Cost-Utility Analysis of Prenatal Diagnosis of Congenital Cardiac Diseases using Deep Learning

Ginsberg Gm[1,2], Drukker L[3], Pollak U[4], and Mayer Brezis[5]

[1]Braun School of Public Health, Hebrew University -Hadassah Hospital
[2]HECON, Health Economics Consultancy
[3]Department of Obstetrics and Gynecology, Rabin Medical Center, Tel-Aviv University
[4]Faculty of Medicine, Pediatric Critical Care Sector, Hadassah University Medical Center, The Hebrew University of Jerusalem
[5]Affiliation not available

March 14, 2024

Cost-Utility Analysis of Prenatal Diagnosis of Congenital Cardiac Diseases using Deep Learning.

Ginsberg GM[1,2]*, Drukker L[3], Pollak U[4], Brezis M[5].

1. Braun School of Public Health, Hebrew University – Hadassah Hospital, Ein Kerem, Israel.
2. HECON, Health Economics Consultancy, Jerusalem, Israel
3. Department of Obstetrics and Gynecology, Rabin Medical Center, Tel-Aviv University, Israel
4. Pediatric Critical Care Sector, Hadassah University Medical Center and the Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, Israel.
5. Center for Quality and Safety, Hadassah Medical Center & Braun School of Public Health, Hebrew University, Jerusalem, Israel.

* Corresponding Author: gmginsberg@gmail.com

Word Count (main text): 4,886 Abstract: 286 words

References (main text): 102 Tables: 3

Appendices: 15

Cost-Utility Analysis of Prenatal Diagnosis of Congenital Cardiac Diseases using Deep Learning.

Abstract:

Background

Deep learning (DL) is a new technology that can assist prenatal ultrasound (US) in the detection of congenital heart disease (CHD) at the prenatal stage. Hence, an economic-epidemiologic evaluation (aka Cost-Utility Analysis) is required to assist policymakers in deciding whether to adopt the new technology.

Methods

The cost-utility ratios (CUR) were calculated for the current provision of US plus pulse oximetry (POX) and with DL-assisted ultrasound (DL-US) plus POX by means of a spreadsheet model integrating demographic, economic epidemiological, health service utilization, screening performance, survival and lifetime quality of life data based on the standard formula:

CUR = (Intervention Costs - Treatment Savings)/ Quality Adjusted Life Years (QALY) gained

US screening data were based on data from real-world operational routine reports (as opposed to research studies). The DL screening cost of 145 USD was based on Israeli US costs plus 20.54 USD for reading and recording screens.

Results

The addition of DL-US, which is associated with increased sensitivity (95% vs 58.1%), resulted in far fewer undiagnosed infants (16 vs 102 [or 2.9% vs 15.4% of the 560 and 659 births, respectively). Adoption of DL-US will add 1,204 QALYs. The increased screening costs of DL-US (23.2 million USD) are largely offset by decreased treatment costs (20.8 million NIS). Therefore, the new DL-US technology is considered "very cost-effective", costing only 6,441 NIS per QALY. For most performance combinations (sensitivity > 80%, specificity >90%), the adoption of DL-US is either cost effective or very cost effective. For specificities greater than 98% (with sensitivities above 94%), DL-US (& POX) is said to "dominate" US (& POX) by providing more QALYs at a lower cost.

Conclusion

Our exploratory CUA calculations indicate the feasibility of DL-US as being at least cost-effective.

Keywords: Prenatal Screening, Ultrasound, Congenital Cardiac Disease, Deep Learning, Cost-Utility Analysis

Cost-Utility Analysis of Prenatal Diagnosis of Congenital Cardiac Diseases using Deep Learning

Introduction.

Congenital heart diseases (CHD) are the most common type of congenital defect, accounting for nearly one-third of all major congenital anomalies (1,2). CHD, and most notably, critical CHD (cCHD), is the leading cause of mortality and morbidity from birth abnormalities worldwide (3,4), accounting for more than 200,000 deaths annually (3). In developed countries, more than half of the total cost attributed to all birth defects combined is currently associated with care of CHD (5).

CHD is considered major if it requires cardiac surgery or catheter intervention or results in death in the first year of life. It is defined as critical if these occur in the first 28 days of life (6). Critical CHD conditions include valvular atresia or severe stenosis, coarctation of the aorta, transposition of the great arteries, total anomalous pulmonary venous connection (4) and many others. Within this group, the outcome varies considerably, from a guarded outlook, such as in hypoplastic left heart syndrome (HLHS) or interruption of the aortic arch, to conditions with better outcomes, such as complete transposition of the great arteries (TGA) (6).

Primary prevention of CHD is possible to some extent via improved diabetic control, switching to nonteratogenic medicine for treating epilepsy and possibly iron and folic acid supplementation. Unfortunately, less than one half of CHD (especially minor CHD) are detected prenatally (7-19), although detection rates vary depending on the type of defect (17,19,20,21), the examiner Skill (22), and specific population (23,24).

Increasing detection via prenatal diagnosis of CHD (and subsequent possible timely treatment), should result in a lower morbidity and mortality (25-30,31), only partly due to possible elective terminations of pregnancy, Prenatal diagnosis allows for family preparation, facilitates counselling, shared decision-making, planning for optimal neonatal intervention and medical care after delivery (31,32), including the transfer of deliveries to a tertiary care center

with resources to manage critically ill newborns (14,33,34), resulting in fewer and less severe accompanying neurodevelopmental disabilities (20, 31) and improved childhood developmental milestones.

Almost 30% of newborns affected with CHD are diagnosed late (35) and are more likely to experience hemodynamic compromise, resulting in prolonged hypoxemia to vital organs. The resultant untimely medical-surgical intervention results in elevated morbidity and mortality rates, including irreversible pulmonary hypertension (36,37).

A study of a pediatric population with pulmonary hypertension reported high readmission rates and use of expensive intensive care unit resources (38). Overall children with CHD incur 23% of total hospitalization costs globally, while accounting for only 4.4% of all hospital admissions (39). Importantly, the distance between the place of birth and a cardiac center has been shown to be correlated with neonatal death rates (40). Clearly, delayed diagnosis of CHD imposes a large cost burden on health services.

Adding universal echocardiographic screening of newborns (to routine prenatal screening) is unlikely to be cost effective. This is due not only to the high screening costs associated with echocardiography but also to the diminished pool (because of initial prenatal screening) of as-yet undetected cardiac abnormalities. Adding low-cost universal pulse oximetry (POX) screening to newborns is more likely to be cost effective. A UK modelling study (4) reported an incremental cost of approximately 41,000 USD (at 2009 price levels) per timely diagnosis of POX and a routine clinical examination in a population in which antenatal screening for CHDs already existed.

Routine implementation of POX was expected to be cost-effective in many studies (4,41-46), including a Dutch study where homebirths were predominant (47). However, resultant potential treatment cost savings and quality of life improvements, which would have resulted in a full cost utility analysis, were rarely included in such studies. Likewise, many previous cost-effective ultrasound (US) studies were limited to reporting either the cost per detected CHD case (47-49) and/or to the diagnosis of a specific ailment, such as coarctation of the aorta (50). The cost per

detected case was as high as $113,000 USD (at 2012 price levels) in the USA (48), with an antenatal ultrasound that includes five cardiac axial screening views having the lowest cost per detected case (51,52).

For our study purposes, we defined severe congenital heart disease (sCHD) as a diagnosis of either critical or major CHD. For the sake of completeness in measuring all the potential benefits, we also included screening effects on Minor CHD (mCHD), which include ventricular septal defects, atrial septal defects and bicuspid aortic valves and are more challenging to diagnose prenatally, in addition to prenatal diagnosis possibly having little impact on morbidity and neonatal mortality.

Recently, artificial intelligence-driven deep learning has been explored as a complement to and as an enhancement of routine US (referred to as DL-US) through its ability to increase the sensitivity of prenatal discovery of CHD (53,54). As a guide for policy-makers policymakers in deciding whether to adopt the new technology (DL-US), this study aimed to carry out full cost-utility analyses (CUAs) of various combinations of US (see Appendix Ia for a fuller description), POX (Appendix Ib) and artificial intelligence-driven DL-US [see Appendix Ic] modelling the many diagnoses, specific survival gains, quality of life gains and treatment costs.

METHODS

The Cost-Utility Ratio (CUR) was based on applying the interventions to the Israeli population on a national level. CURs (compared to the "null" of no screening) of various combinations of US, POX and DL-US, were calculated by means of an Excel based spreadsheet model that integrated demographic (55-57), economic (58-63), epidemiologic (64-67), screening efficacy (41,68-70), health service utilization (71), survival (57,58) and quality of life (51,58,72-75) data based on the standard formula:

$$\text{CUR} = \frac{\text{Intervention Costs - Treatment Savings}}{\text{Averted Quality Adjusted Life Year (QALY) losses due to Mortality and Morbidity}}$$

All costs are in USD (at 2022 price levels), based on the average exchange rate of 3.36 NIS to USD (76). Future costs and utilities were discounted using a rate of 3% per annum. In the absence of Israeli specific guidelines, interventions were deemed to be cost saving, very cost effective, cost effective, or not cost effective when treatment savings exceeded intervention costs, CUR < GDP (Gross Domestic Product) per capita, GDP per capita <CUR < 3 x GDP per capita and CUR> 3 x GDP per capita (77), based on Israel's GDP per capita in 2022 of approximately 54,800 USD (78). The full details of the complex modelling methodology are described in Appendix III.

We assumed a "baseline" based on the few reported US studies (68-70) that were carried out in routine settings in busy primary care units. These were characterized by lower standard operators working and devoting less than adequate time to the US. These will be subsequently referred to as "routine reports", which contrast with higher standard US reports carried out under "research study" conditions that are characterized by prospective supervised academic research in referral centers.

In this model, the baseline DL-US sensitivity and specificity were assumed to be 95% and 96%, respectively. The baseline definition enabled the exploration of the CUR of all the potential strategies, viz: null (i.e., doing nothing), US, POX at birth, US plus POX, DL-US, and DL-US plus POX.

Our major focus was evaluating a possible future operational change where the current operational screening of US plus POX would be replaced in the future by DL-US plus POX. Since there is a dearth of studies reporting operational data for DL-US, we ran the model over a wide range of expected DL-US sensitivities (from 80% to 99%) and expected specificities (from 90% to 100%) for the following three scenarios:

A. Routine: Based on data from the few "routine reports" of US studies that were based on actual real operational data. A cost per DL-US screen of $144.82 was based on Israeli US costs of $124.42 plus $20.40 for reading and recording the screen.
B. Routine High Cost: Based on data from the few "routine reports" of US studies. The cost per DL-US screen of $248.84 was assumed to be double that of US screens to reflect the pricing of the new technology to cover development costs.

C. Routine High Costs & High Performance: Based on data from the numerous US performed under "research study" conditions, that reported better operational data (i.e. higher sensitivities) than did those reported from the few real-life "routine reports" of retrospective studies. To achieve these higher operational standards, we assumed that double the amount of time would be allocated for the US screen (costing $248.84) plus an additional 25% of the original time for extra supervision ($31.25), for a total screen cost of $280.09. A cost per DL-US screen of $300.57 was based on the $280.09 US cost plus $20.48 for reading and recording the screen.

RESULTS.

Demographics

Based on a backwards calculation from birth data, in 2022, there were an estimated 199,935 pregnancies, with an early pregnancy loss of 12% (55) resulting in 175,943 viable pregnancies by the end of the first trimester, when the nuchal translucency scan is offered and taken up by nearly all women in Israel. There were an additional 3,151 elective terminations of pregnancies (55,78,79), 3% (55) foetal losses after the US and 0.345% stillbirths (78), resulting in 167,031 birth episodes and 181,269 new-borns (56).

Survival in CHD patients

A sample of just over half of all sCHDs in Israel was used (Appendix IV). Weighted survival rates based on prenatal diagnosis were non-significantly greater than those based on postnatal diagnosis (88.3% vs 87.1%; not sig). However, a survival advantage was found in favor of prenatal (vs. postnatal) diagnosis for several but not all CHDs:- Left heart obstruction (93.3% vs 80.9%; not sig), HLHS (71.1% vs 61.8%; $p<.001$) and TGA (96.2% vs 92.0%; $p< .001$). Conversely, for truncus arteriosus survival, there was a paradoxically lower survival rate for prenatal diagnosis (57.9% vs 91.5%; $p< .0001$), - possibly because more severe conditions may be more easily detected in utero. The postnatal survival rate was split into 87.2% and 87.0% for diagnoses $\leq$24 hours or > 24 hours, respectively.

For mCHD patients, the one-year survival rates were 93.4%, 96.44% and 96.40% for prenatal, postnatal <=24 hours and postnatal > 24 hours, respectively.

Treatment Costs

For the first year of life, treatment costs were $13,657 and $8,232 for sCHD and mCHD, respectively. The lifetime discount costs for sCHD patients diagnosed prenatally, < 24 hours and >= 24 hours were $220,570, $214,249 and $213,259, respectively, for males but were greater for females, $242,294, $236,014 and $235,551, respectively (due to increased life expectancy).

For mCHD, the discounted lifetime treatment costs were $163,672, $186,079, and $185,646 for males and $176,843, $198,636 and $198,216 for females diagnosed prenatally, < 24 hours and >= 24 hours, respectively.

