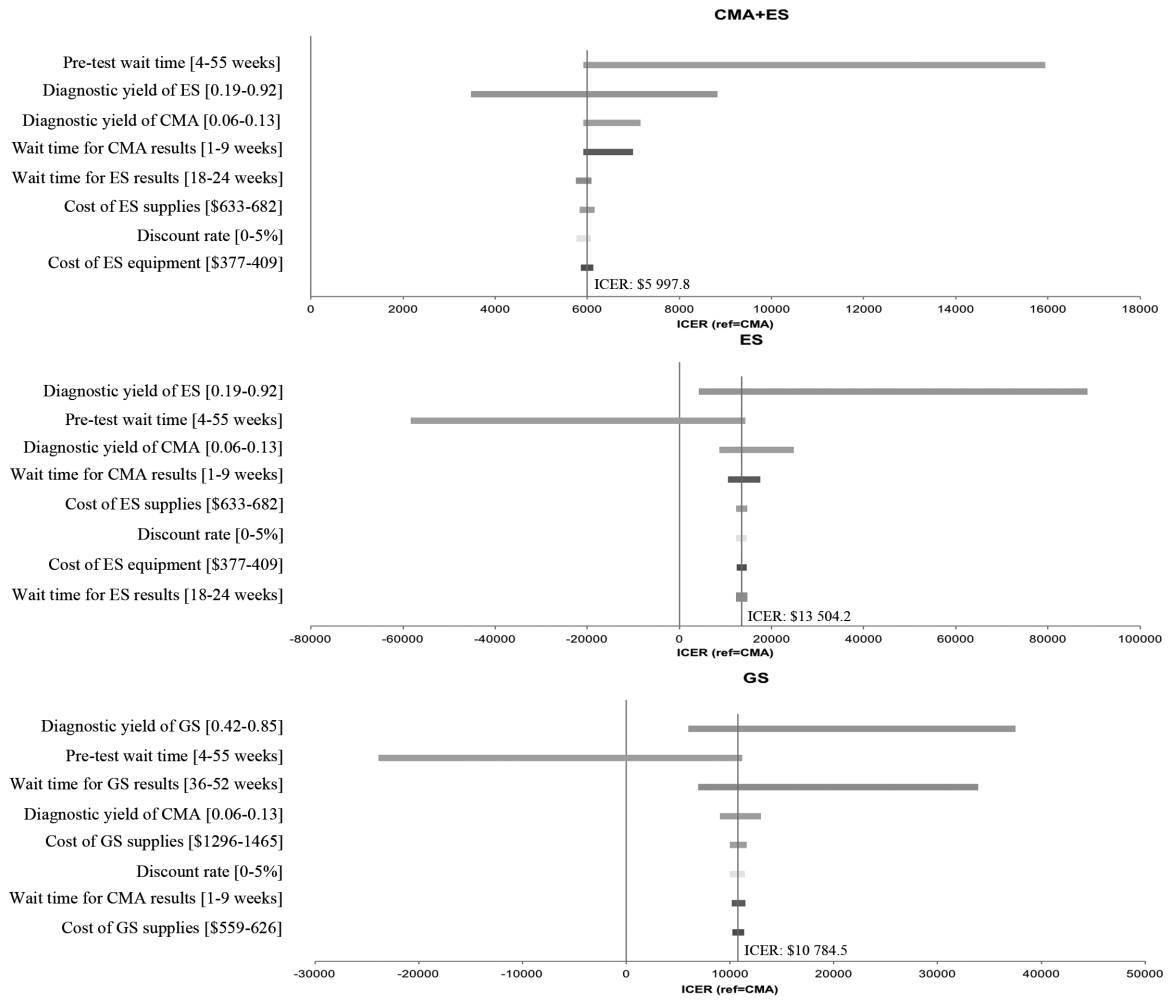


Figure 0.3 One-way sensitivity analysis under the societal perspective.



Appendix 0.1 Input parameters used in the microsimulation model.

Patient Characteristics	Distributions	Source
Sex (Male)	Bernoulli(0.85)	(Centers for Disease Control and Prevention, 2016; Coo et al., 2012)
Intellectual disability	Male: Bernoulli (0.20) Female: Bernoulli (0.55)	(Banach et al., 2009; Charman et al., 2011; Yeargin-Allsopp et al., 2003)
Congenital anomalies	Bernoulli (0.11)	(S. Dawson, Glasson, Dixon, & Bower, 2009; Timonen-Soivio et al., 2015; Wier, Yoshida, Odouli, Grether, & Croen, 2006)
Macrocephaly (head circumference >97 th percentile)	Bernoulli(0.17)	(Fombonne, Rogé, Claverie, Courty, & Frémolle, 1999; Lainhart et al., 2006)
Microcephaly (head circumference <3 rd percentile)	Bernoulli (0.03)	(Lainhart et al., 2006)
Wait Times (weeks)		
Pre-testing genetic counselling	Gamma(36,1.5)	(Ontario Genetics Secretariat, 2014)
CMA test results	Trun N(5, 2, 3, 7)	(Ny Hoang, MSc, email communication, April 2016).
ES test results- primary or negative	Trun N(20, 2, 18, 24)	(Melissa Carter, MD, email communication, 2016).
ES test results- secondary	ES primary + 2 weeks	
GS test results- primary or negative	Trun N(48, 2, 44, 52)	(Ny Hoang, MSc, email communication, April 2016).
GS test results- secondary	GS primary + 2 weeks	
Diagnostic Yield		
Chromosomal microarray	N(0.09,0.02)	(McGrew, Peters, Crittendon, & Veenstra-Vanderweele, 2012; Shen et al., 2010; Tammimies et al., 2015)
Exome sequencing (all children with ASD)		(Tammimies et al., 2015)
Primary variants only	N(0.08,0.02)	
Secondary variants only	N(0.06,0.02)	
Both primary and secondary	N(0.024,0.01)	
Exome sequencing (syndromic children with ASD and negative CMA)*		(Tammimies et al., 2015)
Primary variants only	N(0.13,0.024)	
Secondary variants only	N(0.10,0.03)	
Both primary and secondary	N(0.04,0.01)	

Genome sequencing		(Yuen et al., 2015)
Primary variants only	N(0.16,0.05)	
Secondary variants only	N(0.21,0.09)	
Both primary and secondary	N(0.05,0.02)	

*N(0.6, 0.1) of syndromic children with ASD with a rare genetic variant are assumed to be detected by CMA (Tammimies et al., 2015).

CMA: chromosomal microarray; ES: exome sequencing; GS: genome sequencing.

Bernoulli(p) denotes a Bernoulli distribution where p is the probability of having the trait. Gamma(α, λ) represents the Gamma distribution, where α is the shape parameter and λ is the scale parameter. Trun N($\mu, \sigma, \alpha, \beta$) represents truncated normal distribution where μ is the mean, σ is the standard deviation, α is the minimum value and β is the maximum value. N(μ, σ) denotes the normal distribution where μ is the mean and σ is the standard deviation.

Appendix 0.2 Unit price and resource use of cost items in the microsimulation model.

Cost Items	Distribution Unit Cost (\$)	Source	Resource Use	Source
Single Gene Tests				
Fragile X	N(325,2.5)	(Hospital for Sick Children, 2015)	In children with ID	
Chromosomal Microarray (CMA)				
Pre-testing counselling				
Medical geneticist	N(90.9,8.7) ^a	(Ministry of Health and Long-Term Care, 2015; Régie de l'assurance maladie du Québec, 2015)	1 session of N(30,7) minutes	(Ny Hoang, MSc, email communication, April 2016)
Genetic counsellor	N(40.1, 2.13) ^b	(Cheryl Shuman, MSc, email communication, January 2016)	1 session N(60,15) minutes	(Ny Hoang, MSc, email communication, April 2016)
Cost of 1 test				
Labour	N(141.6, 4.8)	(Tsiplova et al., 2017)	Per test	
Overhead	N(39.5,1.1)	(Tsiplova et al., 2017)	Per test	
Equipment	N(30.1,1.0)	(Tsiplova et al., 2017)	Per test	
Supplies	N(434.6,3.5)	(Tsiplova et al., 2017)	Per test	
Validation (qPCR)	N(223.9, 12.6)	(Tsiplova et al., 2017)	Positive finding in child: 1 test	(Tsiplova et al., 2017)
Follow-up genetic testing (FISH)	N(671.1,8.5)	(Tsiplova et al., 2017)	Positive finding in child: 1 trio	(Tsiplova et al., 2017)
Post-testing counselling				
Medical geneticist	N(90.98,8.7) ^a	(Ministry of Health and Long-Term Care, 2015; Régie de l'assurance maladie du Québec, 2015)	Positive finding in child: 1 session of N(30,7) minutes Negative finding in child: 1 session of N(15,3) minutes	(Ny Hoang, MSc, email communication, April 2016)
Exome Sequencing				
Pre-testing counselling				
Medical geneticist	N(90.9,8.7) ^a	(Ministry of Health and Long-Term Care, 2015; Régie de l'assurance maladie du Québec, 2015)	90% children: 1 session of N(30,7) minutes	(Ny Hoang, MSc, email communication, April 2016)

Cost Items	Distribution Unit Cost (\$)	Source	Resource Use	Source
Genetic counsellor	N(40.1, 2.13) ^b	(Cheryl Shuman, MSc, email communication, January 2016)	10% children: 2 sessions of N(30,7) minutes each 90% children: 1 session N(60,15) minutes 10% children: 2 sessions N(60,15) minutes each	(Ny Hoang, MSc, email communication, April 2016)
Cost of 1 test				
Labour	N(318.4,12.4)	(Tsiplova et al., 2017)	Per test	
Overhead	N(165.5,3.4)	(Tsiplova et al., 2017)	Per test	
Equipment	N(394.5,8.0)	(Tsiplova et al., 2017)	Per test	
Bioinformatics	N(6.2,0.4)	(Tsiplova et al., 2017)	Per test	
Supplies	N(657.7,12.2)	(Tsiplova et al., 2017)	Per test	
Validation (Sanger sequencing)	N(38.5, 0.8)	(Tsiplova et al., 2017)	Positive finding in child: 2 tests 30% negative finding: 2 tests	(Tsiplova et al., 2017)
Follow-up genetic testing (Sanger sequencing)	N(38.5, 0.8)	(Tsiplova et al., 2017)	Primary or secondary finding in child: 4 tests (2 for each parent)	(Tsiplova et al., 2017)
Post-testing counselling				
Medical geneticist	N(90.9,8.7) ^a	(Ministry of Health and Long-Term Care, 2015; Régie de l'assurance maladie du Québec, 2015)	Primary or secondary finding in child: 1 session of N(30,7) minutes Negative finding: 1 session of N(15,3) minutes	(Ny Hoang, MSc, email communication, April 2016)
Genetic counsellor	N(40.1, 2.13) ^b	(Cheryl Shuman, MSc, email communication, January 2016)	Primary finding in child: 0 Secondary finding in child: 1 session N(60,15) minutes	(Ny Hoang, MSc, email communication, April 2016)
Genome Sequencing				
Pre-test counselling				
Medical geneticist	N(90.98,8.7) ^a	(Ministry of Health and Long-Term Care, 2015; Régie de l'assurance maladie du Québec, 2015)	90% children: 1 session of N(30,7) minutes 10% children: 2 sessions of N(30,7) minutes each	(Ny Hoang, MSc, email communication, April 2016)
Genetic counsellor	N(40.1, 2.13) ^b	(Cheryl Shuman, MSc, email communication, January 2016)	90% children: 1 session N(60,15) minutes	(Ny Hoang, MSc, email communication, April 2016)