Screening Performance

Three "routine reports" (from 2015-23) for sCHD reported (68-70) sensitivities ranging from 33.3% to 79.3% (weighted average 58.1%), alongside reported specificities of 100%. This performance was far lower than the 79.9% sensitivity, and a similar 99.95% specificity that were found in many publications (80) based on the use of the US and carried out under "research study" conditions (see Appendix V).

For mCHD (from 2015-2023), we excluded the two lone sensitivities of mCHD from "routine reports" due to lack of homogeneity (reporting 50% and 2.7% sensitivities). Instead, our model estimated a sensitivity for mCHD of 23.0%, based on the relative magnitudes of sensitivity for sCHD reported under "research study" (58.1%) and "routine reports" (79.9%) conditions multiplied by 31.6%. being the sensitivity for mCHD under "research study" conditions. (31.6% x 58.1%/79.9%). The Specificity of the "routine reports" was assumed to be the same as the results under "research study" conditions (99.97%) (Appendix VI).

For DL-US, our baseline screening sensitivity and specificity for sCHD were based on 95% and 96% respectively (72). The baseline sensitivity and specificity of DL-US for mCHD were assumed to be the same as for "routine" US, 23.0% and 99.7% respectively. For POX screening

at birth, the sensitivity and specificity for sCHD were 70.95% and 98.43% respectively (Appendix

VII).

Healthy adjusted life expectancy (HALE).

The resultant discounted (and undiscounted) HALE for males with sCHD was 14.17 (24.2), 13.83 (23.4) and 13.78 (23.2) for prenatal diagnosis, diagnosis ≤24 hours and diagnosis > 24 hours. For sCHD females, the HALEs were 15.14 (27.3), 14.82 (26.5) and 14.79, (26.4) for prenatal, ≤24 hours, > 24 hours diagnoses respectively.

Due to their lower average lifetime disability weights (DWs) (0.061 vs 0.241 for sCHD), HALES were greater for mCHD. For males, the discounted (and undiscounted) HALEs were 20.87 (38.3), 22.81 (44.7) and 22.77 (44.6) for prenatal diagnosis, diagnosis ≤24 hours and diagnosis > 24 hours respectively. For mCHD females, HALEs were 21.92 (42.2), 23.62 (48.2), and 23.59 (48.06) for prenatal diagnosis, diagnosis ≤24 hours and diagnosis > 24 hours. HALE losses were calculated by subtracting these from the average populations discounted (and undiscounted HALES of 29.66 (72.5) for males and 29.31 (73.0) for females (81).

Cost Utility Ratios (CUR)

In our base line situation, the assumed higher sensitivity (95%) and lower specificity (96%) of DL-US (with and without POX) generated elevated usage of electrocardiograms and elective abortions, respectively. However, the effect of different interventions on miscarriages and stillbirths was minimal (Table 1). When no screening was undertaken (Appendix VIII), the 905 sCHD fetuses that were viable at 12 weeks underwent 14 abortions, 87 miscarriages and 2 stillbirths, resulting in 802 live births with undiscovered sCHD (and similarly 1485 with mCHD). The use of only the US or POX alone led to 319 prenatal or 569 postnatal discoveries, respectively, of sCHD, resulting in 346 (48%) and 233 (29%) sCHD cases, respectively, being undiscovered before the infant was two days old (Appendix IX). The current Israeli practice of screening by both US & POX, results in only 102 (or 15%) undiscovered cases out of 659 live births with sCHD (Appendix VI). Use of DL-US has an expected higher sensitivity resulting in only 49 (8.7%) or 16 (2.9%) undiscovered cases with or without POX respectively.

Due to its inherent influence on the learning process, DL-US likely to eventually have a higher specificity than US alone. However, if DL-US has a lower relative specificity, this would result in higher abortion rates (Appendix VIII), which could cause the intervention costs of DL-US to be approximately 27% higher than those of US (Table 1), despite unit screening costs being only 6.1% higher (80). Again, the increased sensitivity of DL-US results in lower QALY losses from CHD. These are offset by the increased QALY losses from abortions due to the possible lower specificity (Table 2). All interventions (except for POX) are both cost saving and add QALYs compared to doing nothing ("the null") - that is, they "dominate" the null.

POX, on its own costs approximately $52,000 per QALY (Table 2), deeming it to be marginally very cost-effective. The recent introduction of POX to prenatal US, increased costs by $3,304,000, and added 31 QALYs at a cost-effective incremental cost effectiveness ratio (ICER) of $106,600 per added QALY.

Substituting DL-US (& POX) for the current US protocol (& POX) would cost an extra $2,308,000 but provide 1,204 more QALYs (Table 2) at a cost of $1,917 per QALY, which renders the intervention very cost-effective.

SENSITIVITY ANALYSES

Based on data from the few "routine reports" on US that were based on actual real operational data, Option A ("Routine DL-US"), Table 3A shows us where the advantage of DL-US (in terms of higher sensitivity) outweighs its possible disadvantage (due to possible lower specificities) versus US alone. Among all the combinations, where its sensitivity is >94%, DL-US (& POX) is either very cost-effective by providing more QALYs at a relatively low extra cost (see Appendix IX.A) or dominates US (& POX) by providing more QALYs at a lower cost.

In Scenario B ("Routine High DL-US Screening Cost") where the price of DL-US was double that of US (Table 3B), DL-US (& POX) only dominated when the DL sensitivity was $\geq$99% and specificity was 100%. Despite their higher costs, DL (& POX) are still mainly cost-effective or very cost-effective (Appendix IX.B).

In Scenario C ("Routine, High & Costly US Performance"), the relative advantage of DL-US is reduced, as it assumes greater achievements in the field of US screening efficiency levels attained under "research study" conditions. However, this higher US performance comes at a higher cost due to increased US screening time and supervision. US (& POX) dominates in many cells by providing additional QALYs at a lower cost (Table 3C). At higher specificity levels, cost-effectiveness and even very high cost-effectiveness are achievable by DL-US (& POX). Indeed, for some combinations (with a sensitivity and specificity of 99% and where a specificity level of 100% is accompanied by a sensitivity above 92%), DL-US (& POX) dominates US (& POX) because it is less expensive (Appendix IX. C) in addition to providing more QALYs (Appendix X.C.).

DISCUSSION

Our CUA focused on prospectively evaluating the anticipated substitution of DL-US for US in the future. The use of artificial intelligence-based DL-US in the diagnosis, risk stratification, and management of CHD is a promising future possibility given the current advancements in machine learning and knowledge of neural networks (56), paving the way for extremely efficient human error-free health care (82). The evaluation of DL-US images is currently severely hampered by the lack of clinical trial data on the sensitivity and specificity of DL-US for identifying sCHD and mCHD. Expected gains in sensitivity (and subsequent survival of live births) will result in increases in the number of elective abortions.

If DL-assisted US screening is found to have a lower specificity than US alone, this might result in more voluntary abortions accompanied by fewer miscarriages and stillbirths. There is, however, currently no clear-cut evidence about the lower specificity of this tool, and if it is, it is likely to be corrected in the future as part of the learning process. This concern emphasizes the need to use this tool to support sonographer clinicians, who must have a final say in the diagnostic process.

Because of these limitations, we used our model to perform a range of sensitivity analyses, including some relating to an increased cost of DL-US screening to double that of US. Our study

contributes to mapping out in advance the cost per QALY of various combinations of sensitivities and specificities, whose values are not yet known. Of course, oligopolistic suppliers of DL-US might use these data to increase DL-US costs up to the point where the intervention remains just cost-effective.

The most extensive meta-analysis of results from the "research study" perspective cannot overcome an inherent bias: that not only were the operators subject to more stringent quality controls of performance skills but also the time allocated to US performance (approximately 30 minutes) was greater than the 20 minutes devoted by busy community clinicians under "routine-reported" conditions. The potential comparative sensitivity advantage of DL-US compared to US increased (by 21.8% for sCHD and by 32.0% for mCHD) when US data were based on the three "routine" studies that were identified (68-70). For this reason, in our baseline and first two analyses (Tables 3A, 3B), we relied on data from "routine reports".

"Routine reports" show greater resemblance to real-life routine practices than studies that are operated under prospective "research study" protocols. However, higher sensitivities have been reported in routine (reported) practices from a thoroughly organized national screening program with well-defined ultrasound protocols (17). Therefore, the fact that someone cares, in routine practice, about quality control can provide an impetus toward better results. Uniform training and quality assessments of ultra sonographers within an integrated managed care consortium are additional factors for achieving greater sensitivities in both urban and rural areas (34).

The level of experience of the person performing or interpreting the scan (64,83), as well as maternal characteristics [e.g., body mass index, abdominal scars] (3,81,84), affect the detection of foetal heart malformations. However, it is possible that the use of DL-US will ameliorate these problems. If this decrease occurs, then this will at least narrow the gap between DL-US sensitivities that will be reported under clinical trial and actual field conditions.

The option of primary prevention of CHD is unlikely to be feasible since 80% of CHD cases occur in foetuses of mothers without any risk factors (85,86). However, one should be open to exploring (via CUA) the feasibility of options such as adding additional US or DL-US screening

in the second or third trimesters to mothers to be in any identified high-risk group. However, third-semester screening is unlikely to be cost effective due to the low incidence and severity of detectable defects (87,88).

The decision to recommend adding POX to the existing US protocol was made without any ex ante cost-utility analysis based on an Israeli setting. Cost-effectiveness analyses from other countries resulted in decisions to implement POX (i.e., Israel was in comparative need of this intervention), in addition to the logical assumption that the benefits of postnatal diagnoses via POX can be achieved at a very low cost. It should be noted that in some of the other countries, evaluative studies of the POX did not even factor in the cost of nursing due to the short time needed to complete the screening (41). Indeed, our retrospective (ex-post) CUA showed the original decision to be cost-effective and correct from a health economic viewpoint. The diffusion of this cheap technology appears to be far faster than was initially anticipated (89). Following the national policy decision to adopt the technology in 2021, a recent survey reported that it had been implemented by all Israeli hospitals in 2023.

However, hospitals that have implemented POX screening have been reported to be able to do so using existing nursing staff and do not incur additional staff costs. From the hospital perspective, the cost of staff time need not be included. From a societal perspective, the inclusion of staff time makes sense if the nursing time used for POX screening could have been used for other tasks. If nursing time could not be reallocated, the fact that our estimates included a costing of nursing time would cause an overestimation of the CUR for POX screening (90).

Falling outside the domain of this paper are machine learning algorithms, which include the perfusion index, heart rate, pulse delay and photoplethysmography characteristics; these algorithms have been reported to improve the sensitivity of cCHD detection by ten percentage points over pulse oximetry screening alone (91).

The calibration and structure of the model were constrained by the availability of the data. Unfortunately, for CHD patients diagnosed >24 hours after birth, no mortality, QALY or cost data have been published by age (in weeks or months) at CHD discovery. The delayed discovery of CHD associated with pulmonary hypertension and increased neurodevelopmental morbidity may lead to higher lifetime treatment costs and undesirably higher mortality rates. Early diagnosis and treatment can reduce the incidence of irreversible and intractable pulmonary hypertension through its associated morbidity, treatment costs and complications. The availability of such data would have enabled us to calculate the cost-effectiveness of adding additional screening strategies after the infant is discharged from the hospital.

The impact of disease on families of patients has often gone unrecognized and is therefore underestimated (92). Measurements of the impact are usually disease specific (92) and have been expressed only in very rare instances in utility values, such as the caregiver burden of spouses with dementia (93). Therefore, we attempted to estimate the impact of CHD on the quality of life of one (for single parents) or both parents.

An Egyptian study reported that parents of children with heart disease scored worse on QOL scales in all dimensions except bodily pain (94). Mothers have been reported to have greater stress (95) and to report feelings of anger, sadness, loneliness, helplessness, numbness, and confusion (96). In contrast to one study (52) in which QALY loss was ceased from the mother's perspective after her death, we applied these values to the child over the child's expected lifetime.

We also added the expected QALY losses of the father (if present), who is more likely to report feelings of shock, such as when first learning about the diagnosis at the postnatal stage, treatment plan or unexpected complications (96). Fathers often described their stress as not being able to protect their infant from CHD and from difficulties balancing employment (despite coworker support and being allowed flexible scheduling) with support for their partner and care of their child when hospitalized (96).

A prospective longitudinal study (97) [based on the Assessment of Quality of Life (AQoL)-8D Multi-Attribute Utility Instrument (98)] of the quality of life in parents of seriously Ill/injured

children hospitalized in cardiology, oncology or intensive care wards was performed. The study reported decreased quality of life (compared with that of parents of healthy children) of 0.0376 and 0.0048 after four weeks and seven months, respectively. The figure for four weeks was close to the 0.03 loss we used in our model based on parents of CHD children. If the WHO DWs that we used for child and adult CHD were based on parental valuations, then these are likely to have under-estimated the DWs as felt by the child or adult with CHD (99).

For both the few "routine reports" and the many "research study" reports, the data were extracted from a recent meta-analysis of first trimester screening (80). Despite a great deal of caution used in the estimation of false positives (80), there is a possibility that specificities were overestimated, leading to underestimates of the potential for improvement by adopting DL and hence upwardly biased CUA ratios.

Other factors that caused an upwards bias in our cost-utility analysis (towards higher costs per QALY) include the following:

i) We excluded parental QOL losses on account of children who were aged 18 and older. QOL losses are especially likely to still occur in the parents of young adults with sCHD.
ii) We did not attempt to estimate the impact on the quality of life of siblings (100,101) or members of the extended family (92), especially grandparents.
iii) Our perspective did not include work losses, transport costs, out-of-pocket expenses or premature burial costs resulting from the screening.
iv) A prenatal diagnosis has been found to increase the level of parental distress from diagnosis to six months after birth (102). We did not impute the QOL effects of parental worry from fetal diagnosis (or misdiagnosis) until abortion, mis-carriage or birth.
v) If the WHO DWs for CHD that we used were based on parental valuations, then these DWs are likely to have underestimated the DWs as felt by the child or adult with CHD (40).

vi) The added costs of litigation in connection with CHD were not included. These include not only the direct costs of litigation (such as lawyers and possible court costs) but also increased insurance premiums, defensiveness reactions and burn-out from misdiagnoses in the current adversarial legal system.

vii) One of the three "routine reports", was carried out in a high-quality setting, with physicians performed US taking a long time (approximately 30 minutes) with additional (transvaginal) views as required (68). This results in underestimation of the potential for improvement by adopting DL-US.

Factors causing a downwards bias in our cost–utility analysis (towards lower costs per QALYs) include the following:

i) If the WHO DWs for CHD were based on health professionals' valuations, then these DWs are likely to have over-estimated the DWs as felt by the child or adult with CHD (99).

ii) The extent to which CHD treatment costs were associated with conditions might be underestimated in our model. Lifetime CHD disease-specific costing is essential for improving these estimates.

iii) Clearly not all persons losing a pregnancy due to miscarriage, stillbirth or abortion would try to replace their loss by having another pregnancy.

A factor whose direction of bias is unknown is that we did not account for the impact of a false-positive CHD diagnosis because the effect of the initial parental stress is hard to quantify (an additional question can be asked if the mothers' stress could affect the foetus) and is offset partly or more than totally by the relief obtained once patients learn that the foetus is indeed unaffected. However, given the very high specificity of both initial heart screening (US or DL-US) and confirmation by foetal echocardiography, the number of pregnant women (and indeed their spouses) experiencing this issue would be rather small (52).

Because of lack of available data, our analysis was unable to model cost-saving and improved outcomes by DL-US related early CHD detection and prevention of irreversible and/or

intractable pulmonary hypertension. We failed to find literature data with separation of outcomes for CHD detected at birth from those diagnosed several months or years later (leading us to include detections at >24 hours as one variable). Since late diagnosis incurs high mortality and costly morbidity - including permanent neurodevelopmental defects, it is likely that incorporating this issue in CUA would have made the adoption of DL-US even more advantageous.

The adoption of DL-US can improve health systems not only in the administrative (e.g., eligibility) and operational (e.g., operating room and ER management) domains but also in the clinical domain (90). We believe that even early analyses (i.e., before all DL-US performance information is available), such as those we have undertaken, can accelerate the adoption of this new technology.

Unless there is a substantial decrease in relative specificity, the increase in clinical sensitivity provides a great impetus for the adoption of DL-US. Our exploratory CUA calculations point to the possibility of DL-US being cost-effective, despite the weakness of the data in that they were not based on screening characteristics from meta-analyses of clinical trials using DL-US.

List of Abbreviations

AQoL	Assessment of Quality of Life
cCHD	Critical Congenital Heart Disease
mCHD	Minor Congenital Heart Disease
sCHD	Serious Congenital Heart Disease
CHD	Congenital Heart Disease
CUR	Cost Utility Ratio
DL	Deep Learning
DW	Disability Weight
GDP	Gross Domestic Product
HALE	Healthy Adjusted Life Expectancy
HLHS	Hypoplastic Left Heart Syndrome
NIS	New Israel Shekles
POX	Pulse Oximetry
QALY	Quality Adjusted Life Year
QOL	Quality of Life
TGA	Transposition of Great Arteries
US	Ultrasound
USD	United States Dollars

Declarations

Ethics approval and consent to participate

'Not applicable'

Consent for publication

'Not applicable'

Availability of data and materials

Data sharing is not applicable to this article, as no datasets were generated or analysed during the current study.

Spreadsheet calculations that support the findings of this study are available upon reasonable request from the corresponding author.

Competing interests

“The authors declare that they have no competing interests”.

Funding

This paper was funded by Prof. Brezis’ (Hebrew University-Hadassah Hospital) research fund.

Authors' contributions

Study conception and design: GG, LD, UP, MB. Analytic design: GG. Published Data collection: GG, LD, MB. Economic-Epidemiological Spreadsheet analysis: GG. Manuscript draft: GG, MB. Substantial Manuscript revisions: GG, LD, UP, MB. The authors read and approved the final manuscript.

Acknowledgement (Dedication)

This paper is dedicated to the memory of Amit A. Brezis, who died at the age of three years and six months from resistant severe pulmonary hypertension. Amit had a sCHD that was missed by prenatal ultrasound. He was diagnosed 14 months after birth and underwent surgery soon after birth. Each year in Israel, approximately 70 to 80 infants suffer from severe CHD with delayed diagnosis associated with early death and/or major morbidity. Following Amit’s tragedy, a national policy of universal pulse oximetry survey in newborns was adopted in Israel in 2021,

and according to a recent survey by the Neonatologist's Union, nearly 100% compliance has been achieved. In the aftermath of the loss of Amir, lawyers rejected a suggestion to add in the settlement a sentence stating, "The provider will consider the option of using Artificial Intelligence to improve the detection of CHD in pregnancy by US". Unfortunately, in the current litigation system, money is spent on "deny and defend" rather than on effective prevention.

References

1. van der Linde D, Konings EEM, Slager MA, Witsenburg M, Helbing WA, Takkenberg JJM, et al. Birth Prevalence of Congenital Heart Disease Worldwide. A Systematic Review and Meta-Analysis. Journal of the American College of Cardiology 2011;58:2241-7.

2. do Amaral Lopes SAV, Guimarães ICB, de Oliva Costa SF, Acosta AX, Sandes KA et al. Mortality for Critical Congenital Heart Diseases and Associated Risk Factors in Newborns. A Cohort Study. Arq Bras Cardiol 2018; 111: 666-73.

3. Su Z, Zou Z, Hay SI, Liu Y, Li S, Chen H, et al. Global, regional, and national time trends in mortality for congenital heart disease, 1990−2019: An age-period-cohort analysis for the Global Burden of Disease 2019 study. EClinicalMedicine 2022;43:101249.

4. Ewer AK, Furmston AT, Middleton LJ, Deeks JJ, Daniels JP, Pattison HM, et al. Pulse oximetry as a screening test for congenital heart defects in newborn infants: a test accuracy study with evaluation of acceptability and cost-effectiveness. Health Technol Assess 2012;16(2) :v-xiii, 1-184.

5. Russo CA, Elixhauser A. Hospitalizations for Birth Defects, 2004: Statistical Brief #24. 2007 Jan. In: Healthcare Cost and Utilization Project (HCUP) Statistical Briefs [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2006.

6. Carvalho JS. Antenatal diagnosis of critical congenital heart disease. Optimal place of delivery is where appropriate care can be delivered. Arch Dis Child 2016;101:505-7.

7. Khoo NS, Van Essen P, Richardson M, Robertson T. Effectiveness of prenatal diagnosis of congenital heart defects in South Australia: a population analysis. Aust N Z J Obstet Gynaecol. 2008;48:6:559-63.

8. Stoll C, Garne E, Clementi M; EUROSCAN Study Group. Evaluation of prenatal diagnosis of associated congenital heart diseases by fetal ultrasonographic examination in Europe. Prenat Diagn 2001;21:243–52.

9. Marek J, Tomek V, Skovranek J, Povysilova V, Samanek M. Prenatal ultrasound screening of congenital heart disease in an unselected national population: a 21-year experience. Heart 2011;97:124–30.

10. McBrien A, Sands A, Craig B, Dornan J, Casey F. Major congenital heart disease: antenatal detection, patient characteristics and outcomes. J Matern Fetal Neonatal Med 2009;22:101–5.

11. Pinto NM, Keenan HT, Minich LL, Puchalski MD, Heywood M, Botto LD. Barriers to prenatal detection of congenital heart disease: a population-based study. Ultrasound Obstet Gynecol 2012;40:418–25.

12. Acharya G, Sitras V, Maltau JM, Bedrup L, Kaarsen PI, Hanssen TA et al. Major congenital heart disease in northern Norway: shortcomings of pre- and postnatal diagnosis. Acta Obstet Gynecol Scand 2004;83:1124–9.

13. Bull C. Current and potential impact of fetal diagnosis on prevalence and spectrum of serious congenital heart disease at term in the UK. British Paediatric Cardiac Association. Lancet 1999;354:1242–7.

14. Friedberg MK, Silverman NH, Moon-Grady AJ, Tong E, Nourse J, Sorenson B, et al. Prenatal detection of congenital heart disease. J Pediatr 2009;155:26–3.

15. Garne E, Stoll C, Clementi M, Euroscan G. Evaluation of prenatal diagnosis of congenital heart diseases by ultrasound: experience from 20 European registries. Ultrasound Obstet Gynecol 2001;17:386–91.

16. Jaeggi ET, Sholler GF, Jones OD, Cooper SG. Comparative analysis of pattern, management and outcome of pre- versus postnatally diagnosed major congenital heart disease: a population-based study. Ultrasound Obstet Gynecol 2001;17:380–5.

17. Parikh LI, Grantz KL, Iqbal SN, Huang CC, Landy HJ, Fries MH et al. Neonatal outcomes in fetuses with cardiac anomalies and the impact of delivery route. Am J Obstet Gynecol 2017;217: ;217(4):469.e1-469.e12. doi: 10.1016/j.ajog.2017.05.049.

18. Ailes EC, Gilboa SM, Riehle-Colarusso T, et al. Prenatal diagnosis of nonsyndromic congenital heart defects. Prenat Diagn 2014;34:214–22.

19. Quartermain MD, Pasquali SK, Hill KD, et al. Variation in prenatal diagnosis of congenital heart disease in infants. Pediatrics 2015; 136(2):e378-85.

20. Khoshnood B, Lelong N, Houyel L, Bonnet D, Ballon M, Jouannic JM, et al. Impact of prenatal diagnosis on survival of newborns with four congenital heart defects: a prospective, population-based cohort study in France (the EPICARD Study). BMJ Open 2017; 8:e018285. doi: 10.1136/bmjopen-2017-018285.

21. van Velzen CL, Clur SA, Rijlaarsdam ME, Bax CJ, Pajkrt E, Heymans MW, et al. Prenatal detection of congenital heart disease—results of a national screening programme. BJOG 2016;123:400–7.

22. Tegnander E, Eik-Nes SH. The examiner's ultrasound experience has a significant impact on the detection rate of congenital heart defects at the second-trimester fetal examination. Ultrasound Obstet Gynecol 2006;28: 8–14.

23. Tegnander E, Williams W, Johansen OJ, Blaas HG, Eik-Nes SH. Prenatal detection of heart defects in a non-selected population of 30,149 fetuses—detection rates and outcome. Ultrasound Obstet Gynecol 2006;27:252–65.

24. Lytzen R, Vejlstrup N, Bjerre J, Petersen OB, Leenskjold S, Dodd JK, et al. Live-born major congenital heart disease in Denmark: incidence, detection rate, and termination of pregnancy rate from 1996 to 2013. JAMA Cardiol. 2018;3::829-37.

25. Levey A, Glickstein JS, Kleinman CS, Levasseur SM, Chen J, Gersony WM, et al. The impact of prenatal diagnosis of complex congenital heart disease on neonatal outcomes. Pediatr Cardiol. 2010; 31:587-97.

26. Cloete E, Bloomfield FH, Sadler L, de Laat MWM, Finucane AK, Gentles TL. Antenatal detection of treatable critical congenital heart disease is associated with lower morbidity and mortality. J Pediatr 2019;204:66–70.

27. Calderon J, Angeard N, Moutier S, Plumet MH, Jambaque I, Bonnet D. Impact of prenatal diagnosis on neurocognitive outcomes in children with transposition of the great arteries. J Pediatr 2012;161:94-8..

28. Khoshnood B, Lelong N, Houyel L, Thieulin AC, Jouannic JM, Magnier S, et al. Prevalence, timing of diagnosis and mortality of newborns with congenital heart defects: a population-based study. Heart 2012;98:1667–73.

29. Brown MD, Wernovsky G, Mussatto KA, Berger S. Long-term and developmental outcomes of children with complex congenital heart disease. Clin Perinatol 2005;32:1043–57.

30. Tworetzky W, McElhinney DB, Reddy VM, Brook MM, Hanley FL, Silverman NH. Improved surgical outcome after fetal diagnosis of hypoplastic left heart syndrome. Circulation 2001;103:1269–73.

31. Patient Safety and Quality Committee, Society for Maternal-Fetall Medicine. Combs CA, Hameed AB, Friedman AM, Hoskins IA. Proposed quality metrics to assess accuracy of prenatal detection of congenital heart defects. Am J Obstet Gynecol 2020;222:6:B2-B9.

32. Morris SA and Lopez KN. Deep learning for detecting congenital heart disease in the fetus. Nature Medicine 2021;27: 759–65

33. Berndl A, Pearce K, Mondal T. Fetal echocardiography assists in determining optimal delivery site. J Obstet Gynaecol Can 2014;36:210–5.

34. Levy DJ, Pretorius DH, Rothman A, et al. Improved prenatal detection of congenital heart disease in an integrated health care system. Pediatr Cardiol 2013;34:670–9.