Cost Items	Distribution Unit Cost (\$)	Source	Resource Use	Source
			10% children: 2 sessions N(60,15) minutes each	
Cost of 1 test				
Labour	N(250.5,12.2)	(Tsiplova et al., 2017)	Per test	
Overhead	N(241.7,5.3)	(Tsiplova et al., 2017)	Per test	
Equipment	N(592.7,17.2)	(Tsiplova et al., 2017)	Per test	
Bioinformatics	N(207.5,8.9)	(Tsiplova et al., 2017)	Per test	
Supplies	N(1381.1,43.1)	(Tsiplova et al., 2017)	Per test	
Validation genetic testing	N(38.5,0.8)	(Tsiplova et al., 2017)	Positive finding: 2 tests 20% negative finding: 2 tests	(Tsiplova et al., 2017)
Follow-up genetic testing				
Sanger sequencing	N(38.5,0.8)	(Tsiplova et al., 2017)	Primary or secondary finding in child: 4 tests (2 for each parent)	(Tsiplova et al., 2017)
qPCR	N(684.9, 22.8)	(Tsiplova et al., 2017)	In 10% of positive (primary or secondary): 1 trio	(Tsiplova et al., 2017)
Post-testing counselling				
Medical geneticist	N(90.98,8.7) ^a	(Ministry of Health and Long-Term Care, 2015; Régie de l'assurance maladie du Québec, 2015)	Primary or secondary finding in child: 1 session of N(30,7) minutes Negative finding: 1 session of N(15,3) minutes	(Ny Hoang, MSc, email communication, April 2016)
Genetic counsellor	N(40.1, 2.13) ^b	(Cheryl Shuman, MSc, email communication, January 2016)	Primary finding in child: 0 Secondary finding in child: 1 session N(60,15) minutes	(Ny Hoang, MSc, email communication, April 2016)
Productivity Cost (per hour)				
Mother	N(21.6,0.8)	(Statistics Canada, 2015b)	Pre-testing counselling (90 minutes) + post-testing counselling (30 minutes for primary, 90 minutes for secondary)	(Ny Hoang, MSc, email communication, April 2016)

Cost Items	Distribution Unit Cost (\$)	Source	Resource Use	Source
Father	N(25.5, 0.8)	(Statistics Canada, 2015b)	Pre-testing counselling (90 minutes) + post-testing counselling (30 minutes for primary, 90 minutes for secondary)	(Ny Hoang, MSc, email communication, April 2016)

^aBased on physician fee code K016 in Ontario and 09056 in Quebec. ^bBased on hourly rate at Hospital for Sick Children.

FISH: fluorescence in situ hybridization; ID: intellectual disability; qPCR: real-time polymerase chain reaction

N(μ , σ) denotes the normal distribution where μ is the mean and σ is the standard deviation.

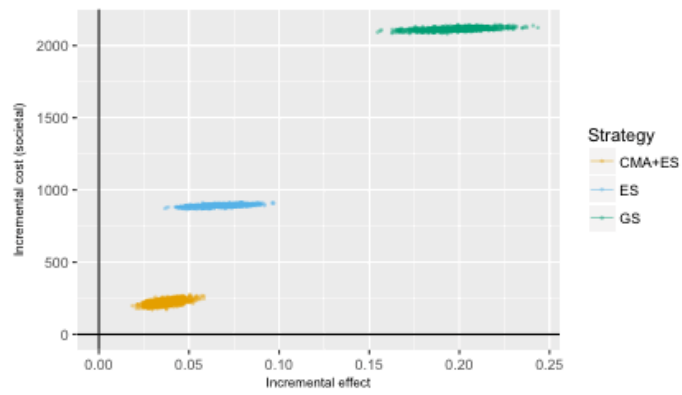
Appendix 0.3 Values used in one-way sensitivity analysis and alternate parameter distributions for microsimulation model.

	Range	"Best Case"	"Worst Case"
Diagnostic Yield			
Chromosomal microarray	0.06-0.13	N(0.12,0.02)	N(0.06,0.02)
Exome sequencing (all children with ASD)			
Primary variants only	0.04-0.16	N(0.13,0.02)	N(0.03,0.02)
Secondary variants only	0.01-0.20	N(0.11,0.02)	N(0.02,0.02)
Both primary and secondary	0.01-0.15	N(0.04,0.01)	N(0.006,0.01)
Exome sequencing (syndromic children with ASD and negative CMA)			
Primary variants only	0.09-0.37	N(0.18,0.02)	N(0.08,0.02)
Secondary variants only	0.05-0.30	N(0.15,0.03)	N(0.04,0.03)
Both primary and secondary	0.05-0.25	N(0.06,0.01)	N(0.02,0.01)
Genome sequencing			
Primary variants only	0.082-0.25	N(0.26,0.05)	N(0.006,0.05)
Secondary variants only	0.33-0.50	N(0.31,0.09)	N(0.11,0.09)
Both primary and secondary	0.01-0.20	N(0.08,0.02)	N(0.001,0.02)
Wait Time (weeks)			
Pre-testing genetic counselling	4-55	N(14.06,3.48)	N(27.71,3.48)
CMA test results	1-9	N(3,2)	N(7, 2)
ES test results- primary or negative	18-24	N(18,2)	N(24, 2)
ES test results- secondary	20-26	ES primary + 2	ES primary + 2
GS test results- primary or negative	36-52	N(44,2)	N(52,2)
GS test results- secondary	38-54	GS primary + 2	GS primary + 2
Costs			
Cost of ES			
Equipment	377-409		
Supplies	633-682		
Cost of GS			
Equipment	559-626		
Supplies	1298-1465		
Discount Rate	0-5%		

ASD: autism spectrum disorder; CMA: chromosomal microarray; ES: exome sequencing; GS: genome sequencing.

$N(\mu, \sigma)$ denotes the Normal distribution where μ is the mean and σ is the standard deviation.

Appendix 0.4 Incremental costs and effects, from bootstrap simulation, of each alternate genetic testing strategy compared chromosomal microarray only in the societal perspective.



5. DISCUSSION

5.1 Summary of Main Findings

The overall goal of this thesis was to estimate the monetary and health impact of introducing two diagnostic and screening services along the ASD clinical pathway. The meta-analysis in Chapter 2 summarized the accuracy of the M-CHAT and concluded that it performs with low-to-moderate accuracy in identifying ASD among children with developmental concerns. The pooled sensitivity was 0.83 (95% CrI: 0.75, 0.90), specificity was 0.51 (95% CrI: 0.41, 0.61), PPV was 0.55 (95% CrI: 0.45, 0.66) in high-risk children and 0.07 (95% CrI: <0.01, 0.16) in low-risk children. Findings from the meta-regressions suggest that clinicians should account for a child's age (sensitivity was higher at 30 months compared to 24 months) and existing developmental concerns (PPV was higher in high-risk compared to low-risk children) when considering using the M-CHAT and interpreting its score. Quality assessment also identified potential bias in sample selection, implementation of the M-CHAT and/or ASD clinical diagnosis in all 13 included studies. Moreover, studies in low-risk samples did not follow-up with children with a negative screen, thus sensitivity and specificity of the M-CHAT in low-risk children could be not estimated. Although the M-CHAT was designed to screen high- and low-risk children ages 16 to 30 months (Robins et al., 2001), there is a lack of evidence supporting its use as part of universal screening at 18 and 24 months. Given the low pooled specificity in high-risk children and low ASD prevalence, unnecessary referral to in-depth assessment due to false positive screens would be high. Validation studies with methodological rigor in both low- and high-risk populations are needed before it can be recommended to be used on a population level.

Using results from the meta-analysis, the CEA in Chapter 3 quantified the incremental benefits and costs of universal or high-risk screening compared to surveillance monitoring in ASD. The simulation study demonstrated that universal screening would greatly burden the healthcare system by increasing the demand for ASD diagnostic services and by increasing the need for treatment services for those who may not benefit. Although universal screening was 3 times more effective compared to surveillance, the number of children referred for diagnostic assessment was also 6-fold higher and yielding ICERs of \$65000-140000/child diagnosed or initiated treatment earlier. In turn, some children waited more than 12 months for diagnostic

assessment. As some clinicians who carry out ASD diagnostic assessment also attend to children with other neuropsychiatric, behavioural or developmental conditions, this would delay access for children other than those with suspected ASD. Rather, targeting ASD screening towards children at heightened risk of ASD (e.g. children with one or more full sibling diagnosed with ASD) could be considered cost-effective at a willingness-to-pay threshold of ~\$2000 per additional child diagnosed or initiated treatment earlier. Eliminating or reducing wait time for ASD intervention had the highest impact on the effectiveness of screening, particular for surveillance. Given the hypothesized benefits of early diagnosis is mediated through early treatment initiation, resource allocated towards reducing wait times for ASD services appear be more worthwhile.

Results from the CEA on genetic testing in Chapter 4 also demonstrated that a strategic approach to resource allocation is most efficient. Compared to all children with ASD undergoing CMA, the addition of ES for children with negative CMA and syndromic symptoms was the most cost-effective genetic testing strategy, resulting in ICERs ranging from \$5800-6000 per additional child with pathogenic variants within 18 months of ASD diagnosis. If CMA was to be replaced by a new sequencing platform, GS would be a more cost-effective option compared to ES (ICERs \$10700-10800 vs. \$13400-13500). The cost-effectiveness of ES and GS were highly sensitive to changes in wait times for pre-test genetic counselling and for test results. Despite rapid decrease in the costs of GS and ES, the clinical and personal utility of genetic test results for children with ASD needs to be better established prior to clinical implementation. While this study assumed test results received within 18 months of ASD diagnosis is an acceptable timeline, a much shorter time interval should be targeted if sequencing is to be introduced broadly across clinical settings. Moreover, the potential benefit of test findings, both ASD and non-ASD specific, is contingent on availability and accessibility of appropriate follow-up assessment, preventive intervention and treatment services. In turn, accompanying services, such as clinical genetics and laboratory services, need to be scaled up to meet the anticipated surge in service demand in order to ensure children have timely access.

5.2 Implications for ASD Clinical Care

The studies in this thesis show that strategic resource allocation, such that targeted populations who would most benefit from additional health services receive them, is most efficient in terms of screening and diagnosis in ASD. Due to the cascade of medical, psychosocial and educational services required by individuals with ASD, changes in one component of care can lead to drastic increases in demand, wait time and public expenditure. This is particularly important given that the type and number of services recommended as standard care for individuals with ASD will likely increase as we continue to learn more about the condition.

An example would be clinical genetic services. Clinical guidelines published in recent years (Anagnostou et al., 2014; Miller et al., 2010; Schaefer & Mendelsohn, 2013; Volkmar et al., 2014) recommend the use of CMA in all children with ASD. The integration of ES and GS as part of recommended ASD care could occur in the near future, especially given the pace of emerging research on their ability to identify additional pathogenic genetic variants. Although findings from Chapter 4 indicate that the use of ES in children with syndromic features after a negative CMA could be considered cost-effective, policies on provincial health coverage of ES and GS are still being developed. In Ontario, ES is currently covered for select individuals on a case by case basis using a special authorization program. In 2017, in recognition of the rapid evolution of sequencing technology and the need for evidence to inform policy decision-making, Health Quality Ontario established the Ontario Genetics Testing Advisory Committee as a special sub-committee of the Ontario Health Technology Assessment Committee (OHTAC), “to advise on the clinical utility, validity, and value for money of new and existing genetic and genomic tests in Ontario to support OHTAC’s role in making recommendations” (Health Quality Ontario, 2016).

Other than the cost of the test, the clinical and personal utility (ACMG Board of Directors, 2015; Foster, Mulvihill, & Sharp, 2009) of genetic test results need to be better established prior to implementation to understand the value of adding genetic sequencing to the ASD pathway of care. While ES and GS hold promise for more personalized ASD intervention, the evidence is not yet available. Timing is another critical component to be considered as some families can feel overwhelmed when first learning of their child’s ASD diagnosis and may not want to be

informed of their child's other, potentially adult-onset (i.e. incidental findings), conditions at the same time. Conversely, the benefits perceived by patients and families from knowing the genetic basis of the condition, having timely information on one's health risk and the ability to use the test results to inform family planning may decrease given the long time interval (predicted mean of 11 months for ES and 17 months for GS from Chapter 4) between referral and receiving sequencing results.

In terms of ASD screening, the addition of targeted screening could improve identification of children who potentially have ASD, leading to earlier diagnosis and treatment. Who should undergo standardized screening, however, is likely dependent on context. Children at heightened risk for ASD (e.g. children with a full sibling diagnosed with ASD, premature babies, infants with specific genetic conditions (Grønberg et al., 2013; Limperopoulos et al., 2008; Ozonoff et al., 2011; Richards et al., 2015)) are likely to benefit from active screening in order to capture subtle abnormalities in development. While there are reports of under and/or delayed ASD diagnosis in ethnic minorities (Jo et al., 2015; Mandell et al., 2009), whether they would benefit from active screening is uncertain. For example, disparities in diagnosis due to differences in ASD clinical profiles or decreased likelihood to undergo well-child visits (Chi et al., 2013; Jhanjee et al., 2004; Tek & Landa, 2012) would not be ameliorated by more frequent screening.

The results of the two CEAs in this thesis also emphasized the need to reduce or eliminate wait times for services used by individuals with ASD as opposed to putting resources into universal screening. The impact was particularly evident when wait times for ABA-based therapy and EIBI was eliminated, which led to a 50% and 100% increase in the number of children who started treatment by 48 months. A recent Canadian study (Piccininni et al., 2017) further estimated the potential gain in health outcomes and cost savings in the long run if wait time for EIBI was reduced or eliminated. Given the rapid development in young children, timely diagnosis, of ASD and of genetic conditions, and access to appropriate intervention is critical to ensure their developmental trajectory could be reverted to a more age-appropriate level.

Although this thesis generated evidence to inform policies on ASD pathways of care, actual implementation might be difficult. Provinces struggle with introducing new standards and

policies in ASD due to the interests of different parties (i.e. parents, clinicians, educators, community) involved. In turn, effective and efficient policies in ASD care will need to balance societal values and scientific evidence.

5.3 Implications for the Healthcare System

The interventions examined in this thesis can have drastic impacts on the healthcare system, but in different ways. A negative consequence of universal screening is not just the cost of screening, but the increase in demand, wait time and expenditure for downstream services. Genetic sequencing test, on the other hand, are high in cost but test results could potentially guide prevention efforts for conditions with onset later in life, which could reduce healthcare use in the long run.

Although the focus of these studies was in ASD, the simulation models demonstrated that the impact of increasing services for one patient group can influence access for others. Availability of genetic counsellors and medical geneticists is limited and wait time for consultations can be up to 8 months for individuals with developmental delay (Ontario Genetics Secretariat, 2014). Implementation of ES and GS in ASD would require longer or more consultation sessions which would further prolong wait times for all individuals who need clinical genetic services. Similarly, some clinicians who carry out ASD diagnostic assessment also attend to individuals with other developmental, behavioural or psychiatric concerns. In turn, the delay in ASD diagnosis resulting from universal screening would be experienced by other children who require in-depth clinical assessment.

In both instances, ASD screening and implementation of genetic sequencing could potentially widen the equity gap due to differential access to health services. GS and ES are currently available in specific tertiary hospitals, in comparison to standard care which is offered by more laboratories. Families in remote regions or urban cities with low service coverage would have limited access, even if GS and ES were covered by provincial health insurance plans and were offered clinically. The same could be apply to the specialized follow-up care needed to treat and monitor conditions associated with the identified pathogenic variants. In terms of ASD screening, families who can pay out-of-pocket for behavioural or developmental interventions at

first sign of atypical development, thus bypassing the long wait times for diagnostic services and publically funded interventions, would benefit more from frequent screening than those without access to or the means to pay for private care.

Overall, policy decisions on ASD resource allocation should anticipate the potential gaps in services that would be introduced. One of the benefits of system-wide simulation models is the ability to estimate the impact of policy changes on the individual-level and on the healthcare system-level. While both models in Chapters 3 and 4 focused on the healthcare system, services from other sectors (e.g. social services, education) could also be included. This is particularly relevant for ASD given the spectrum of service needs across lifespan.

5.4 Future Research

The meta-analysis in Chapter 3 summarized published evidence on the M-CHAT and also identified a lack of quality research on screening tools for children without developmental concerns. In the limited studies on low-risk children, clinical diagnosis was not established for children who screened negative on the M-CHAT and the sensitivity and specificity could not be estimated. Methodological flaws, such as lack of blinding, high drop-out rates and selection bias, were also identified in the published studies, which could have biased the estimated accuracy of the M-CHAT. If a tool is to be recommended for use on a population level, it must be validated in both low- and high-risk children by studies with methodological rigour. From an analytic standpoint, this is one of the few meta-analyses where a bivariate regression model under a Bayesian framework was used. Although it was used to jointly summarize the sensitivity and specificity of a screening tool, the same statistical method would be applicable to areas where outcomes are correlated.

The outcomes of CEAs on genetic testing should move beyond health benefits (i.e. morbidity and mortality) and include the clinical and personal utility a family gains from genetic test results. How health service utilization changes after receiving a genetic diagnosis is not only critical for establishing the clinical utility, but also for the government and healthcare providers to anticipate changes in demand for services. Given the novelty of the construct, qualitative studies using a variety of sample populations and genetic testing scenarios are likely needed to

better define what personal utility entails. For example, personal utility gained from genetic sequencing in children, where results could affect parents (e.g. explanation for ASD by establishing a genetic etiology, family planning) and the patient themselves (e.g. awareness of health risks), will differ from perceived personal utility from testing in adult populations. After delineating what personal utility is, the construct could then be accurately measured and potentially be integrated in economic evaluations.

As simulation models require large amounts of detailed information to accurately reflect clinical reality, several areas with limited high-quality evidence were identified. Data on the longitudinal outcomes for children with ASD identified by active screening compared to surveillance and between children who underwent ASD intervention at different ages is crucial to better estimate the long-term impact of more vigorous screening. Information on out-of-pocket expenditures in parents of children with ASD is needed not only for CEA, but for the government to plan for changes in demand for publically funded services. The studies in this thesis add to the limited published simulation studies in ASD (Mavranouzouli et al., 2014; Motiwala et al., 2006; Penner et al., 2015; Piccininni et al., 2017), and the first to use DES. This research demonstrates that dynamic modelling could better describe the complexity of developmental trajectories in children and how service use varies depending on each child's presenting symptoms. Another advantage of using the DES is that it could quantify changes in wait time, a widely used performance benchmark in the Canadian healthcare system, in relation to service use and referral. Given the spectrum of services that individuals with ASD use, future models should also broaden the clinical pathway in order to capture health, psychosocial and educational services. With more accurate and longitudinal inputs, simulation models could better predict the short- and long-term impact of policy decisions on the individual-level, for patients and families, and on the system-level across sectors.

5.5 Conclusion

Given the network of services necessary to care for individuals with ASD, changes in one area could have a large impact on the healthcare system overall. Strategic resource allocation is critical to ensure that introduction of new services is efficient and effective. Newer genetic sequencing platforms are available, but existing clinical genetics and laboratory facilities may

not be able to support the increase in service demand from offering sequencing to all individuals with ASD. Rather, offering ES to children with a clinical indication that a more thorough examination of their genome is necessary could be a cost-effective option.

Although universal screening could lead to more children diagnosed or initiated treatment earlier compared to surveillance monitoring, it would also greatly burden the healthcare system and further delay access for children who require in-depth behavioural or psychiatric evaluation. A more cost-effective and efficient strategy would be to screen children at heightened risk for ASD, but criterion to define this high-risk population requires further study. Other than the lack of evidence supporting universal screening, studies with methodological rigour are needed to validate the use of a commonly used screening tool, M-CHAT, on a population level. Additional research is also needed prior to clinical implementation of either genetic sequencing or ASD screening to ensure timely and equitable access to services for individuals with ASD.