35. Peterson C, Ailes E, Riehle-Colarusso T, Oster ME, Olney RS, Cassell CH, et al. Late detection of critical congenital heart disease among US infants: estimation of the potential impact of proposed universal screening using pulse oximetry. JAMA Pediatr. 2014;168:361–70.

36. Plana_MN, Zamora_J, Suresh G, Fernandez-Pineda_L, Thangaratinam_S, Ewer_AK. Pulse oximetry screening for critical congenital heart defects. Cochrane Database of Systematic Reviews 2018, Issue 3. Art. No.: CD011912. DOI: 10.1002/14651858.CD011912.pub2.

37. Mahle WT, Newburger JW, Matherne GP, Smith FC, Hoke TR, Koppel R, et al. Role of pulse oximetry in examining newborns for congenital heart disease: a scientific statement from the American heart association and American Academy of Pediatrics. Circulation. 2009;120:447–58.

38. Awerbach JD, Mallory Jr GB, Kim S, Cabrera AG. Hospital Readmissions in Children with Pulmonary Hypertension: A Multi-Institutional Analysis. J Pediatr 2018;195:95-101.

39. Faraoni D, Nasr VG, DiNardo JA. Overall Hospital Cost Estimates in Children with Congenital Heart Disease: Analysis of the 2012 Kid's Inpatient Database. Pediatr Cardiol 2016; 37:37–43.

40. Eckersley L, Saddler L, Parry E, Finucane K, Gentels TL. Timing of diagnosis affects mortality in critical congenital heart disease. Arch Dis Child 2016;101:516–20.

41. Mukerji A, Shafey A, Jain A, Cohen E, Shah PS, Sander B, et al. Pulse oximetry screening for critical congenital heart defects in Ontario, Canada: a cost-effectiveness analysis. Canadian Journal of Public Health 2020; 111:804–11 (Suppl 2).

42. Griebsch I, Knowles R L, Brown J, Bull C, Wren C, Dezateux CA. Comparing the clinical and economic effects of clinical examination, pulse oximetry, and echocardiography in newborn screening for congenital heart defects: a probabilistic cost effectiveness model and value of information analysis. International Journal of Technology Assessment in Health Care 2007; 23: 192–204.

43. Knowles R, Griebsch I, Dezateux C, Brown J, Bull C, Wren C. Newborn screening for congenital heart defects: a systematic review and cost-effectiveness analysis. Health Technology Assessment (Winchester) 2005; 9:1–152 iii-iv.

44. de-Wahl Granelli A, Wennergren M, Sandberg K, Mellander M, Bejlum C, Inganäs L, et al. Impact of pulse oximetry screening on the detection of duct dependent congenital heart disease: a Swedish prospective screening study in 39,821 newborns. BMJ (Clinical Research ed) 2009; doi: 10.1136/bmj.a3037.

45. Roberts TE, Barton PM, Auguste PE, Middleton LJ, Furmston AT, Ewer AK. Pulse oximetry as a screening test for congenital heart defects in newborn infants: a cost-effectiveness analysis. Archives of Disease in Childhood 2012; 97: 221–6.

46. Peterson C, Grosse SD, Oster ME, Olney RS, Cassell C H. Cost-effectiveness of routine screening for critical congenital heart disease in US newborns. Pediatrics 2013; 132: e595–e603.

47. Narayen IC, te Pas AB, Blom NA, van den Akker-van Marie E. Cost-effectiveness analysis of pulse oximetry screening for critical congenital heart defects following homebirth and early discharge. European Journal of Pediatrics 2019; 178: 97–103.

48. Pinto NM, Nelson R, Pulchalski M, Metz TD, Smith KJ. Cost-effectiveness of prenatal screening strategies for congenital heart disease Ultrasound. Obstet Gynecol 2014; 44: 50–7

49. Trujillo DL, Reyes NFS, Restrepo AT, Velasquez CLC , Torres MTD, Ducuara SVR et al. Cost-effectiveness analysis of newborn pulse oximetry screening to detect critical congenital heart disease in Colombia. Cost Eff Resour Alloc. 2019;17:11.

50. Evers PD, Ranade D, Lewin M, Arya B. Diagnostic Approach in Fetal Coarctation of the Aorta: A Cost-Utility Analysis. Journal of the American Society of Echocardiography 2017; 30:589-94.

51. Bak GS, Shaffer BL, Madriago E, Allen A, Kelly B, Caughey AB, et al. Detection of fetal cardiac anomalies: cost-effectiveness of increased number of cardiac views. Ultrasound Obstet Gynecol 2020; 55: 758–67

52. Carvalho JS, Allan LD, Chaoui R, Copel JA, DeVore GR, Hecher K, et al. ISUOG Practice Guidelines (updated): sonographic screening examination of the fetal heart. International Society of Ultrasound in Obstetrics and Gynecology, Ultrasound Obstet Gynecol 2013; 41: 348–59

53.Drukker L, Sharma H, Karim JN, Droste R, Noble JA, Papageorghiou AT. Clinical workflow of sonographers performing fetal anomaly ultrasound scans: deep-learning-based analysis. Ultrasound Obstet Gynecol. 2022;60:759-65

54.Drukker L, Noble JA, Papageorghiou AT. Introduction to artificial intelligence in ultrasound imaging in obstetrics and gynecology. Ultrasound Obstet Gynecol. 2020;56:498-505

55. Ghosh J, Papadopoulou A, Devall AJ, Jeffery HC, Beeson LE, Do V et al. Methods for managing miscarriage: a network meta-analysis. Cochrane Database Syst Rev. 2021 Jun 1;6(6):CD012602. doi: 10.1002/14651858.CD012602.pub2.

56. Central Bureau of Statistics. Statistical Abstract of Israel 2022 no. 73, Jerusalem. Statistical Abstract of Israel 2022 - No.73. https://www.cbs.gov.il/en/publications/Pages/2022/Statistical-Abstract-of-Israel-2022-No-73.aspx (Accessed 22nd January 2024).

57. Central Bureau of Statistics. Statistical Abstract of Israel 2016 no. 67, Jerusalem. Statistical Abstract of Israel 2016 - No.67. https://www.cbs.gov.il/en/publications/Pages/2016/Statistical-Abstract-of-Israel-2016-No-67.aspx (Accessed 22nd January 2024).

58. Abouk R, Grosse SD, Ailes EC, Oster ME. Association of US State Implementation of Newborn Screening Policies for Critical Congenital Heart Disease with Early Infant Cardiac Deaths JAMA. 2017;318:2111-8.

59. Best KE, Rankin J . Long-Term Survival of Individuals Born with Congenital Heart Disease: A Systematic Review and Meta-Analysis. J Am Heart Assoc. 2016;5: e002846

60. Ginsberg G, Kark J, Einav S. Is out-of-hospital resuscitation worthwhile? Cost-Utility Analysis of Cardiac Resuscitation Services in Jerusalem. Resuscitation 2015; 86:54-61.

61. Ministry of Health Price List for Ambulatory and Hospitalization Services. Effective date: 01/10/2022. Ministry of Health, Jerusalem. https://www.gov.il/BlobFolder/dynamiccollectorresultitem/spl-moh-price-list-030722/he/files_databases_moh-price-list_2022_moh-Price-List-010722.xlsx (Accessed 22nd January 2024)

62. Reeder MR, Kim J, Nance A, Krikov S , Feldkamp ML, Randall H, et al. Evaluating Cost and Resource Use Associated with Pulse Oximetry Screening for Critical Congenital Heart Disease: Empiric Estimates and Sources of Variation Birth Defects Research (Part A) 2015:103:962–71.

63. Quenby S, Gallos ID, Dhillon-Smith RK, Podesek M, Stephenson MD, Fisher J, et al. Miscarriage matters: the epidemiological, physical, psychological, and economic costs of early pregnancy loss. Lancet 2021; 397: 1658–67.

64. Peterson C, Dawson A, Grosse SD, Riehle-Colarusso T, Olney RS, Tanner JP, et al. Hospitalizations, costs, and mortality among infants with critical congenital heart disease: how important is timely detection? Birth Defects Research. Part A, Clinical and molecular teratology. 2013;97:664-72.

65. Israel Ministry of Health. Congenital Malformations in Israel (2000-2014), Quality checks and trends over time. Ministry of Health, Jerusalem, 2018.

66. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, et al. American Heart Association Statistics Committee; Stroke Statistics Subcommittee. AHA statistical Update. Heart Disease and Stroke Statistics—2016 Update. A Report from the American Heart Association. Circulation 2016; 133: 447-54.

67. Brown KL, Ridout DA, Hoskote A, Verhulst L, Ricci M, Bull C. Delayed diagnosis of congenital heart disease worsens preoperative condition and outcome of surgery in neonates. Heart 2006; 92: 1298–1302.

68. Syngelaki A, Hammami A, Bower D, Zidere V, Akolekar R, Nicolaides KH. Diagnosis of fetal non chromosomal abnormalities on routine ultrasound examination at 11–13 weeks' gestation. Ultrasound Obstet Gynecol 2019; 54: 468–76

69. Andrew C, Gopal S, Ramachandran H, Suvika M. First Trimester Ultrasound: Addition of Anatomical Screening Adds Value to the Examination: A Retrospective Case Series Journal of Evolution of Medical and Dental Sciences 2015; 4: 2690-9.

70. Vayna AM, Veduta A, Duta S, Panaitescu AM, Stoica S, BuinoiuN, et al. Diagnosis of Fetal Structural Anomalies at 11 to 14 Weeks. J Ultrasound Med 2018; 37:2063–73.

71. Benderly M, Buber J, Kalter-Leibovici O, Blieden L, Dadashev A, Lorber A, et al. Health Service Utilization Patterns Among Adults With Congenital Heart Disease. : A Population-Based Study. J Am Heart Assoc. 2021;10:e018037. DOI: 10.1161/JAHA.120.018037

72. Global Health Data Exchange GHDx. Global Burden of Disease Study 2019 (GBD 2019) Disability Weights. https://ghdx.healthdata.org/record/ihme-data/gbd-2019-disability-weights. Accessed (accessed 2nd May 2022)

73. Gravensteen IK, Helgadottir LB, Jacobsen E-M, Sandset PM, Ekeberg Ø. Long-term impact of intrauterine fetal death on quality of life and depression: a case–control study. BMC Pregnancy Childbirth 2012; 12: 43

74. Turton P, Evans C, Hughes P. Long-term psychosocial sequelae of stillbirth: phase II of a nested case control study. Arch Womens Ment Health 2009; 12: 35–41.

75. Kuppermann M, Nease RF, Learman LA, Gates E, Blumberg B, Washington AE. Procedure-related miscarriages and Down syndrome-affected births: implications for prenatal testing based on women's preferences. Obstet Gynecol 2000; 96: 511–6.

76. Central Bureau of Statistics. Israel Statistics Monthly. August 2023, CBS, Jerusalem. https://www.cbs.gov.il/he/publications/Pages/2023. Accessed 14th September 2023.

77. WHO Commission on MacroEconomics and Health: Macroeconomics and health: investing in health for economic development. Report of the commission on Macroeconomics and Health Geneva: World Health Organization 2001.

78. Central Bureau of Statistics. Statistical Abstract of Israel 2021 no. 72, Jerusalem. Statistical Abstract of Israel 2021 - No.72 (cbs.gov.il). https://www.cbs.gov.il/en/publications/Pages/2021/Statistical-Abstract-of-Israel-2021-No-72.aspx (Accessed 16th April 2023).

79. Central Bureau of Statistics. Applications for Pregnancy Termination in 2019, 29th December 2020, Jerusalem. 05_20_434b.pdf (cbs.gov.il) https://www.cbs.gov.il/en/mediarelease/Pages/2020/Applications-for-Pregnancy-Termination-in-2019.aspx (Accessed 22nd January 2024). (In Hebrew)

80. Karim JN, Bradburn E, Roberts N, Papageorghiou AT for the ACCEPTS study. First-trimester ultrasound detection of fetal heart anomalies: systematic review and meta-analysis. Ultrasound Obstet Gynecol 2022; 59: 11–25.

81. Areia M, Mori Y, Correale L, Repici A, Bretthauer M, Sharma P, et al. Cost-effectiveness of artificial intelligence for screening colonoscopy: a modelling study. Lancet Digit Health 2022; 4: e436–44

82. Sethi Y, Patel N, Kaka N, Desai A, Kaiwan O, Sheth M, et al. Artificial Intelligence in Pediatric Cardiology: A Scoping Review. J. Clin. Med 2022; 11: 7072

83. Guo F, Tang S, Guo T, Bartell S, Detrano R. Revised threshold values for neonatal oxygen saturation at mild and moderate altitudes. Acta Paediatr 2020; 109:2: 321-6

84. Martin GR, Ewer AK, Gaviglio A, Hom LA, Saarinen A, Sontag M, et al. Updated Strategies for Pulse Oximetry Screening for Critical Congenital Heart Disease. Pediatrics 2020;146:1:e20191650

85. Li Y-F, Zhou K-Y, Fang J, Wang C, Hua Y-M, Mu D-Z. Efficacy of prenatal diagnosis of major congenital heart disease on perinatal management and perioperative mortality: a meta-analysis. World J Pediatr 2016; 12: 298-307

86. Agarwal A, Vittinghoff E, Myers JJ, Dudley RA, Khan A, John A, et al. Ambulatory Health Care Service Use and Costs Among Commercially Insured US Adults With Congenital Heart Disease. Research Letter . JAMA Network Open 2020;3:9:e2018752.