BIBLIOGRAPHY

- ACMG Board of Directors. (2015). Clinical utility of genetic and genomic services: a position statement of the American College of Medical Genetics and Genomics. *Genetics in Medicine*, 17(6), 505–507. <https://doi.org/10.1038/gim.2015.41>
- Al-Qabandi, M., Gorter, J. W., & Rosenbaum, P. (2011). Early autism detection: Are we ready for routine screening? *Pediatrics*, 128(1), e211–e217. <https://doi.org/10.1542/peds.2010-1881>
- American Academy of Pediatrics. (2015, October). Recommendations for preventive pediatric health care—Periodicity schedule. Retrieved from https://www.aap.org/en-us/Documents/PeriodicitySchedule2015_Visionscreening.pdf
- American College of Medical Genetics and Genomics. (2014, April 1). ACMG updates recommendation on “Opt Out” for genome sequencing return of results. Retrieved from https://www.acmg.net/docs/Release_ACMGUpdatesRecommendations_final.pdf
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (Fifth Edition). American Psychiatric Association. Retrieved from <http://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425596>
- Amiet, C., Couchon, E., Carr, K., Carayol, J., & Cohen, D. (2014). Are there cultural differences in parental interest in early diagnosis and genetic risk assessment for autism spectrum disorder? *Frontiers in Pediatrics*, 2. <https://doi.org/10.3389/fped.2014.00032>
- Anagnostou, E., Zwaigenbaum, L., Szatmari, P., Fombonne, E., Fernandez, B. A., Woodbury-Smith, M., ... Scherer, S. W. (2014). Autism spectrum disorder: Advances in evidence-based practice. *Canadian Medical Association Journal*, 186(7), 509–519. <https://doi.org/10.1503/cmaj.121756>
- Auditor General of Ontario. (2013). Autism services and supports for children. Office of Auditor General of Ontario.
- Auditor General of Ontario. (2015, December). Autism services and supports for children. Office of Auditor General of Ontario.
- Banach, R., Thompson, A., Szatmari, P., Goldberg, J., Tuff, L., Zwaigenbaum, L., & Mahoney, W. (2009). Brief Report: Relationship between non-verbal IQ and gender in autism. *Journal of Autism and Developmental Disorders*, 39(1), 188–193. <https://doi.org/10.1007/s10803-008-0612-4>
- Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *The British Journal of Psychiatry*, 161, 839–843.
- Barrett, B., Mosweu, I., Jones, C. R., Charman, T., Baird, G., Simonoff, E., ... Byford, S. (2015). Comparing service use and costs among adolescents with autism spectrum disorders, special needs and typical development. *Autism*, 19(5), 562–569. <https://doi.org/10.1177/1362361314536626>
- Bayley, N. (2006). *Bayley scales of infant and toddler development* (3rd ed.). San Antonio, Texas: Harcourt Assessment, Inc.
- Bethell, C., Reuland, C., Schor, E., Abrahms, M., & Halfon, N. (2011). Rates of parent-centered developmental screening: Disparities and links to services access. *Pediatrics*, 128(1), 146–155. <https://doi.org/10.1542/peds.2010-0424>
- Bossuyt, P., Davenport, C., Deeks, J., Hyde, C., Leeflang, M., & Scholten, R. (2013). Chapter 11: Interpreting results and drawing conclusions. In *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (version 0.9). The Cochrane Collaboration.

- Boycott, K., Hartley, T., Adam, S., Bernier, F., Chong, K., Fernandez, B. A., ... on behalf of the Canadian College of Medical Geneticists. (2015). The clinical application of genome-wide sequencing for monogenic diseases in Canada: Position statement of the Canadian College of Medical Geneticists. *Journal of Medical Genetics*, 52(7), 431–437.
<https://doi.org/10.1136/jmedgenet-2015-103144>
- British Columbia Psychological Association. (2014). Suggested current market rate 2014-15. British Columbia Psychological Association.
- Bryson, S. E., Rogers, S. J., & Fombonne, E. (2003). Autism spectrum disorders: early detection, intervention, education, and psychopharmacological management. *Canadian Journal of Psychiatry*, 48(8), 506–516.
- Buescher, A. V. S., Cidav, Z., Knapp, M., & Mandell, D. S. (2014). Costs of autism spectrum disorders in the United Kingdom and the United States. *JAMA Pediatrics*, 168(8), 721.
<https://doi.org/10.1001/jamapediatrics.2014.210>
- Cairney, J., Clinton, J., Veldhuizen, S., Rodriguez, C., Missiuna, C., Wade, T., ... Kertoy, M. (2016). Evaluation of the revised Nipissing District Developmental Screening (NDDS) tool for use in general population samples of infants and children. *BMC Pediatrics*, 16(1).
<https://doi.org/10.1186/s12887-016-0577-y>
- Canal-Bedia, R., García-Primo, P., Martín-Cilleros, M. V., Santos-Borbujo, J., Guisuraga-Fernández, Z., Herráez-García, L., ... Posada-de la Paz, M. (2011). Modified Checklist for Autism in Toddlers: Cross-cultural adaptation and validation in Spain. *Journal of Autism and Developmental Disorders*, 41(10), 1342–1351. <https://doi.org/10.1007/s10803-010-1163-z>
- Carter, M., & Scherer, S. (2013). Autism spectrum disorder in the genetics clinic: a review: Autism spectrum disorder in the genetics clinic. *Clinical Genetics*, 83(5), 399–407.
<https://doi.org/10.1111/cge.12101>
- CCMG Cytogenetics Committee. (2010). CCMG guidelines for genome microarray testing. Canadian College of Medical Geneticists.
- Centers for Disease Control and Prevention. (2015). Developmental milestone checklist. Division of Birth Defects, National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention. Retrieved from <http://www.cdc.gov/ncbddd/actearly/milestones/>
- Centers for Disease Control and Prevention. (2016). Prevalence and characteristics of autism spectrum disorder among children aged 8 years- Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States 2012. *Morbidity and Mortality Weekly Report*, 65(3).
- Charman, T., Baird, G., Simonoff, E., Chandler, S., Davison-Jenkins, A., Sharma, A., ... Pickles, A. (2016). Testing two screening instruments for autism spectrum disorder in UK community child health services. *Developmental Medicine & Child Neurology*, 58(4), 369–375.
<https://doi.org/10.1111/dmcn.12874>
- Charman, T., Pickles, A., Simonoff, E., Chandler, S., Loucas, T., & Baird, G. (2011). IQ in children with autism spectrum disorders: data from the Special Needs and Autism Project (SNAP). *Psychological Medicine*, 41(3), 619–627.
<https://doi.org/10.1017/S0033291710000991>
- Chawarska, K., Klin, A., Paul, R., & Volkmar, F. (2007). Autism spectrum disorder in the second year: Stability and change in syndrome expression. *Journal of Child Psychology and Psychiatry*, 48(2), 128–138. <https://doi.org/10.1111/j.1469-7610.2006.01685.x>

- Chen, L. S., Xu, L., Huang, T.-Y., & Dhar, S. U. (2013). Autism genetic testing: a qualitative study of awareness, attitudes, and experiences among parents of children with autism spectrum disorders. *Genetics in Medicine, 15*(4), 274–281. <https://doi.org/10.1038/gim.2012.145>
- Chi, D. L., Momany, E. T., Jones, M. P., Kuthy, R. A., Askelson, N. M., Wehby, G. L., & Damiano, P. C. (2013). An explanatory model of factors related to well baby visits by age three years for Medicaid-enrolled infants: a retrospective cohort study. *BMC Pediatrics, 13*(1), 158. <https://doi.org/10.1186/1471-2431-13-158>
- Chlebowski, C., Robins, D. L., Barton, M. L., & Fein, D. (2013). Large-scale use of the Modified Checklist for Autism in low-risk toddlers. *Pediatrics, 131*(4), e1121–e1127. <https://doi.org/10.1542/peds.2012-1525>
- Ch'ng, C., Kwok, W., Rogic, S., & Pavlidis, P. (2015). Meta-analysis of gene expression in autism spectrum disorder: Meta-analysis of gene expression in ASD. *Autism Research, n/a-n/a*. <https://doi.org/10.1002/aur.1475>
- Chu, H., & Cole, S. R. (2006). Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *Journal of Clinical Epidemiology, 59*(12), 1331-1332-1333. <https://doi.org/10.1016/j.jclinepi.2006.06.011>
- Cidav, Z., Lawer, L., Marcus, S. C., & Mandell, D. S. (2013). Age-related variation in health service use and associated expenditures among children with autism. *Journal of Autism and Developmental Disorders, 43*(4), 924–931. <https://doi.org/10.1007/s10803-012-1637-2>
- Cidav, Z., Marcus, S. C., & Mandell, D. S. (2012). Implications of childhood autism for parental employment and earnings. *Pediatrics, 129*(4), 617–623. <https://doi.org/10.1542/peds.2011-2700>
- Cogan-Ferchalk, J. (2013). *The utility of the Modified Checklist for Autism in Toddlers in a preschool-age special education sample*. Indiana University of Pennsylvania, USA.
- Coo, H., Ouellette-Kuntz, H., Lam, M., Yu, C. T., Dewey, D., Bernier, F. P., ... Holden, J. J. (2012). Correlates of age at diagnosis of autism spectrum disorders in six Canadian regions. *Chronic Diseases and Injuries in Canada, 32*(2), 90–100.
- Croen, L. A., Najjar, D. V., Ray, G. T., Lotspeich, L., & Bernal, P. (2006). A comparison of health care utilization and costs of children with and without autism spectrum disorders in a large group-model health plan. *Pediatrics, 118*(4), e1203–e1211. <https://doi.org/10.1542/peds.2006-0127>
- Currie, L., Dodds, L., Shea, S., Flowerdew, G., McLean, J., Walker, R., & Vincer, M. (2012). Investigation of test characteristics of two screening tools in comparison with a gold standard assessment to detect developmental delay at 36 months: A pilot study. *Paediatrics & Child Health, 17*(10), 549–552.
- Dahinten, S., & Ford, L. (2004). Validation of the Nipissing District Developmental Screen For Use With Infants and Toddlers. *Consortium for Health, Intervention, Learning and Development*. Retrieved from [http://www.earlylearning.ubc.ca/documents/NDDS%204.1%20Dev%20Screening%20Research%20Report%20\(2004%2011%2015\)%20Final.pdf](http://www.earlylearning.ubc.ca/documents/NDDS%204.1%20Dev%20Screening%20Research%20Report%20(2004%2011%2015)%20Final.pdf)
- Daniels, A. M., & Mandell, D. S. (2013). Children's compliance with American Academy of Pediatrics' well-child care visit guidelines and the early detection of autism. *Journal of Autism and Developmental Disorders, 43*(12), 2844–2854. <https://doi.org/10.1007/s10803-013-1831-x>

- Dawson, G. (2016). Why it's important to continue universal autism screening while research fully examines its impact. *JAMA Pediatrics*, *170*(6), 527. <https://doi.org/10.1001/jamapediatrics.2016.0163>
- Dawson, S., Glasson, E. J., Dixon, G., & Bower, C. (2009). Birth defects in children with autism spectrum disorders: A population-based, nested case-control study. *American Journal of Epidemiology*, *169*(11), 1296–1303. <https://doi.org/10.1093/aje/kwp059>
- Deeks, J., Bossuyt, P., & Gatsonis, C. (2013). *Cochrane handbook for systematic reviews of diagnostic test accuracy*. The Cochrane Collaboration.
- Devlin, B., & Scherer, S. W. (2012). Genetic architecture in autism spectrum disorder. *Current Opinion in Genetics & Development*, *22*(3), 229–237. <https://doi.org/10.1016/j.gde.2012.03.002>
- Dosreis, S., Weiner, C. L., Johnson, L., & Newschaffer, C. J. (2006). Autism spectrum disorder screening and management practices among general pediatric providers. *Journal of Developmental and Behavioral Pediatrics*, *27*(2 Suppl), S88-94.
- Duby, J. C., & Johnson, C. P. (2009). Universal screening for autism spectrum disorders: A snapshot within the big picture. *Pediatric Annals*, *38*(1), 36–41.
- Eaves, L. C. (2006). Screening for autism: Agreement with diagnosis. *Autism*, *10*(3), 229–242. <https://doi.org/10.1177/1362361306063288>
- Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child and Adolescent Psychology*, *38*(3), 439–450. <https://doi.org/10.1080/15374410902851739>
- Elsabbagh, M., & Johnson, M. H. (2010). Getting answers from babies about autism. *Trends in Cognitive Sciences*, *14*(2), 81–87. <https://doi.org/10.1016/j.tics.2009.12.005>
- Fein, D. (2016). Commentary on USPSTF final statement on universal screening for autism: *Journal of Developmental & Behavioral Pediatrics*, *37*(7), 573–578. <https://doi.org/10.1097/DBP.0000000000000345>
- Fessenden, V. M. (2013). *Comparison of the Pervasive Developmental Disorders Screening Test and Modified Checklist for Autism in Toddlers: Which is the better predictor of autism in toddlers?* University of Nevada, Nevada, USA.
- Filipek, P. A., Accardo, P. J., Ashwal, S., Baranek, G. T., Cook, E. H., Dawson, G., ... Volkmar, F. R. (2000). Practice parameter: screening and diagnosis of autism: report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. *Neurology*, *55*(4), 468–479.
- Fletcher-Watson, S., McConnell, F., Manola, E., & McConachie, H. (2014). Interventions based on the Theory of Mind cognitive model for autism spectrum disorder (ASD). In The Cochrane Collaboration (Ed.), *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd. Retrieved from <http://doi.wiley.com/10.1002/14651858.CD008785.pub2>
- Fombonne, E., Rogé, B., Claverie, J., Courty, S., & Frémolle, J. (1999). Microcephaly and macrocephaly in autism. *Journal of Autism and Developmental Disorders*, *29*(2), 113–119.
- Foster, M. W., Mulvihill, J. J., & Sharp, R. R. (2009). Evaluating the utility of personal genomic information. *Genetics in Medicine*, *11*(8), 570–574. <https://doi.org/10.1097/GIM.0b013e3181a2743e>
- Frenette, P., Dodds, L., MacPherson, K., Flowerdew, G., Hennen, B., & Bryson, S. (2013). Factors affecting the age at diagnosis of autism spectrum disorders in Nova Scotia, Canada. *Autism*, *17*(2), 184–195. <https://doi.org/10.1177/1362361311413399>

- Ganz, M. L. (2007). The lifetime distribution of the incremental societal costs of autism. *Archives of Pediatrics & Adolescent Medicine*, 161(4), 343–349. <https://doi.org/10.1001/archpedi.161.4.343>
- Goodwin, S. (2010). *M-CHAT and PDDST-2 as predictors of autism spectrum disorders in young children living in rural Canada*. Fielding Graduate University, California, USA.
- Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders*, 37(4), 613–627. <https://doi.org/10.1007/s10803-006-0280-1>
- Green, R. C., Berg, J. S., Grody, W. W., Kalia, S. S., Korf, B. R., Martin, C. L., ... American College of Medical Genetics and Genomics. (2013). ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genetics in Medicine*, 15(7), 565–574. <https://doi.org/10.1038/gim.2013.73>
- Grønberg, T. K., Schendel, D. E., & Parner, E. T. (2013). Recurrence of autism spectrum disorders in full- and half-siblings and trends over time: A population-based cohort study. *JAMA Pediatrics*, 167(10), 947. <https://doi.org/10.1001/jamapediatrics.2013.2259>
- Guevara, J. P., Gerdes, M., Localio, R., Huang, Y. V., Pinto-Martin, J., Minkovitz, C. S., ... Pati, S. (2013). Effectiveness of developmental screening in an urban setting. *Pediatrics*, 131(1), 30–37. <https://doi.org/10.1542/peds.2012-0765>
- Gura, G. F., Champagne, M. T., & Blood-Siegfried, J. E. (2011). Autism spectrum disorder screening in primary care. *Journal of Developmental & Behavioral Pediatrics*, 32(1), 48–51. <https://doi.org/10.1097/DBP.0b013e3182040aea>
- Gurrieri, F. (2012). Working up autism: The practical role of medical genetics. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, 160C(2), 104–110. <https://doi.org/10.1002/ajmg.c.31326>
- Guthrie, W., Swineford, L. B., Nottke, C., & Wetherby, A. M. (2013). Early diagnosis of autism spectrum disorder: Stability and change in clinical diagnosis and symptom presentation. *Journal of Child Psychology and Psychiatry*, 54(5), 582–590. <https://doi.org/10.1111/jcpp.12008>
- Hagan, J. F., Shaw, S. M., & Duncan, P. (2008). *Bright futures: Guidelines for health supervision of infants, children, and adolescents. Pocket Guide* (3rd ed.). Elk Grove Village, IL: American Academy of Pediatrics.
- Health Quality Ontario. (2016). Ontario Genetic Advisory Committee terms of reference version 3. Health Quality Ontario.
- Higgins, J. P. T. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327(7414), 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
- Horlin, C., Falkmer, M., Parsons, R., Albrecht, M. A., & Falkmer, T. (2014). The cost of autism spectrum disorders. *PLoS ONE*, 9(9), e106552. <https://doi.org/10.1371/journal.pone.0106552>
- Hospital for Sick Children. (2015). Molecular genetic lab- Billing and CPT codes. Hospital for Sick Children.
- Huerta, M., Bishop, S. L., Duncan, A., Hus, V., & Lord, C. (2012). Application of DSM-5 criteria for autism spectrum disorder to three samples of children with DSM-IV diagnoses of pervasive developmental disorders. *The American Journal of Psychiatry*, 169(10), 1056–1064. <https://doi.org/10.1176/appi.ajp.2012.12020276>
- Iossifov, I., O’Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., ... Wigler, M. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature*, 515(7526), 216–221. <https://doi.org/10.1038/nature13908>

- Järbrink, K. (2007). The economic consequences of autistic spectrum disorder among children in a Swedish municipality. *Autism, 11*(5), 453–463. <https://doi.org/10.1177/1362361307079602>
- Järbrink, K., Fombonne, E., & Knapp, M. (2003). Measuring the parental, service and cost impacts of children with autistic spectrum disorder: a pilot study. *Journal of Autism and Developmental Disorders, 33*(4), 395–402.
- Jhanjee, I., Saxeena, D., Arora, J., & Gjerdingen, D. K. (2004). Parents' health and demographic characteristics predict noncompliance with well-child visits. *The Journal of the American Board of Family Practice, 17*(5), 324–331.
- Jiang, Y., Yuen, R. K. C., Jin, X., Wang, M., Chen, N., Wu, X., ... Scherer, S. W. (2013). Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. *American Journal of Human Genetics, 93*(2), 249–263. <https://doi.org/10.1016/j.ajhg.2013.06.012>
- Jo, H., Schieve, L. A., Rice, C. E., Yeargin-Allsopp, M., Tian, L. H., Blumberg, S. J., ... Boyle, C. A. (2015). Age at autism spectrum disorder (ASD) diagnosis by race, ethnicity, and primary household language among children with special health care needs, United States, 2009–2010. *Maternal and Child Health Journal, 19*(8), 1687–1697. <https://doi.org/10.1007/s10995-015-1683-4>
- Johnson, C. P., Myers, S. M., & the Council on Children With Disabilities. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics, 120*(5), 1183–1215. <https://doi.org/10.1542/peds.2007-2361>
- Jones, E. J. H., Gliga, T., Bedford, R., Charman, T., & Johnson, M. H. (2014). Developmental pathways to autism: A review of prospective studies of infants at risk. *Neuroscience & Biobehavioral Reviews, 39*, 1–33. <https://doi.org/10.1016/j.neubiorev.2013.12.001>
- Kalia, S. S., Adelman, K., Bale, S. J., Chung, W. K., Eng, C., Evans, J. P., ... Miller, D. T. (2016). Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): A policy statement of the American College of Medical Genetics and Genomics. *Genetics in Medicine. https://doi.org/10.1038/gim.2016.190*
- Kamio, Y., Inada, N., Koyama, T., Inokuchi, E., Tsuchiya, K., & Kuroda, M. (2014). Effectiveness of using the Modified Checklist for Autism in Toddlers in two-stage screening of autism spectrum disorder at the 18-month health check-up in Japan. *Journal of Autism and Developmental Disorders, 44*(1), 194–203. <https://doi.org/10.1007/s10803-013-1864-1>
- Kara, B., Mukaddes, N. M., Alt nkaya, I., Guntepe, D., Gokcay, G., & Ozmen, M. (2014). Using the Modified Checklist for Autism in Toddlers in a well-child clinic in Turkey: Adapting the screening method based on culture and setting. *Autism, 18*(3), 331–338. <https://doi.org/10.1177/1362361312467864>
- Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., ... Fein, D. (2008). The Modified Checklist for Autism in Toddlers: A follow-up study investigating the early detection of autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*(5), 827–839. <https://doi.org/10.1007/s10803-007-0450-9>
- Koh, H. C., Lim, S. H., Chan, G. J., Lin, M. B., Lim, H. H., Choo, S. H. T., & Magiati, I. (2014). The clinical utility of the Modified Checklist for Autism in Toddlers with high risk 18–48 month old children in Singapore. *Journal of Autism and Developmental Disorders, 44*(2), 405–416. <https://doi.org/10.1007/s10803-013-1880-1>
- Lai, M. C., Lombardo, M. V., & Baron-Cohen, S. (2014). Autism. *The Lancet, 383*(9920), 896–910. [https://doi.org/10.1016/S0140-6736\(13\)61539-1](https://doi.org/10.1016/S0140-6736(13)61539-1)