87. Odibo AO, Coassolo KM, Stamilio DM, Ural SH, Macones GA. Should all pregnant diabetic women undergo a fetal echocardiography? A cost-effectiveness analysis comparing four screening strategies. Prenat Diagn 2006; 26: 39–44.

88. Drukker L, Bradburn E, Rodriguez GB, Roberts NW, Impey L, Papageorghiou AT. How often do we identify fetal abnormalities during routine third-trimester ultrasound? A systematic review and meta-analysis. BJOG 2021;128(2):259-69.

89. Bin-Nun A, Hammerman C, Mimouni FB, Wasserteil N, Kasirer YM.The Saga of Pulse Oximetry Screening for Critical Congenital Heart Disease in Israel: A Historical Perspective. IMAJ 2021; 23: 229-32.

90. Grosse SD, Peterson C, Abouk R, Glidewell J, Oster ME. Cost and Cost-Effectiveness Assessments of Newborn Screening for Critical Congenital Heart Disease Using Pulse Oximetry: A Review. Int J Neonatal Screen. 2017; 3:34

91. Lai Z, Vadlaputi P, Tancredi DJ, Garg M, Koppel RI, Goodman M, et al. Enhanced Critical Congenital Cardiac Disease Screening by Combining Interpretable Machine Learning Algorithms. Annu Int Conf IEEE Eng Med Biol Soc 2021; 1403–6.

92. Golics CJ, Basra MKA , Finlay AY, Salek S. The impact of disease on family members: a critical aspect of medical care. Journal of the Royal Society of Medicine 2013; 106: 399–407.

93. Tu JY, Jin G, Chen J-H, Chen Y-CJ. Caregiver Burden and Dementia: A Systematic Review of Self-Report Instruments. Alzheimers Dis . 2022;86:1527-43.

94. Arafa MA, Zaher SR, El-Dowaty AA, Moneeb DE. Quality of life among parents of children with heart disease. Health and Quality of Life Outcomes 2008; 6:91.

95. Jackson AC, Frydenberg E, Liang R P–T, Higgins RO, Murphy BM. Familial Impact and Coping with Child Heart Disease: A Systematic Review. Pediatr Cardiol: 2015;36:695–712

96. Sood E, Karpyn A, Demianczyk AC et al, Mothers and fathers experience stress of congenital heart disease differently: Recommendations for pediatric critical care. Pediatr Crit Care Med. 2018; 19: 7: 626–634.

97. Graj E, Muscara F, Anderson V, Hearps S, McCarthy M. Quality of life in parents of seriously Ill/injured children: a prospective longitudinal study. Quality of Life Research 2021: 30:193–202.

98. Richardson, JR, Iezzi, A, Khan, MA, Maxwell, A. Validity and reliability of the Assessment of Quality of Life (AQoL)-8D multiattribute utility instrument. Patient 2014;7:85-96.

99. Saigal S, Stoskopf BL, Feeny D, Furlong W, Burrows E, Rosenbaum PL, et al. Differences in preferences for neonatal outcomes among health care professionals, parents and adolescents. JAMA 1999; 281:1991-7.

100. Parker R , Houghton S, Bichard E, McKeever S. Impact of congenital heart disease on siblings: A review Journal of Child Health Care 2020; 24: 297–316

101. O'Brien I, Duffy A and Nicholl H. Impact of childhood chronic illnesses on siblings: a literature review. Br J Nurse 2009; 18: 1358–65.

102. Brosig CL, Whitstone BN, Frommelt MA, Frisbee SJ, Leuthner SR. Psychological distress in parents of children with severe congenital heart disease: the impact of prenatal versus postnatal diagnosis. J Perinatol 2007; 27: 687–92.

Table 1: Inputs, Effects and Costs by Interventions type.

	Null	US Routine (a)	POX	US Routine & POX	Deep Learning - US	Deep Learning -US &POX
Echocardiogram (b)	0	672	3,394	3,742	8,013	10,760
Ultrasounds	0	190,317	0	190,317	190,317	190,317
Abortions	2,907	3,084	2,907	3,084	5,689	5,689
Miscarriages	5,637	5,621	5,637	5,621	5,610	5,610
Stillbirths	567.2	566.7	567.2	566.7	566.4	566.4
POX tests	0	0	181,206	181,051	0	180,952
COSTS: (million USD)						
Intervention Costs	56	81	59	84	104	107
Treatment Costs	461.5	422.9	461.9	423.1	402.2	402.3

(a) Based on retrospective "routine reports"
(b) on true positive and false positive infants
Screening costs: US ($124.42), DL ($144.82)
DL Sensitivity (95%), Specificity (96%)

Table 2: QALYs gained, Intervention and Treatment Costs by intervention.

	Null	US Routine (a)	POX	US Routine & POX	Deep Learning -US	Deep Learning -US & POX
Costs (million nis)						
Intervention Cost	56	81	59	84	104	107
Treatment Cost	461.5	422.9	461.9	423.1	402.2	402.3
Total Costs	**517**	**504**	**521**	**512**	**506**	**509**
Net Costs cf: null		**-13**	**4**	**-5**	**-11**	**-8**
QALY losses						
Mothers: Abortions	431.9	458.2	431.9	458.2	845.3	845.3
Mothers: Miscarriages	4,088	4,076	4,088	4,076	4,069	4,069
Mothers: Stillbirths	411	410.9	411.3	410.9	410.7	410.7
Mothers: Neonatal Mortality	91	47	43	27	22	19
Parents: due to Childs CHD	50	47	50	47	44	44
Patient: CHD lifetime	21,573	19,748	21,549	19,738	18,165	18,164
TOTAL QALY loss	**26,646**	**24,787**	**26,573**	**24,756**	**23,556**	**23,552**
QALYs gained cf null		**1,858**	**73**	**1,889**	**3,089**	**3,093**
Average Cost-Effectiveness Ratio (USD per QALY gained)		**dom**	**50,882**	**dom**	**dom**	**dom**

Notes:
(a) based on retrospective "routine reports"
dom denotes intervention dominates the null by providing more QALYS at no additional cost
Screening costs: US ($124.42), DL ($144.82)
DL Sensitivity (95%), Specificity (96%)

Table 3: Costs per QALY (USD at 2022 price levels) of DL-US (& POX) vs US (& POX)
(based on US operational data: Sensitivity 58.1%, Specificity 100%)
"Routine DL-US" : Screening Costs: US (418 nis) DL (487nis)

A.	DL Specificity						
DL Sensitivity	**90%**	**92%**	**94%**	**96%**	**98%**	**99%**	**100%**
80%	5,205,000	151,000	52,000	18,752	1,651	dom	dom
84%	190,000	74,000	32,000	11,336	dom	dom	dom
88%	98,000	46,000	21,000	6,780	dom	dom	dom
92%	61,000	32,000	14,875	3,670	dom	dom	dom
96%	43,000	23,000	10,420	1,413	dom	dom	dom
99%	34,000	19,000	7,919	88	dom	dom	dom

(based on US operational data: Sensitivity 58.1%, Specificity 100%)
"Routine High DL-US Screening Costs" : US ($122.42) DL ($144.82)

B.	DL Specificity						
DL Sensitivity	**90%**	**92%**	**94%**	**96%**	**98%**	**99%**	**100%**
80%	7,990,000	254,000	105,000	28,000	28,000	19,347	12,554
84%	314,000	129,000	68,000	11,310	21,000	13,003	8,621
88%	154,000	83,000	49,000	8,632	15,180	10,129	5,752
92%	99,000	60,000	37,000	6,548	11,456	7,238	3,561
96%	72,000	46,000	29,000	5,138	8,562	4,993	1,832
99%	58,000	38,000	25,000	4,312	6,808	3,613	758

(based on US research data: Sensitivity 79.9%, Specificity 100%)
"Routine High DL-US Costs & Performance":
Screening Costs: US (418 nis) DL (1,009 nis)

C.	DL Specificity						
DL Sensitivity	**90%**	**92%**	**94%**	**96%**	**98%**	**99%**	**100%**
80%	US dom	US dom	US dom	US dom	US dom	US dom	671,000
84%	US dom	US dom	US dom	US dom	US dom	87,000	15,525
88%	US dom	US dom	US dom	US dom	60,000	69,000	1,658
92%	US dom	US dom	US dom	114,000	23,000	23,914	DL dom
96%	US dom	US dom	179,000	46,000	10,890	6,096	DL dom
99%	US dom	412,000	85,000	29,000	20,481	DL dom	DL dom

Notes:	
US dom: US (& POX) dominate DL-US (& POX) by supplying more QALYS at a lower cost	
DL dom: DL-US (& POX) dominate US (& POX) by supplying more QALYS at a lower cost	
0< Cost per QALY < $54,800 denotes very cost-effective intervention	
$54,800 <Cost per QALY< $164,400 denotes cost-effective intervention	

$164,400 <Cost per QALY denotes intervention is not cost-effective

Appendix Ia: Major Diagnostic Modes

Ultrasound (US)

Over the years, improvement of obstetric ultrasound (US) is attributed to many factors, including developments in education, accreditation, guidelines, quality assurance, anatomical and physiological knowledge, and imaging quality (1).

There is however, a large difference between sensitivities attained under optimal (usually "research study" conditions) and routine everyday practices. Optimal US results were by and large generated by referral institutions affiliated with academic centers often under prospective conditions. Far lower US sensitivities were reported from a few studies (2-4) from routine, overloaded community-based practices (5) that allocated far less time to the screening than under optimal conditions. Thus, the potential for improvement and the incentive to improve post-natal screenings (see pulse oximetry and deep learning below) is, in reality, far greater than what has been reported in the literature dominated by research studies with high quality control conditions.

In Israel, until 2023, routine screening for both sCHDs and mCHDs relied on a mid-trimester anomaly scan in pregnant women, involving basic sectional imaging of heart anatomy and function as well as a postnatal auscultation by stethoscope (6). US detection rates show a strong correlation with CHD severity (7). However, like some other countries, there is no required accreditation for US in Israel despite the evidence that the lack of accreditation is highly correlated with poorer performance quality (8,9). This results in US sensitivity in Israel falling still further below the sensitivities reported in optimal prospective trials and in retrospective studies (especially for mCHD), resulting in many newborns being discharged from hospital before CHD was diagnosed (6).

Even some members, though by no means all, of specific groups, (ultra-orthodox Jews and Muslims) who are strictly opposed to aborting even severely malformed foetuses (10), opt to undergo an US scan for anomalies, which can at least give them time to prepare for a possibly challenged infant.

Postnatal echocardiography is the established gold standard for diagnosing CHDs. However, echocardiography may also contribute to an apparent rising incidence of CHDs mainly as a result of the detection of abnormalities which are of no functional or clinical significance (11,12). As a result, echocardiography is likely to have significant limitations as a screening tool, not only due to elevated false positive rates (13,14), but also as a result of cost and qualified manpower constraints (6).

Appendix Ib: Major Diagnostic Modes

Pulse Oximetry (POX)

Infants not detected with sCHD before discharge leaving hospital had far lower hospitalization days and hospital costs during infancy (up to 12 months old) than persons with timely detection between birth and hospital discharge. However, after adjusting for mothers age, race, education, payor status and the infants gender gestation period and type of sCHD, a 64% reduction in costs was changed to a 35% increase in relative costs for the undetected group, as a result of the groups 52% higher hospital admission rate and 18% higher utilization of hospitalisation days (15).

In the USA, statewide implementation of mandatory policies for newborn screening for critical CHD was associated with a significant decrease in infant cardiac deaths between 2007 and 2013 compared with states without these policies (16).

A policy of universal pulse oximetry survey in newborns was rejected in 2013 by the National Council for Child Health and Pediatrics in Israel because of the argument that, in this country, an ultrasound is nearly always performed – and with the mistaken assumption that its sensitivity is high. Following the tragic case of the child – to whom this paper is dedicated – a national policy was adopted in 2021 for pulse oximetry (POX) in newborns and a committee of experts was appointed to advise about quality control in pregnancy US with special consideration for Artificial Intelligence (AI) implementation in this field.

POX, measures the percentage of haemoglobin saturated with oxygen (17) in a simple, quick, safe, painless (6), non-invasive way (18) that is acceptable to both parents and staff (6). In addition, POX may also prove to be beneficial in conditions that cannot be identified before birth, including coarctation of the aorta that occurs with duct closure – 48 hours after birth.

Besides not being distressing for the baby, and reassuring for parents. POX has the potential to detect problems soon after birth, before discharge, allowing treatment to be started and lives to be saved (22). POX levels usually below 95% [subject to adjustment by altitude (10,19)] indicate possible sCHD in the new-born. Those not achieving predetermined oxygen saturation thresholds are usually referred for echocardiography (6).

In a national survey, around 78% of the neonatal units in the UK that used POX, responded that they felt that screening did not lead to an increase in the number of unnecessary investigations, while 10% of the neonatal units felt that any small increase was justified and offset by the benefits of identifying considerable cardiac and non-cardiac pathology (20).

However, even the combined sensitivity of US (and POX) is insufficient (21,22) and a significant proportion of both sCHD and mCHD are diagnosed post-discharge (23).

Appendix Ic: Major Diagnostic Modes

Deep Learning & Ultrasound (DL-US)

It is hoped that in the future, the evolution of deep learning (DL) based on artificial intelligence (AI) and data science, will be integrated into mechanizing several aspects of medical care requiring critical thinking: including diagnosis, risk stratification, and management, thus reducing both physician's burden and the likelihood of human error. The use of neural networks and machine learning may significantly improve the diagnostic value (24) of cardiac magnetic resonance imaging, echocardiograms, computer tomography scans and electrocardiographs, in consequence augmenting and improving the diagnostic accuracy of detecting foetal CHD (25-28).

AI models have been found to be statistically superior to standard foetal biometry–based gestational age estimates derived from images captured by expert sonographers in estimating gestational age (29). AI applied to ultrasound examination of the foetal heart has been explored to improve diagnostic accuracy in the context of foetal CHD (24,26,30), matching "expert" performance levels (26). However, there is a paucity of information about the impact that AI might have on national screening for CHD. Another study reported that electrocardiograms using DL-US outperformed diagnostically (i.e. greater sensitivity and specificity), cardiologists reading of electrocardiograms for atrial septal defects (31).

DL-US has the potential to level the playing field for centers that obtain access to this technology. Particularly for those hospitals that are exposed to a low volume of CHD and/or have sonographers with limited exposure to and training on fetal CHD. Also, for those in populations with a lower socioeconomic status, whose maternal rates of pre-existing diabetes additionally puts their patients at higher risk for having a child with CHD (32).

References to Appendices 1a,1b & 1c.

1. Drukker L, Sharma H, Karim N, Droste R, Noble JA, Papageorghiou AT. Clinical workflow of sonographers performing fetal anomaly ultrasound scans: deep-learning-based analysis, Ultrasound Obstet Gynecol 2022;60:759-65.

2. Syngelaki A, Hammami A, Bower D, Zidere V, Akolekar R, Nicolaides KH. Diagnosis of fetal non-chromosomal abnormalities on routine ultrasound examination at 11–13 weeks' gestation. Ultrasound Obstet Gynecol 2019; 54: 468–76.

3. Andrew C, Gopal S, Ramachandran H, Suvika M. First Trimester Ultrasound: Addition of Anatomical Screening Adds Value to the Examination: A Retrospective Case Series Journal of Evolution of Medical and Dental Sciences 2015; 4: 2690-2699.

4. Vayna AM, Veduta A, Duta S, Panaitescu AM, Stoica S, BuinoiuN, et al. Diagnosis of Fetal Structural Anomalies at 11 to 14 Weeks. J Ultrasound Med 2018; 37:2063–73.

5. Krishnan A, Donofrio MT. Impact of Socioeconomic Status, Race and Ethnicity, and Geography on Prenatal Detection of Hypoplastic Left Heart Syndrome and Transposition of the Great Arteries. Circulation 2021;143:2049–60.

6. Ewer AK, Furmston AT, Middleton LJ, Deeks JJ, Daniels JP, Pattison HM et al.
Pulse oximetry as a screening test for congenital heart defects in newborn infants:
a test accuracy study with evaluation of acceptability and cost-effectiveness. Health Technol Assess 2012;16: 2::v-xiii, 1-184.

7. van Velzen CL, Ket JCF, van de Ven PM, Blom NA, Haak MC. Systematic review and meta-analysis of the performance of second-trimester screening for prenatal detection of congenital heart defects. Int J Gynecol Obstet 2018; 140: 137–45.

8. Sorrentino K. Accreditation, Credentialing, and Quality Improvement in Diagnostic Medical Sonography: A Literature Review. Journal of Diagnostic Medical Sonography 2019; 35: 401–11.

9. Levy DJ, Pretorius DH, Rothman A, Gonzales M, Rao C, Nunes ME et al. Improved Prenatal Detection of Congenital Heart Disease in an Integrated Health Care System. Pediatr Cardiol 2013;34:670–9.

10. Bin-Nun A, Hammerman C, Mimouni FB, Wasserteil N, Kasirer YM. The Saga of Pulse Oximetry Screening for Critical Congenital Heart Disease in Israel: A Historical Perspective. IMAJ 2021; 23: 229-32.

11. Hoffman JIE, Kaplan S. The incidence of congenital heart disease. J Am Coll Cardiol 2002;39:1890–900.

12. Wren C, Richmond S, Donaldson L. Temporal variability in birth prevalence of cardiovascular malformations. Heart 2000;83:414–19.

13. Knowles R, Griebsch I, Dezateux C, Brown J, Bull C, Wren C. Newborn screening for congenital heart defects: a systematic review and cost-effectiveness analysis. Health Technol Assess 2005; 9(44):1-152, iii-iv.

14. Mahle WT, Newburger JW, Matherne GP, Smith FC, Hoke TR, Koppel R, et al. Role of pulse oximetry in examining newborns for congenital heart disease: a scientific statement from the AHA and AAP. Pediatrics 2009;124:823–36

15. Peterson C, Dawson A, Grosse SD, Riehle-Colarusso T, Olney RS, Tanner JP, et al. Hospitalizations, costs, and mortality among infants with critical congenital heart disease: how important is timely detection? Birth Defects Research. Part A, Clinical and molecular teratology. 2013;97:664-72.

16. Abouk R, Grosse SD, Ailes EC, Oster ME. Association of US State Implementation of Newborn Screening Policies for Critical Congenital Heart Disease With Early Infant Cardiac Deaths JAMA. 2017;318:2111-8.

17. Plana_MN, Zamora_J, Suresh G, Fernandez-Pineda_L, Thangaratinam_S, Ewer_AK. Pulse oximetry screening for critical congenital heart defects. Cochrane Database of Systematic Reviews 2018, Issue 3. Art. No.: CD011912. DOI: 10.1002/14651858.CD011912.pub2.

18. Roberts TE, Barton PM, Auguste PE, Middleton LJ, Furmston AT, Ewer AK. Pulse oximetry as a screening test for congenital heart defects in newborn infants: a cost-effectiveness analysis, Arch Dis Child 2012;97:221–6.

19. Guo F, Tang S, Guo T, Bartell S, Detrano R. Revised threshold values for neonatal oxygen saturation at mild and moderate altitudes. Acta Paediatr 2020; 109:2: 321-6

20. Brown S, Liyanage S, Mikrou P, Singh A, Ewer AK. Newborn pulse oximetry screening in the UK: a 2020 survey. The Lancet September 26, 2020; 396;881.

21. Campbell MJ, Quarshie WO, Faerber J, Goldberg DJ, Mascio CE, Blinder J. Pulse Oximetry Screening Has Not Changed Timing of Diagnosis or Mortality of Critical Congenital Heart Disease Pediatr Cardiol. 2020 Jun;41:899-904..

22. Chakraborty A, Gorla SR, Swaminathan S. Impact of prenatal diagnosis of complex congenital heart disease on neonatal and infant morbidity and mortality Prenatal Diagnosis 2018;38: 958‑63

23. Bonnet D. Impacts of prenatal diagnosis of congenital heart diseases on outcomes Transl Pediatr 2021;10(8):2241-49.

24. Day TG, Kainz B, Hajnal J, Razavi R, Simpson JM. Artificial intelligence, fetal echocardiography, and congenital heart disease. Prenatal Diagnosis 2021;41:733–42.

25. Sethi Y, Patel N, Kaka N, Desai A, Kaiwan O, Sheth M et al. Artificial Intelligence in Pediatric Cardiology: A Scoping Review. J. Clin. Med 2022; 11: 7072

26. Arnaout R, Curran L, Zhao Y, Levine JC, Chinn E, Moon-Grady AJ. An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease. Nature Medicine 2021; 27: 882–91.

27. Carroll AE, Downs SM. Improving decision analyses: parent preferences (utility values) for pediatric health outcomes. J Pediatr 2009; 155: 21–5.

28. Day TG, Kainz B, Hajnal J, Razavi R, Simpson JM. Artificial intelligence, fetal echocardiography, and congenital heart disease. Prenat Diagn 2021; 41: 733–42.

29. Lee C, Willis A, Chen C, Sieniek M, Watters A, Stetson B et al. Development of a Machine Learning Model for Sonographic Assessment of Gestational Age. JAMA Network Open 2023;6:1:e2248685.

30. Garcia-Canadilla P, Sanchez-Martinez S, Crispi F, Bijnens B. Machine Learning in Fetal Cardiology: What to Expect. Fetal Diagn Ther 2019 : 149–53.

31. Mori H, Inai K, Sugiyama H, Muragaki Y. Diagnosing Atrial Septal Defect from Electrocardiogram with Deep Learning. Pediatric Cardiology 2021; 42:1379–87.

32. Morris SA and Lopez KN. Deep learning for detecting congenital heart disease in the fetus. Nature Medicine 2021;27; 759–65.

Appendix II. Modelling Survival

Modelling possible survival gains due to pre or timely post-natal diagnosis is particularly complex. Comparing one-year survival rate among those with noncritical CHDs alone (n = 2,455) showed no difference between prenatal and postnatal diagnoses (96% vs 98%, respectively, p = 0.26), whereas among those with critical CHDs (n = 691), prenatally diagnosed infants had significantly lower survival rate, 71% vs 86%, respectively (1).

Among infants with critical CHDs, the adjusted hazard ratio for one -year mortality of those prenatally versus postnatally diagnosed was 2.51 (95% CI 1.72 to 3.66). Prenatal diagnosis was associated with lower one-year survival rates for infants with isolated critical CHDs but showed no change for those with isolated noncritical CHDs. The precise explanation as to why those whose critical CHDs are diagnosed earlier seem to have poorer survival is likely a reflection of the severity of disease, i.e. those that have more severe disease (even within one diagnostic category) are more likely to be diagnosed earlier and are also more likely to have poorer survival. (1-3).

In a similar way, severe diagnoses are more likely to be discovered during the first 24 hours after birth (as opposed to > 24 hours) and therefore can explain the higher infant mortality rate (82.5% vs 71.7% in babies diagnosed after 24 hours) (4). A similar gradient is found in that infants diagnosed after hospital discharge had lower mortality rates than those diagnosed before discharge, who in turn had lower mortality rates than those with a prenatal diagnosis (5).

The superiority of a prenatal diagnosis manifests itself in allowing surgical procedures to be carried out in the early neonatal period (6). When high-risk infants and comfort care infants are excluded, infants with a prenatal diagnosis had far lower pre-operational mortality rates. However, there were more high-risk and comfort-care patients in the prenatal compared with the postnatal diagnosis group (7). One cost-effectiveness analysis was driven by experts' opinion that assumed a 20% higher mortality rate for diagnoses based on the post- as opposed to pre- natal stages (8).

Early diagnosis of CHD is infants is imperative since delayed diagnosis of congenital heart disease worsens the preoperative condition and outcome of surgery in neonates (9). However, survival was found to be lower in those who were diagnosed prenatally than postnatally (10). The explanation being that the more severe defects (with higher mortality rates) are easier to diagnose and consequently minor defects (with lower mortality rates) are harder to diagnose. Therefore, we attempted to control for the diagnosis in order to measure the benefits of diagnosing early (eg: prenatally or in the first day of life). However, even after controlling for diagnosis, more severe cases of the same diagnosis are still be more likely to be discovered before less severe cases, thereby being prone to worst outcomes (11). This phenomena is evidenced by a study (11) that reported lower infant survival rates for prenatal diagnosis (compared to postnatal diagnosis) for Single Ventricular [SV] (46.9% vs 57.1%) and d-transposition of the great arteries (TGA) (91.2% vs 95.8%). However, the study (12) also reported similar survival rates for COA (93.1% vs 93.2%) and non-significant improved survival for TOF (97.2% vs 88.9%).

When analyses are diagnosis-specific there is less (though not zero) selection bias of easily detectable severe cases at the prenatal stage. Overall, after taking into account, those who refused surgery, there was a higher survival rate (63.6% vs 45.5%) in live births, that had a prenatal diagnosis of HLHS (13). Another study reported that prenatal diagnosis resulted in remarkably reduced the pre-operative (22%, 95% CI 6% - 80%) and post-operative mortality (11%, 95% CI 1% - 83%) rates in cases (6) with TGA.

A similar advantage to prenatal screening when a group of subtypes (consisting of TGA, HLHS, SV, TOF and double outlet right ventricle (DORV)) with reductions in the pre-operative (41%, 95% CI 18% - 94%) and post-operative mortality (66%, 95% CI 46% - 94%) rates. In India, prenatal diagnosis, and planned delivery of neonates with critical CHD was associated with significantly lower costs of cardiac care (14). In parallel,

reduced severity when controlling for diagnosis was reflected in a retrospective study of infants with TGA from the USA, that reported 22% higher hospitalization costs in infants without a prenatal diagnosis (15).

References to Modelling Survival

1. Oster ME, Kim CH, Kusano AS, Cragan JD, Dressler P, Hales AR et al. A Population-Based Study of the Association of Prenatal Diagnosis With Survival Rate for Infants With Congenital Heart Defects. Am J Cardiol. 2014; 113:1036–40.

2. Bak GS, Shaffer BL, Madriago E, Allen A, Kelly B, Caughey AB et al. Detection of fetal cardiac anomalies: cost-effectiveness of increased number of cardiac views. Ultrasound Obstet Gynecol 2020; 55: 758–67

3. Han B, Tang Y, Qu X, Deng C, Wang X, Li J. Comparison of the 1-year survival rate in infants with congenital heart disease diagnosed by prenatal and postnatal ultrasound: a retrospective study. Medicine 2021;100:4 (e23325).

4. Oster MR, Lee KA, Honein MA, Riehle-Colarusso T, Shin M, Correa A. Temporal Trends in Survival Among Infants With Critical Congenital Heart Defects Pediatrics 2013 ; 131 :5: e1502–8.