- Lainhart, J. E., Bigler, E. D., Bocian, M., Coon, H., Dinh, E., Dawson, G., ... Volkmar, F. (2006). Head circumference and height in autism: a study by the Collaborative Program of Excellence in Autism. *American Journal of Medical Genetics. Part A*, *140*(21), 2257–2274. <https://doi.org/10.1002/ajmg.a.31465>
- Le Couteur, A., Haden, G., Hammal, D., & McConachie, H. (2008). Diagnosing autism spectrum disorders in pre-school children using two standardised assessment instruments: The ADI-R and the ADOS. *Journal of Autism and Developmental Disorders*, *38*(2), 362–372. <https://doi.org/10.1007/s10803-007-0403-3>
- Leslie, D. L., & Martin, A. (2007). Health care expenditures associated with autism spectrum disorders. *Archives of Pediatrics & Adolescent Medicine*, *161*(4), 350–355. <https://doi.org/10.1001/archpedi.161.4.350>
- Levy, S. E., Mandell, D. S., & Schultz, R. T. (2009). Autism. *The Lancet*, *374*(9701), 1627–1638. [https://doi.org/10.1016/S0140-6736\(09\)61376-3](https://doi.org/10.1016/S0140-6736(09)61376-3)
- Limbos, M. M., Joyce, D. P., & Roberts, G. J. (2010). Nipissing District Developmental Screen: patterns of use by physicians in Ontario. *Canadian Family Physician*, *56*(2), e66–72.
- Limperopoulos, C., Bassan, H., Sullivan, N. R., Soul, J. S., Robertson, R. L., Moore, M., ... du Plessis, A. J. (2008). Positive screening for autism in ex-preterm infants: Prevalence and risk factors. *Pediatrics*, *121*(4), 758–765. <https://doi.org/10.1542/peds.2007-2158>
- Lintas, C., & Persico, A. M. (2008). Autistic phenotypes and genetic testing: state-of-the-art for the clinical geneticist. *Journal of Medical Genetics*, *46*(1), 1–8. <https://doi.org/10.1136/jmg.2008.060871>
- Liptak, G. S., Stuart, T., & Auinger, P. (2006). Health care utilization and expenditures for children with autism: Data from U.S. national samples. *Journal of Autism and Developmental Disorders*, *36*(7), 871–879. <https://doi.org/10.1007/s10803-006-0119-9>
- Lord, C., Luyster, R., Gotham, K., & Guthrie, W. (2012). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part II): Toddler Module*. Torrance, CA: Western Psychological Services.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., ... Rutter, M. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205–223.
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part I): Modules 1–4*. Torrance, CA: Western Psychological Services.
- Lord, C., Volkmar, F. R., DiLavore, P. S., & Risi, S. (1999). *Autism Diagnostic Observation Schedule (ADOS)*. Los Angeles, CA: Western Psychological Services.
- Mandell, D. S., & Mandy, W. (2015). Should all young children be screened for autism spectrum disorder? *Autism*, *19*(8), 895–896.
- Mandell, D. S., Novak, M. M., & Zubritsky, C. D. (2005). Factors associated with age of diagnosis among children with autism spectrum disorders. *Pediatrics*, *116*(6), 1480–1486. <https://doi.org/10.1542/peds.2005-0185>
- Mandell, D. S., Wiggins, L. D., Carpenter, L. A., Daniels, J., DiGuseppi, C., Durkin, M. S., ... Kirby, R. S. (2009). Racial/ethnic disparities in the identification of children with autism spectrum disorders. *American Journal of Public Health*, *99*(3), 493–498. <https://doi.org/10.2105/AJPH.2007.131243>

- Manning, M., Hudgins, L., & Professional Practice and Guidelines Committee. (2010). Array-based technology and recommendations for utilization in medical genetics practice for detection of chromosomal abnormalities. *Genetics in Medicine*, *12*(11), 742–745. <https://doi.org/10.1097/GIM.0b013e3181f8baad>
- Marshall, C. R., Noor, A., Vincent, J. B., Lionel, A. C., Feuk, L., Skaug, J., ... Scherer, S. W. (2008). Structural variation of chromosomes in autism spectrum disorder. *The American Journal of Human Genetics*, *82*(2), 477–488. <https://doi.org/10.1016/j.ajhg.2007.12.009>
- MATLAB. (2017). SimEvent (Version R2017a) [MathWorks]. Natick, MA: MathWorks Inc.
- Mavranzouli, I., Megnin-Viggars, O., Cheema, N., Howlin, P., Baron-Cohen, S., & Pilling, S. (2014). The cost-effectiveness of supported employment for adults with autism in the United Kingdom. *Autism*, *18*(8), 975–984. <https://doi.org/10.1177/1362361313505720>
- Mazefsky, C. A., & Oswald, D. P. (2006). The discriminative ability and diagnostic utility of the ADOS-G, ADI-R, and GARS for children in a clinical setting. *Autism*, *10*(6), 533–549. <https://doi.org/10.1177/1362361306068505>
- McGrew, S. G., Peters, B. R., Crittendon, J. A., & Veenstra-Vanderweele, J. (2012). Diagnostic yield of chromosomal microarray analysis in an autism primary care practice: which guidelines to implement? *Journal of Autism and Developmental Disorders*, *42*(8), 1582–1591. <https://doi.org/10.1007/s10803-011-1398-3>
- McPartland, J. C., Reichow, B., & Volkmar, F. R. (2012). Sensitivity and specificity of proposed DSM-5 diagnostic criteria for autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(4), 368–383. <https://doi.org/10.1016/j.jaac.2012.01.007>
- Miles, J. H. (2011). Autism spectrum disorders—A genetics review. *Genetics in Medicine*, *13*(4), 278–294. <https://doi.org/10.1097/GIM.0b013e3181ff67ba>
- Miles, J. H., Takahashi, T. N., Bagby, S., Sahota, P. K., Vaslow, D. F., Wang, C. H., ... Farmer, J. E. (2005). Essential versus complex autism: Definition of fundamental prognostic subtypes. *American Journal of Medical Genetics Part A*, *135A*(2), 171–180. <https://doi.org/10.1002/ajmg.a.30590>
- Miller, D. T., Adam, M. P., Aradhya, S., Biasecker, L. G., Brothman, A. R., Carter, N. P., ... Ledbetter, D. H. (2010). Consensus statement: Chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *The American Journal of Human Genetics*, *86*(5), 749–764. <https://doi.org/10.1016/j.ajhg.2010.04.006>
- Ministry of Health and Long-Term Care. (2015). Schedule of benefits for physician services under the Health Insurance Act. Retrieved from http://www.health.gov.on.ca/english/providers/program/ohip/sob/physerv/physerv_mn.html
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting Items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Molloy, C. A., Murray, D. S., Akers, R., Mitchell, T., & Manning-Courtney, P. (2011). Use of the Autism Diagnostic Observation Schedule (ADOS) in a clinical setting. *Autism*, *15*(2), 143–162. <https://doi.org/10.1177/1362361310379241>
- Monroe, G. R., Frederix, G. W., Savelberg, S. M. C., de Vries, T. I., Duran, K. J., van der Smagt, J. J., ... van Haaften, G. (2016). Effectiveness of whole-exome sequencing and costs of the traditional diagnostic trajectory in children with intellectual disability. *Genetics in Medicine*, *18*(9), 949–956. <https://doi.org/10.1038/gim.2015.200>

- Montes, G., & Halterman, J. S. (2008). Association of childhood autism spectrum disorders and loss of family income. *Pediatrics*, *121*(4), e821–e826. <https://doi.org/10.1542/peds.2007-1594>
- Motiwala, S. S., Gupta, S., Lilly, M. B., Ungar, W. J., & Coyte, P. C. (2006). The Cost-Effectiveness of Expanding Intensive Behavioural Intervention to All Autistic Children in Ontario. *Healthcare Policy*, *1*(2), 135–151.
- Mullen, E. (1995). *Mullen Scales of Early Learning*. Circle Pines, MN: American Guidance Services, Inc.
- Myers, S. M., Johnson, C. P., & the Council on Children with Disabilities. (2007). Management of children with autism spectrum disorders. *Pediatrics*, *120*(5), 1162–1182. <https://doi.org/10.1542/peds.2007-2362>
- Nachshen, J., Garcin, N., Moxness, K., Tremblay, Y., Hutchinson, P., Lachance, A., ... Ruttle, P. (2008). Screening, assessment and diagnosis of autism spectrum disorder in young children: Canadian best practice guidelines. Miriam Foundation.
- Nachshen, J., Garcin, N., Moxness, K., Tremblay, Y., & Hutchison, P. (2008). *Screening assessment and diagnosis of autism spectrum disorders in young children: Canadian best practice guidelines*. Montreal, Quebec: Miriam Foundation.
- Nagy, P., Ryan, B., & Robinson, R. (2002). Nipissing instrument validation report, 2001-2002. In *Evaluation of the Healthy Babies, Healthy Children Program: Project report*. Ontario Ministry of Community, Family and Children's Services.
- National Institute for Health and Care Excellence. (2011). Autism diagnosis in children and young people: Recognition, referral and diagnosis of children and young people on the autism spectrum. National Institute for Health and Care Excellence.
- Newschaffer, C. J., Croen, L. A., Daniels, J., Giarelli, E., Grether, J. K., Levy, S. E., ... Windham, G. C. (2007). The epidemiology of autism spectrum disorders. *Annual Review of Public Health*, *28*(1), 235–258. <https://doi.org/10.1146/annurev.publhealth.28.021406.144007>
- Ontario Association of Speech-Language Pathologists and Audiologists. (2016). Recommended fee schedule for speech language pathology services. Ontario Association of Speech-Language Pathologists and Audiologists.
- Ontario Genetics Secretariat. (2014). Clinical genetics wait time study.
- Ontario Psychological Association. (2015). Guidelines for fees and billing practices.
- Oono, I. P., Honey, E. J., & McConachie, H. (2013). Parent-mediated early intervention for young children with autism spectrum disorders (ASD). In The Cochrane Collaboration (Ed.), *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd. Retrieved from <http://doi.wiley.com/10.1002/14651858.CD009774.pub2>
- Ouellette-Kuntz, H. M., Coo, H., Lam, M., Breitenbach, M. M., Hennessey, P. E., Jackman, P. D., ... Chung, A. M. (2014). The changing prevalence of autism in three regions of Canada. *Journal of Autism and Developmental Disorders*, *44*(1), 120–136. <https://doi.org/10.1007/s10803-013-1856-1>
- Ouellette-Kuntz, H. M., Coo, H., Lam, M., Yu, C. T., Breitenbach, M. M., Hennessey, P. E., ... Crews, L. R. (2009). Age at diagnosis of autism spectrum disorders in four regions of Canada. *Canadian Journal of Public Health*, *100*(4), 268–273.
- Ozonoff, S., Heung, K., Byrd, R., Hansen, R., & Hertz-Picciotto, I. (2008). The onset of autism: Patterns of symptom emergence in the first years of life. *Autism Research*, *1*(6), 320–328. <https://doi.org/10.1002/aur.53>

- Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., ... Stone, W. L. (2011). Recurrence risk for autism spectrum disorders: A baby siblings research consortium study. *Pediatrics*. <https://doi.org/10.1542/peds.2010-2825>
- PACT consortium, Barrett, B., Byford, S., Sharac, J., Hudry, K., Leadbitter, K., ... Green, J. (2011). Service and wider societal costs of very young children with autism in the UK. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-011-1393-8>
- Peacock, G., Amendah, D., Ouyang, L., & Grosse, S. D. (2012). Autism spectrum disorders and health care expenditures: the effects of co-occurring conditions. *Journal of Developmental and Behavioral Pediatrics*, 33(1), 2–8. <https://doi.org/10.1097/DBP.0b013e31823969de>
- Penner, M. (2016). *Policy analysis and evaluation of national clinician-reported practices for diagnosis of autism spectrum disorder*. University of Toronto, Toronto, Canada.
- Penner, M., Rayar, M., Bashir, N., Roberts, S. W., Hancock-Howard, R. L., & Coyte, P. C. (2015). Cost-Effectiveness Analysis Comparing Pre-diagnosis Autism Spectrum Disorder (ASD)-Targeted Intervention with Ontario's Autism Intervention Program. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-015-2447-0>
- Perera, H., Wijewardena, K., & Aluthwelage, R. (2009). Screening of 18-24-month-old children for autism in a semi-urban community in Sri Lanka. *Journal of Tropical Pediatrics*, 55(6), 402–405. <https://doi.org/10.1093/tropej/fmp031>
- Piccininni, C., Bisnaire, L., & Penner, M. (2017). Cost-effectiveness of wait time reduction for intensive behavioral intervention services in Ontario, Canada. *JAMA Pediatrics*, 171(1), 23. <https://doi.org/10.1001/jamapediatrics.2016.2695>
- Pierce, K., Courchesne, E., & Bacon, E. (2016). To screen or not to screen universally for autism is not the question: Why the task force got it wrong. *The Journal of Pediatrics*, 176, 182–194. <https://doi.org/10.1016/j.jpeds.2016.06.004>
- Pinto-Martin, J. A., Young, L. M., Mandell, D. S., Poghossyan, L., Giarelli, E., & Levy, S. E. (2008). Screening strategies for autism spectrum disorders in pediatric primary care. *Journal of Developmental & Behavioral Pediatrics*, 29(5), 345–350. <https://doi.org/10.1097/DBP.0b013e31818914cf>
- Powell, C. M. (2016). Autism screening or smoke screen and mirrors? *JAMA Neurology*, 73(4), 386. <https://doi.org/10.1001/jamaneurol.2016.0126>
- Radecki, L., Sand-Loud, N., O'Connor, K. G., Sharp, S., & Olson, L. M. (2011). Trends in the use of standardized tools for developmental screening in early childhood: 2002-2009. *Pediatrics*, 128(1), 14–19. <https://doi.org/10.1542/peds.2010-2180>
- Régie de l'assurance maladie du Québec. (2015). Manuel de facturation. rémunération à l'acte. Table B-Tarifification des visites. Régie de l'assurance maladie Quebec. Retrieved from <http://www.ramq.gouv.qc.ca/fr/professionnels/medecins-specialistes/manuels/Pages/facturation.aspx>
- Regier, D., Friedman, J., Makela, N., Ryan, M., & Marra, C. (2009). Valuing the benefit of diagnostic testing for genetic causes of idiopathic developmental disability: willingness to pay from families of affected children. *Clinical Genetics*, 75(6), 514–521. <https://doi.org/10.1111/j.1399-0004.2009.01193.x>
- Regier, D., Friedman, J., & Marra, C. (2010). Value for money? Array genomic hybridization for diagnostic testing for genetic causes of intellectual disability. *The American Journal of Human Genetics*, 86(5), 765–772. <https://doi.org/10.1016/j.ajhg.2010.03.009>
- Reichow, B., Barton, E. E., Boyd, B. A., & Hume, K. (2012). Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD). In *The*

- Cochrane Collaboration (Ed.), *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd. Retrieved from <http://doi.wiley.com/10.1002/14651858.CD009260.pub2>
- Reichow, B., & Wolery, M. (2009). Comprehensive synthesis of early intensive behavioral interventions for young children with autism based on the UCLA young autism project model. *Journal of Autism and Developmental Disorders*, *39*(1), 23–41. <https://doi.org/10.1007/s10803-008-0596-0>
- Richards, C., Jones, C., Groves, L., Moss, J., & Oliver, C. (2015). Prevalence of autism spectrum disorder phenomenology in genetic disorders: A systematic review and meta-analysis. *The Lancet Psychiatry*, *2*(10), 909–916. [https://doi.org/10.1016/S2215-0366\(15\)00376-4](https://doi.org/10.1016/S2215-0366(15)00376-4)
- Ringwood, K. (2010). *Predicting autism spectrum disorder in early childhood clinical assessments*. Alliant International University, California, USA.
- Robins, D. (2008). Screening for autism spectrum disorders in primary care settings. *Autism*, *12*(5), 537–556. <https://doi.org/10.1177/1362361308094502>
- Robins, D., Adamson, L. B., Barton, M., Connell, J. E., Dumont-Mathieu, T., Dworkin, P. H., ... Vivanti, G. (2016). Universal autism screening for toddlers: Recommendations at odds. *Journal of Autism and Developmental Disorders*, *46*(5), 1880–1882. <https://doi.org/10.1007/s10803-016-2697-5>
- Robins, D., Casagrande, K., Barton, M., Chen, C.-M. A., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the Modified Checklist for Autism in Toddlers, Revised With Follow-up (M-CHAT-R/F). *Pediatrics*, *133*(1), 37–45. <https://doi.org/10.1542/peds.2013-1813>
- Robins, D., Fein, D., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *31*(2), 131–144.
- Sanders, G. D., Neumann, P. J., Basu, A., Brock, D. W., Feeny, D., Krahn, M., ... Ganiats, T. G. (2016). Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: Second panel on cost-effectiveness in health and medicine. *JAMA*, *316*(10), 1093. <https://doi.org/10.1001/jama.2016.12195>
- Scarpa, A., Reyes, N. M., Patriquin, M. A., Lorenzi, J., Hassenfeldt, T. A., Desai, V. J., & Kerkering, K. W. (2013). The Modified Checklist for Autism in Toddlers: Reliability in a diverse rural American sample. *Journal of Autism and Developmental Disorders*, *43*(10), 2269–2279. <https://doi.org/10.1007/s10803-013-1779-x>
- Schaefer, G. B., & Mendelsohn, N. J. (2013). Clinical genetics evaluation in identifying the etiology of autism spectrum disorders: 2013 guideline revisions. *Genetics in Medicine*, *15*(5), 399–407. <https://doi.org/10.1038/gim.2013.32>
- Scherer, S. W., & Dawson, G. (2011). Risk factors for autism: Translating genomic discoveries into diagnostics. *Human Genetics*, *130*(1), 123–148. <https://doi.org/10.1007/s00439-011-1037-2>
- Seif Eldin, A., Habib, D., Noufal, A., Farrag, S., Bazaid, K., Al-Sharbati, M., ... Gaddour, N. (2008). Use of M-CHAT for a multinational screening of young children with autism in the Arab countries. *International Review of Psychiatry*, *20*(3), 281–289. <https://doi.org/10.1080/09540260801990324>
- Shashi, V., McConkie-Rosell, A., Rosell, B., Schoch, K., Vellore, K., McDonald, M., ... Goldstein, D. B. (2014). The utility of the traditional medical genetics diagnostic evaluation in the context of next-generation sequencing for undiagnosed genetic disorders. *Genetics in Medicine*, *16*(2), 176–182. <https://doi.org/10.1038/gim.2013.99>

- Shea, L., Newschaffer, C. J., Xie, M., Myers, S. M., & Mandell, D. S. (2014). Genetic testing and genetic counseling among Medicaid-enrolled children with autism spectrum disorder in 2001 and 2007. *Human Genetics*, *133*(1), 111–116. <https://doi.org/10.1007/s00439-013-1362-8>
- Shen, Y., Dies, K. A., Holm, I. A., Bridgemohan, C., Sobeih, M. M., Caronna, E. B., ... Autism Consortium Clinical Genetics/DNA Diagnostics Collaboration. (2010). Clinical genetic testing for patients with autism spectrum disorders. *Pediatrics*, *125*(4), e727-735. <https://doi.org/10.1542/peds.2009-1684>
- Sherman, S., Pletcher, B. A., & Driscoll, D. A. (2005). Fragile X syndrome: Diagnostic and carrier testing. *Genetics in Medicine*, *7*(8), 584–587. <https://doi.org/10.1097/01.GIM.0000182468.22666.dd>
- Silverstein, M., & Radesky, J. (2016). Embrace the complexity: The US Preventive Services Task Force recommendation on screening for autism spectrum disorder. *JAMA*, *315*(7), 661. <https://doi.org/10.1001/jama.2016.0051>
- Siu, A. L., and the US Preventive Services Task Force (USPSTF), Bibbins-Domingo, K., Grossman, D. C., Baumann, L. C., Davidson, K. W., ... Pignone, M. P. (2016). Screening for autism spectrum disorder in young children: US Preventive Services Task Force recommendation statement. *JAMA*, *315*(7), 691. <https://doi.org/10.1001/jama.2016.0018>
- Sivberg, B. (2003). Parents' detection of early signs in their children having an autistic spectrum disorder. *Journal of Pediatric Nursing*, *18*(6), 433–439.
- Snow, A. V., & Lecavalier, L. (2008). Sensitivity and specificity of the Modified Checklist for Autism in Toddlers and the Social Communication Questionnaire in preschoolers suspected of having pervasive developmental disorders. *Autism*, *12*(6), 627–644. <https://doi.org/10.1177/1362361308097116>
- Sparrow, S., Balla, D., & Cicchetti, D. (2005). *Vineland Adaptive Behavior Scales* (2nd ed.). Bloomington, MN: Pearson Assessment.
- Statistics Canada. (2015a). Table 102-4509. Live births, by birth weight and sex, Canada, provinces and territories. Retrieved from <http://www5.statcan.gc.ca/cansim/a26>
- Statistics Canada. (2015b). Table 282-0074. Labour force survey estimates (LFS), wages of employees by job permanence, union coverage, sex and age group. Retrieved from <http://www5.statcan.gc.ca/cansim/a26?lang=eng&id=2820074>
- Sunita, & Bilszta, J. L. (2013). Early identification of autism: A comparison of the Checklist for Autism in Toddlers and the Modified Checklist for Autism in Toddlers. *Journal of Paediatrics and Child Health*, *49*(6), 438–444. <https://doi.org/10.1111/j.1440-1754.2012.02558.x>
- Tammimies, K., Marshall, C. R., Walker, S., Kaur, G., Thiruvahindrapuram, B., Lionel, A. C., ... Fernandez, B. A. (2015). Molecular diagnostic yield of chromosomal microarray analysis and whole-exome sequencing in children with autism spectrum disorder. *JAMA*, *314*(9), 895–903. <https://doi.org/10.1001/jama.2015.10078>
- Tek, S., & Landa, R. J. (2012). Differences in autism symptoms between minority and non-minority toddlers. *Journal of Autism and Developmental Disorders*, *42*(9), 1967–1973. <https://doi.org/10.1007/s10803-012-1445-8>
- The Centre for Applied Genomics. (2015). Cytogenomics and Genome Resources. Retrieved from <http://www.tcag.ca/facilities/cytogenomicsGenome.html>