5. Brown KL, Ridout DA, Hoskote A, Verhulst L, Ricci M, Bull C. Delayed diagnosis of congenital heart disease worsens preoperative condition and outcome of surgery in neonates. Heart 2006;92:1298–1302.

6. Li Y-F, Zhou K-Y, Fang J, Wang C, Hua Y-M, Mu D-Z. Efficacy of prenatal diagnosis of major congenital heart disease on perinatal management and perioperative mortality: a meta-analysis. World J Pediatr 2016; 12: 298-307.

7. Holland BJ, Myers JA, Woods CR, Prenatal diagnosis of critical congenital heart disease reduces risk of death from cardiovascular compromise prior to planned neonatal cardiac surgery: a meta-analysis. Ultrasound Obstet Gynecol 2015; 45: 631–8.

8. Mukerji A, Shafey A, Jain A, Cohen E, Shah PS, Sander B,Shah V. Pulse oximetry screening for critical congenital heart defects in Ontario, Canada: a cost-effectiveness analysis. Canadian Journal of Public Health 2020; 111:804–11.

9. Brown KL, Ridout DA, Hoskote A, Verhulst L, Ricci M, Bull C. Delayed diagnosis of congenital heart disease worsens preoperative condition and outcome of surgery in neonates. Heart 2006;92:1298–1302.

10. Han B, Tang Y, Qu X, Deng C, Wang X, Li J. Comparison of the 1-year survival rate in infants with congenital heart disease diagnosed by prenatal and postnatal ultrasound: a retrospective study. Medicine 2021;100:4(e23325).

11. Wright LK, Ehrlich A, Stauffer N, Samai C, Kogon B, Oster ME. Relation of Prenatal Diagnosis With One-Year Survival Rate for Infants With Congenital Heart Disease. Am J Cardiol 2014;113:1041e1044,

12. Khoshnood B, Lelong N, Houyel L, Bonnet D, Ballon M, Jouannic J-M et al Impact of prenatal diagnosis on survival of newborns with four congenital heart defects: a prospective, population-based cohort study in France (the EPICARD Study). BMJ Open 2017;7:e018285.

13 Tworetzky W, McElhinney DB, Reddy VM, Brook MM, Hanley FL. Silverman NH, Bristow J. Improved surgical outcome after fetal diagnosis of hypoplastic left heart syndrome. Circulation 2001; 103:1269-73.

14. Vaidyanathan B, Rani K, Kunde F, Thomas S, Sudhakar A, Kumar RK, Zheleva B. Prenatal diagnosis lowers neonatal cardiac care costs in resource-limited settings. Cardiology in the Young 2022; 32: 1754–60.

15. Gupta D, Mowitz ME, Lopez-Colon D, Nixon CS, Vyas HV, Co-Vu JG. Effect of prenatal diagnosis on hospital costs in complete transposition of the great arteries. Prenatal Diagnosis. 2018;38: 67–571.

Appendix III: Methodology of Calculating the Cost-Utility ratios by Screening Interventions.

In all our supporting meta-analyses, we used a dynamic search methodology, where to take the example of survival:- initially 36 articles were identified (using search terms “prenatal diagnosis’ AND “postnatal diagnosis” AND (survival or mortality) AND “congenital heart”). Of these, only seven “core” articles (1-7) had some relevant information on diagnosis specific survival differences. Next, searches made not only on references in the core articles, but also by means of searching PubMED for later articles that referenced the core article. The newly identified references were in turn used to identify further articles. This process was repeated until no new relevant articles were found. A similar dynamic method was used to construct the meta-analyses of sensitivities and specificities of US and POX screening as well as miscarriage and abortion rates.

The back calculation of the null (no intervention scenario) and expected births by CHD status (serious, minor, or non-CHD) for each intervention involved the following steps.

a) The 162,489 births in the mid-year of the 2005-2014 period (referred to as its mid-point years 2010/11), multiplied by the specific incidence rates of 3.42 and 1.147 per 1000 for serious and minor CHD respectively (8), in order to estimate the numbers of live births by CHD status (serious, minor, and non-CHD).

b) CHD specific miscarriage rates of 9.71% (Appendix XI) and 3.14% (Appendix XII) for serious and minor CHD respectively were derived from meta-analyses of the literature. Based on an estimated overall post - first trimester 3% miscarriage rate (9), an estimate of 2.97% was made for non-CHD fetuses.

c) CHD specific stillbirth rates (1.010%, 0.327% and 0.309% for serious, minor and non-CHD were estimated by applying the overall 0.312% stillbirth rate (in 2010/11), the mid - point of the 2005-2014 era) (10) in proportion to the CHD specific miscarriage rates.

d) The number of fetuses viable after terminations of pregnancies (TOPs or abortions) related to first semester ultrasounds were (back-) calculated based on applying still birth (11) and miscarriage rates (9) to the live birth rates.

e) Abortion data from 2019 (12,13) was applied to livebirth data (10) to estimate the number of abortions that were actually performed in 2011 by CHD status.

f) The actual number of fetuses viable before week 13 by CHD status was calculated by adding the abortion data (e) to the fetal data (d).

g) The number of fetuses by CHD status undergoing ultrasound was estimated by applying the assumed 99% percentage of pregnant women undergoing first trimester ultrasound to the fetal numbers in f).

h) For fetuses undergoing ultrasound, the sensitivity (58.7% and 19.9%) and specificities (99.991% and 99.995%) of sCHD (Appendix XIII) and mCHD (Appendix XIV) first trimester ultrasounds respectively (based on a meta-analysis of publications in the 2005-2014 era) were applied to the fetal numbers (g) and the numbers aborting (e). These provided estimates of 32.4%, 11.6% and 1.5% abortion rates (based on the sum of true positives and false positives) for the for serious, minor, and non-CHD categories respectively.

i) Next the abortion, miscarriage and stillbirth rates were applied to those who had undergone ultrasound in order to estimate the live births by CHD status.

j) Only the miscarriage and stillbirth rates were applied to the 1% who did not undergo ultrasound to estimate the live births by CHD status.

k) The data from 2005-14 was applied to the updated (by means of a meta-analysis based on the 2015-2022 period) ultrasound sensitivity of 71.3% and 15.5% and specificity of 99.986% and 99.992% for sCHD (Appendix V) and mCHD (Appendix VI) respectively. The addition of abortion, miscarriage and still birth rates were then used to estimate the CHD status at birth in 2022.

l) Next, the data in k) was adjusted backwards to a situation where no woman undertook an ultrasound, this is in effect the null (do nothing) scenario which formed the basis for future calculations.

m) The following interventions were applied in turn to the null scenario in order to generate estimates of abortions, miscarriages, stillbirths, and CHD specific estimates (serious, minor, or non-CHD) livebirths:

i) First trimester ultrasound based on sensitivity and specificity derived from a meta-analysis of publications (with and without POX).

ii) First trimester ultrasound using deep learning AI based on a preliminary study on sCHD that provided sensitivity and specificity (14) estimates of 98% and 96% respectively (with and without POX). A sensitivity analysis explored parameters for DL for detecting both sCHD and mCHD.

iii) Use of POX alone, was based on a meta-analysis (Appendix VII) of publications from 2000-2022 [including those identified in a supplement to a Canadian study (15)] where the cutoff point was 95% that did not exclude fetuses that had already obtained positive fetal diagnoses from prenatal ultrasounds.

n) Disease specific one year survival rates by prenatal or postnatal diagnosis (see Appendix II) were based on meta-analyses of the post 2000 literature (see Supplementary Materials). In order to control for diagnoses, the percentage prevalence of sCHD for each specific diagnosis in Israel (8) was used to calculate weighted average specific mortality rates (for prenatal and postnatal diagnoses).

o) Using data on the comparative survival of infants diagnosed <= and > 48 hours after birth (10), the postnatal survival rate for sCHD was decomposed into assumed <= and >24 hours after birth rates in order to capture any relative advantage of POX screening during the first day of life.

p) Survival of mCHD infants by prenatal and postnatal diagnosis was based on a single study (16), with the postnatal rates again adjusted into pre- and post- 24-hour diagnoses.

q) The one-year survival rates were extended (Appendix XV) by integrating meta-analysis data on 1,5,10 and 15 years survival (17) by diagnosis weighted by prevalence rates from an Israeli study (18) for mCHD and by prevalence rates from a national study (8) for sCHD.

r) Our model estimated neonatal mortality rates for sCHD, mCHD and infants without CHD of 116, 23 and 1.2 per 1000 births (11) respectively, which were used in the calculation of maternal QALY losses.

s) National age and gender specific mortality rates (11) were applied to the relative survival rates for 15 year - olds by (CHD type and time of diagnosis) to estimate survival rates for 16-19 years old's and were applied to the relative infant mortality survival rates to estimate survival rates for 20-99 years olds.

Treatment Costs

t) The percentage of sCHD with morbidity and treatment costs by age groups [0 and1-17] were obtained for sCHD (with and without morbidity) based on Canadian 2010/11 price levels (15). These were linked using the Canadian price index and then converted to Israeli shekels in 2022 by the purchasing power parity exchange rate for the 75% of costs were assumed to be non-tradeable services (eg: salaries) and by the exchange rate for the remainder of any tradeable goods.

u) Recent Israeli (18) age group specific utilization data for complex congenital heart defects (by age groups 18-24, 25-44, 45-64, 65+) and by type of care (G.P., Emergency Room and Out Patient visits, ICU and non-ICU hospitalization days) was multiplied by unit cost data [of $15.77, $309, $88, $3,001 and $1,005 respectively] (19) to provide treatment costs by age (by means of interpolation) for sCHD. In a similar fashion, utilization rates for Intermediate CHD (19) were used to estimate mCHD costs.

v) Next, mCHD costs were estimated for the 0-17 year old's by multiplying the estimates for sCHD by 60.3%, representing the ratio of mCHD to sCHD costs in 18 year olds.

w) The resultant annual costs by one year age gradations were multiplied by the percentage surviving to that age (see section r) to give the estimated annual costs by age. Finally, this was discounted using a 3% annual rate to calculate the estimated lifetime costs of sCHD and mCHD by gender and time of diagnosis.

Utilities from fetal or neonatal losses.

x) Disability Weights (DW) of 0.08 were imputed from parental averages (20) to the mother to be for ten years for miscarriages (21), stillbirths (21) and neonatal deaths and for two years for abortions (20,23). All these DW were adjusted by the age specific health status of the mother-to-be.

Utilities of parents to having a child with CHD

y) The percentages of CHD births by neurological developmental disorder (NDD) severity (none, mild, moderate and severe) and time of diagnosis (prenatal and postnatal) was obtained from a recent cost-effectiveness study (20), where there was a greater prevalence of NDD in postnatally as opposed to prenatally diagnosed infants (16.7% vs 5.0%). Multiplication by the CHD birthrate in Israel of 1.147 per 1000 (8) provided an estimate of the numbers in each NDD category.

z) QALY losses of the mother were calculated by applying DWs (20) of 0.05, 0.12,0 .10, and 0.27 over the lifetime (adjusting for age specific non-NDD DW) for none, mild, moderate and severe NDD categories respectively.

aa) Disutility weights for surgery was assumed to be zero, as no parent would conceivably opt for a lower quality state related directly to not undergoing potentially life-saving surgery (16)

ab) We assigned a DW in comparison with parents of healthy children of 0.959 and 0.957 (24) for the QOL of fathers and mothers of CHD children. These were subsequently applied to the age and gender specific HSVs of the patients. We limited this application up to the time the child reached 18 years old.

Utilities losses per CHD case.

ac) Based on WHO Global Burden of Disease estimates (25) for persons with CHD who had NDD, DWs of 0.089, 0.144, and 0.220 were applied over the lifetime (adjusting for age specific non-CHD DW) for mild (mCHD), moderate (mCHD) and sCHD categories respectively.

ad) For persons without NDD, DWs (30) of 0.041, 0.072 and 0.251 were applied over the lifetime (adjusting for age specific non-CHD DW) for mild, moderate and severe levels of NDD respectively.

ae) The average DW of persons with sCHD and mCHD was calculated [from ab & ac] and applied to the age and gender specific background DW of the general population. In turn this was multiplied by the percentage of

persons surviving every year as a result of prenatal, postnatal discovery of CHD in less than 24 hours and discovery after 24 hours.

af) Discounted lifetime HALE was calculated for all the gender and discovery of CHD timing categories. These were then subtracted from the discounted lifetime HALE of the average population in order to estimate the morbidity and mortality losses from CHD by discovery timing and gender.

ag) For each screening category (in addition to the null category) the numbers in each timing and gender category were multiplied by the relevant specific QALY losses in order to estimate the overall screening-specific QALY losses as a result of CHD.

Intervention Costs were based as follows:

ah) US and echocardiography intervention costs of $124.42 and $453 respectively were based on Ministry of Health national price data (26). DL costs, were based on a recent study (27) showing them to be 16.1% or $20.40 more (i.e. $144.82) than ultrasound in the diagnosis of colorectal cancer from colonoscopies.

ai) The cost of POX screening newborns, including provision for 13% repeated tests (28), amounted to $9.58, consisting of $0.21 oximeter & $0.08 probe costs in addition to $9.29 of labour costs (based on 9.8 minutes of nurses time (28) at Ministry of Health based on employment costs of a nurse with five years' experience of $50.30 per hour.

aj) Miscarriage and Stillbirth costs of $6,172 and $5,273 were estimated from a UK meta-analysis (9). TOPs were based on Ministry costs (26) averaging 3,031 nis based on 67.3%, 31.0% and 1.7% (29) undergoing induced surgical abortion ($902), pharmaceutical abortion ($708) and late intra-amniotic injection ($1222) respectively. In addition, an estimated $5,273 of extra costs relating to counselling and work productivity losses were incurred (30). We assumed all TOP, whether voluntary or involuntary will be replaced by the mother to be having another child.

ak) We integrated into the final spreadsheet calculations the intervention costs, treatment costs and loss of QALYs for each of the CHD screening strategies in addition to a theoretical null scenario, where no screening is carried out.