- Tick, B., Bolton, P., Happé, F., Rutter, M., & Rijdsdijk, F. (2016). Heritability of autism spectrum disorders: a meta-analysis of twin studies. *Journal of Child Psychology and Psychiatry*, *57*(5), 585–595. <https://doi.org/10.1111/jcpp.12499>
- Timonen-Soivio, L., Vanhala, R., Malm, H., Leivonen, S., Jokiranta, E., Hinkka-Yli-Salomäki, S., ... Sourander, A. (2015). The association between congenital anomalies and autism spectrum disorders in a Finnish national birth cohort. *Developmental Medicine & Child Neurology*, *57*(1), 75–80. <https://doi.org/10.1111/dmcn.12581>
- Trakadis, Y., & Shevell, M. (2011). Microarray as a first genetic test in global developmental delay: a cost-effectiveness analysis: Microarray cost-effectiveness in global developmental delay. *Developmental Medicine & Child Neurology*, *53*(11), 994–999. <https://doi.org/10.1111/j.1469-8749.2011.04080.x>
- TreeAge Software. (2015). *TreeAge Pro 2015*. Williamstown, MA: TreeAge Software.
- Tsiplova, K., Zur, R., Marshall, C. R., Stavropoulos, D. J., Pereira, S. L., Merico, D., ... Ungar, W. J. (2017). A microcosting and cost-consequence analysis of genomic testing strategies in autism spectrum disorder. *Genetics in Medicine*. <https://doi.org/10.1038/gim.2017.47>
- Valicenti-McDermott, M., Hottinger, K., Seijo, R., & Shulman, L. (2012). Age at diagnosis of autism spectrum disorders. *The Journal of Pediatrics*, *161*(3), 554–556. <https://doi.org/10.1016/j.jpeds.2012.05.012>
- Vande Wydeven, K., Kwan, A., Hardan, A. Y., & Bernstein, J. A. (2012). Underutilization of genetics services for autism: The importance of parental awareness and provider recommendation. *Journal of Genetic Counseling*, *21*(6), 803–813. <https://doi.org/10.1007/s10897-012-9494-x>
- Veenstra-VanderWeele, J., & McGuire, K. (2016). Rigid, inflexible approach results in no recommendation for autism screening. *JAMA Psychiatry*, *73*(4), 327. <https://doi.org/10.1001/jamapsychiatry.2016.0143>
- Verde, P. E. (2010). Meta-analysis of diagnostic test data: A bivariate Bayesian modeling approach. *Statistics in Medicine*, *29*(30), 3088–3102. <https://doi.org/10.1002/sim.4055>
- Vermeesch, J. R., Fiegler, H., de Leeuw, N., Szuhai, K., Schoumans, J., Ciccone, R., ... Zuffardi, O. (2007). Guidelines for molecular karyotyping in constitutional genetic diagnosis. *European Journal of Human Genetics*, *15*(11), 1105–1114. <https://doi.org/10.1038/sj.ejhg.5201896>
- Villalobos, M. E. (2011, December). *Screening for autism in toddlers: A follow-up study*. The University of Utah, Utah, USA.
- Volkmar, F., Siegel, M., Woodbury-Smith, M., King, B., McCracken, J., & State, M. (2014). Practice parameter for the assessment and treatment of children and adolescents with autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(2), 237–257. <https://doi.org/10.1016/j.jaac.2013.10.013>
- Wang, L., Mandell, D. S., Lawer, L., Cidav, Z., & Leslie, D. L. (2013). Healthcare service use and costs for autism spectrum disorder: A comparison between Medicaid and private insurance. *Journal of Autism and Developmental Disorders*, *43*(5), 1057–1064. <https://doi.org/10.1007/s10803-012-1649-y>
- Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-VanderWeele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics*, *127*(5), e1303–e1311. <https://doi.org/10.1542/peds.2011-0426>
- Warrier, V., Chee, V., Smith, P., Chakrabarti, B., & Baron-Cohen, S. (2015). A comprehensive meta-analysis of common genetic variants in autism spectrum conditions. *Molecular Autism*, *6*(1). <https://doi.org/10.1186/s13229-015-0041-0>

- Weiss, J., Whelan, M., McMorris, C., Carroll, C., & Canadian Autism Spectrum Disorders Alliance. (2014). *Autism in Canada: National needs assessments survey for families, individuals with autism spectrum disorder and professionals*.
- Wetherby, A., & Prizant, B. (2002). *Infant Toddler Checklist*. Baltimore, MD: Paul H. Brookes Publishing Co Inc.
- Wetterstrand, K. (2015). DNA sequencing costs: Data from the NHGRI Genome Sequencing Program (GSP). Retrieved from <http://www.genome.gov/sequencingcosts/>
- Whiting, P. F. (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 155(8), 529. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>
- Wier, M. L., Yoshida, C. K., Odouli, R., Grether, J. K., & Croen, L. A. (2006). Congenital anomalies associated with autism spectrum disorders. *Developmental Medicine and Child Neurology*, 48(6), 500–507. <https://doi.org/10.1017/S001216220600106X>
- Wiggins, L. D., Baio, J., & Rice, C. (2006). Examination of the time between first evaluation and first autism spectrum diagnosis in a population-based sample. *Journal of Developmental and Behavioral Pediatrics*, 27(2 Suppl), S79-87.
- Williams, R., Clinton, J., & Canadian Pediatric Society. (2011). Getting it right at 18 months: In support of an enhanced well-baby visit. Canadian Pediatric Society.
- Wilson, J. M., & Jungner, G. (1968). *Principles and practice of screening for disease*. Geneva: World Health Organization. Retrieved from http://apps.who.int/iris/bitstream/10665/37650/17/WHO_PHP_34.pdf
- Wordsworth, S., Buchanan, J., Regan, R., Davison, V., Smith, K., Dyer, S., ... Knight, S. J. L. (2007). Diagnosing idiopathic learning disability: a cost-effectiveness analysis of microarray technology in the National Health Service of the United Kingdom. *Genomic Medicine*, 1(1–2), 35–45. <https://doi.org/10.1007/s11568-007-9005-6>
- World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization.
- Xu, J., Zwaigenbaum, L., Szatmari, P., & Scherer, S. (2004). Molecular cytogenetics of autism. *Current Genomics*, 5(4), 347–364. <https://doi.org/10.2174/1389202043349246>
- Yama, B., Freeman, T., Graves, E., Yuan, S., & Karen Campbell, M. (2012). Examination of the properties of the Modified Checklist for Autism in Toddlers (M-CHAT) in a population sample. *Journal of Autism and Developmental Disorders*, 42(1), 23–34. <https://doi.org/10.1007/s10803-011-1211-3>
- Yeargin-Allsopp, M., Rice, C., Karapurkar, T., Doernberg, N., Boyle, C., & Murphy, C. (2003). Prevalence of autism in a US metropolitan area. *JAMA*, 289(1). <https://doi.org/10.1001/jama.289.1.49>
- Yu, T. W., Chahrour, M. H., Coulter, M. E., Jiralerspong, S., Okamura-Ikeda, K., Ataman, B., ... Walsh, C. A. (2013). Using whole-exome sequencing to identify inherited causes of autism. *Neuron*, 77(2), 259–273. <https://doi.org/10.1016/j.neuron.2012.11.002>
- Yuen, R. K. C., Thiruvahindrapuram, B., Merico, D., Walker, S., Tammimies, K., Hoang, N., ... Scherer, S. W. (2015). Whole-genome sequencing of quartet families with autism spectrum disorder. *Nature Medicine*, 21(2), 185–191. <https://doi.org/10.1038/nm.3792>
- Ziegler, A., Rudolph-Rothfeld, W., & Vonthein, R. (2017). Genetic testing for autism spectrum disorder is lacking evidence of cost-effectiveness: A systematic review. *Methods of Information in Medicine*, 56(2). <https://doi.org/10.3414/ME16-01-0082>

- Zimmerman, I., Steiner, V., & Pond, R. (2002). *Preschool Language Scale* (4th ed.). San Antonio, Texas: Pearson Assessment.
- Zuckerman, K. E., Lindly, O. J., & Sinche, B. K. (2015). Parental concerns, provider response, and timeliness of autism spectrum disorder diagnosis. *The Journal of Pediatrics*, *166*(6), 1431–1439.e1. <https://doi.org/10.1016/j.jpeds.2015.03.007>
- Zwaigenbaum, L., Bauman, M. L., Choueiri, R., Kasari, C., Carter, A., Granpeesheh, D., ... Natowicz, M. R. (2015). Early intervention for children with autism spectrum disorder under 3 years of age: Recommendations for practice and research. *Pediatrics*, *136*(Supplement), S60–S81. <https://doi.org/10.1542/peds.2014-3667E>
- Zwaigenbaum, L., Bauman, M. L., Fein, D., Pierce, K., Buie, T., Davis, P. A., ... Wagner, S. (2015). Early screening of autism spectrum disorder: Recommendations for practice and research. *Pediatrics*, *136*(Supplement), S41–S59. <https://doi.org/10.1542/peds.2014-3667D>
- Zwaigenbaum, L., Bryson, S. E., Szatmari, P., Brian, J., Smith, I. M., Roberts, W., ... Roncadin, C. (2012). Sex differences in children with autism spectrum disorder identified within a high-risk infant cohort. *Journal of Autism and Developmental Disorders*, *42*(12), 2585–2596. <https://doi.org/10.1007/s10803-012-1515-y>
- Zwaigenbaum, L., Bryson, S., Lord, C., Rogers, S., Carter, A., Carver, L., ... Yirmiya, N. (2009). Clinical assessment and management of toddlers with suspected autism spectrum disorder: Insights from studies of high-risk infants. *Pediatrics*, *123*(5), 1383–1391. <https://doi.org/10.1542/peds.2008-1606>