References for Appendix III (methodology):

1. Khoshnood B, Lelong N, Houyel L, Bonnet D, Ballon M, Jouannic J-M, et al. Impact of prenatal diagnosis on survival of newborns with four congenital heart defects: a prospective, population-based cohort study in France (the EPICARD Study). BMJ Open 2017;7:e018285.

2. Eapen RS, Rowland DG, Franklin WH. Effect of prenatal diagnosis of critical left heart obstruction on perinatal morbidity and mortality. Am J Perinatol. 1998;15:237-42.

3. Fuchs IB, Müller H, Abdul-Khaliq H, Harder T, Dudenhausen JW, Henrich W. Immediate and long-term outcomes in children with prenatal diagnosis of selected isolated congenital heart defects. Ultrasound Obstet Gynecol 2007;29: 38-43.

4. Slodki M, Rizzo G, Augustyniak A, Seligman NS, Zych-Krekora K, Respondek-Liberska M & The International Prenatal Cardiology Collaboration Group. Retrospective cohort study of prenatally and postnatally diagnosed coarctation of the aorta (CoA): prenatal diagnosis improves neonatal outcome in severe CoA, The Journal of Maternal-Fetal & Neonatal Medicine 2020; 33: 947-51.

5. Wolter A, Gebert M, Enzensberger C, Kawecki A, Stessig R, Degenhardt J et al. Outcome and Associated Findings in Individuals with Pre- and Postnatal Diagnosis of Tetralogy of Fallot (TOF) and Prediction of Early Postnatal Intervention. Ultraschall in Med 2018; doi https://doi.org/10.1055/a-0753-0008 (Accessed 22nd January 2024).

6. Peake LK, Draper ES, Budd JLS, Field D. Outcomes when congenital heart disease is diagnosed antenatally versus postnatally in the UK: a retrospective population-based study. BMC Pediatrics 2015: 15:58.

7. Lim JSL, McCrindle BW, Smallhorn JF, Golding F, Caldarone CA, Taketazu M et al. Clinical Features, Management, and Outcome of Children With Fetal and Postnatal Diagnoses of Isomerism Syndromes. Circulation. 2005; 112: 2454-61.

8. Israel Ministry of Health. Congenital Malformations in Israel (2000-2014), Quality checks and trends over time. Ministry of health, Jerusalem.

9. Ghosh J, Papadopoulou A, Devall AJ, Jeffery HC, Beeson LE, Do V et al. Methods for managing miscarriage: a network meta-analysis. Cochrane Database Syst Rev. 2021 ;6 :CD012602. doi: 10.1002/14651858.CD012602.pub2.

10. Central Bureau of Statistics. Statistical Abstract of Israel 2016 no. 67, Jerusalem. Statistical Abstract of Israel 2016 - No.67. https://www.cbs.gov.il/en/publications/Pages/2016/Statistical-Abstract-of-Israel-2016-No-67.aspx (Accessed 22nd January 2024).

11. Central Bureau of Statistics. Statistical Abstract of Israel 2022 no. 73, Jerusalem. Statistical Abstract of Israel 2022 - No.73. https://www.cbs.gov.il/en/publications/Pages/2022/Statistical-Abstract-of-Israel-2022-No-73.aspx (Accessed 22nd January 2024).

12. Central Bureau of Statistics. Statistical Abstract of Israel 2021 no. 72, Jerusalem. Statistical Abstract of Israel 2022 - No.73. https://www.cbs.gov.il/en/publications/Pages/2021/Statistical-Abstract-of-Israel-2021-No-72.aspx (Accessed 22nd January 2024).

13. . Central Bureau of Statistics. Applications for Pregnancy Termination in 2019, 29th December 2020, Jerusalem. 05_20_434b.pdf (cbs.gov.il) https://www.cbs.gov.il/en/mediarelease/Pages/2020/Applications-for-Pregnancy-Termination-in-2019.aspx (Accessed 22nd January 2024). (In Hebrew)

14. Arnaout R, Curran L, Zhao Y, Levine JC, Chinn E, Moon-Grady AJ. An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease. Nature Medicine 2021;27: 882–91.

15. Mukerji A, Shafey A, Jain A, Cohen E, Shah PS, Sander B,Shah V. Pulse oximetry screening for critical congenital heart defects in Ontario, Canada: a cost-effectiveness analysis. Canadian Journal of Public Health 2020; 111:804–811 Supplement 2.

16. Han B, Tang Y, Qu X, Deng C, Wang X, Li J. Comparison of the 1-year survival rate in infants with congenital heart disease diagnosed by prenatal and postnatal ultrasound: a retrospective study. Medicine 2021;100:4 (e23325).

17. Best KE, Rankin J . Long-Term Survival of Individuals Born with Congenital Heart Disease: A Systematic Review and Meta-Analysis. J Am Heart Assoc. 2016;5: e002846

18. Benderly M, Buber J, Kalter-Leibovici O, Blieden L, Dadashev A, Lorber A, et al. for the Israeli Adult Congenital Heart Disease Research Group. Health Service Utilization Patterns Among Adults With Congenital Heart Disease. : A Population-Based Study. J Am Heart Assoc. 2021;10:e018037.

19. Ginsberg G, Kark J, Einav S. Is out-of-hospital resuscitation worthwhile? Cost-Utility Analysis of Cardiac Resuscitation Services in Jerusalem. Resuscitation 2015; 86;54-61.

20. Bak GS, Shaffer BL, Madriago E, Allen A, Kelly B, Caughey AB, Pereira L. Detection of fetal cardiac anomalies: cost-effectiveness of increased number of cardiac views. Ultrasound Obstet Gynecol 2020; 55: 758–67

21. Gravensteen IK, Helgadottir LB, Jacobsen E-M, Sandset PM, Ekeberg Ø. Long-term impact of intrauterine fetal death on quality of life and depression: a case-control study. BMC Pregnancy Childbirth 2012; 12: 43.

22. Turton P, Evans C, Hughes P. Long-term psychosocial sequelae of stillbirth: phase II of a nested case-control study. Arch Womens Ment Health 2009; 12: 35–41.

23. Kuppermann M, Nease RF, Learman LA, Gates E, Blumberg B, Washington AE. Procedure-related miscarriages and Down syndrome-affected births: implications for prenatal testing based on women's preferences. Obstet Gynecol 2000; 96: 511–6.

24. Lawoko S, Soares JJF. Quality of life among parents of children with congenital heart disease, parents of children with other diseases and parents of healthy children. Quality of Life Research 2003;12: 655–66.

25. Global Health Data Exchange GHDx. Global Burden of Disease Study 2019 (GBD 2019) Disability Weights. https://ghdx.healthdata.org/record/ihme-data/gbd-2019-disability-weights. Accessed (accessed 2nd May 2022)

26. Ministry of Health Price List for Ambulatory and Hospitalization Services. Effective date: 01/10/2022. Ministry of Health, Jerusalem. https:// www.gov.il/BlobFolder/ dynamiccollectorresultitem/spl-moh-price-list-030722/he/files_databases_moh-price-list_2022_moh-Price-List-010722.xlsx (Accessed 22nd January 2024)

27. Areia M, Mori Y, Correale L, Repici A, Bretthauer M, Sharma P et al. Cost-effectiveness of artificial intelligence for screening colonoscopy: a modelling study. Lancet Digit Health 2022; 4: e436–44

28. Reeder MR, Kim J, Nance A, Krikov S , Feldkamp ML, Randall H et al. Evaluating Cost and Resource Use Associated with Pulse Oximetry Screening for Critical Congenital Heart Disease: Empiric Estimates and Sources of Variation Birth Defects Research (Part A) 2015:103:962–71.

29. Efrati I. Israel's Abortion Rate Falls as Birth Control, Traditional Mores Rise. Haaretz. May 13th 2018. https://www.haaretz.com/israel-news/2018-05-13/ty-article/.premium/israeli-abortion-rate-falls-as-birth-control-traditional-mores-rise/0000017f-ed4c-d4a6-af7f-ffce608f0000 (accessed 16th April 2023).

30. Quenby S, Gallos ID, Dhillon-Smith RK, Podesek M, Stephenson MD, Fisher J et al. Miscarriage matters: the epidemiological, physical, psychological, and economic costs of early pregnancy loss. Lancet 2021; 397: 1658–67.

Appendix IV: Infants surviving post-intervention (a) by serious CHD diagnosis (prenatal or postnatal: studies 2000-23)

	Ref	Live Births (n)	Died	% live births surviving	Weight (b)	Live Births (n)	Died	% live births surviving
AORTIC VALVE STENOSIS		1	1	0%	3.3%	1	1	0%
UNIVENTRICULAR HEART		32	17	46.9%	1.2%	7	3	57.1%
TRUNCUS ARTERIOSIS		19	8	57.9%	1.2%	94	8	91.5%
HLHS		433	125	71.1%	4.3%	539	206	61.8%
PULMONARY ATRESIA		59	9	84.7%	3.3%	45	5	88.9%
AVSD		80	11	86.0%	6.1%	135	15	89.0%
TOF		106	8	92.9%	9.1%	134	8	94.4%
COARCTATION OF AORTA		116	8	93.1%	11.2%	221	20	91.0%
LHO		15	1	93.3%	2.3%	47	9	80.9%
TGA		368	14	96.2%	8.2%	951	76	92.0%
TAPVR		0	0	n.a.	1.2%	1	1	0%
TOTAL (c)		1229	202	**88.3%**	51.6%	2175	351	**87.0%**

HLHS: Hypoplastic Left Heart Syndrome
AVSD: Atrioventricular Septal Defect
TOF : Tetralogy of Fallot
LHO: Left Heart Obstruction, excluding HLHS and Coarctation of Aorta.
TGA: Transposition of the Great Arteries
TAPVR : Total Anomalous Pulmonary Venous Return

(a) post intervention, discharge or up to one year. Just one study was based on five year mortality
(b) based on Israel main text reference (37)
(c) Weighted survival rate of all sCHDs.

Appendix V: Sensitivity and Specificity of Ultrasound studies (2015-2023) to detect Serious CHD.

Under Study Conditions		True Pos	False Pos	False Neg	True Neg	Sensitivity	Specificity
Colosi	2015	0	0	3	5921	0.0%	100.00%
Wiechec	2015	29	0	2	1053	93.5%	100.00%
D'Antonio	2016	2	0	8	2118	20.0%	100.00%
Takita	2016	1	0	1	2006	50.0%	100.00%
Tudorache	2016	21	13	5	2869	80.8%	99.55%
De Robertis	2017	24	6	6	5307	80.0%	99.89%
Vellamkondo	2017	7	0	5	428	58.3%	100.00%
Garcia-Fernandez	2018	4	0	0	655	100.0%	100.00%
Kenkhuis	2018	3	0	6	5005	33.3%	100.00%
Sainz	2018	9	0	1	401	90.0%	100.00%
Zheng	2018	28	1	2	1561	93.3%	99.94%
Chen	2019	52	0	14	10228	78.8%	100.00%
Elbrashy	2019	68	2	12	3158	85.0%	99.94%
Erenel	2019	6	1	0	664	100.0%	99.85%
Duta	2021	29	0	6	6877	82.9%	100.00%
Total		283	23	71	48251	79.9%	99.95%

Notes: All screenings were between 11weeks+0 days and 13 weeks+6days gestation

Under routine conditions

		True Pos	False Pos	False Neg	True Neg	Sensitivity	Specificity
Andrew*	2015	1	0	2	4418	33.3%	100.00%
Syngelaki*	2019	112	0	90	100795	55.4%	100.00%
Vayna*	2018	23	0	6	6016	79.3%	100.00%
Total		136	0	98	111229	58.1%	100.0%

Appendix VI: Sensitivity and Specificity of Ultrasound studies (2015-2023) to detect Minor CHD.

Under study conditions		True Pos	False Pos	False Neg	True Neg	Sensitivity	Specificity
Wiechec	2015	2	0	2	1080	50.0%	100.00%
Takita	2016	1	1	9	1997	10.0%	99.95%
Tudorache	2016	7	4	8	2889	46.7%	99.86%
De Robertis	2017	3	0	2	5338	60.0%	100.00%
Vellamkondo	2017	0	0	11	429	0.0%	100.00%
Kenkhuis	2018	1	1	2	5010	33.3%	99.98%
Sainz	2018	1	0	1	409	50.0%	100.00%
Chen	2019	13	1	49	10231	21.0%	99.99%
Elbrashy	2019	11	3	5	3221	68.8%	99.91%
Erenel	2019	3	1	1	666	75.0%	99.85%
Duta	2021	1	0	3	6908	25.0%	100.00%
Total		43	11	93	38178	31.6%	99.971%

Notes: All screenings were between 11weeks+0 days and 13 weeks+6days gestation

Under routine conditions		True Pos	False Pos	False Neg	True Neg	Sensitivity	Specificity
Vayna*	2018	3	0	3	6039	50.0%	100.00%
Syngelaki*	2019	5	0	182	100810	2.7%	100.00%
Total		8	0	185	106849	4.1%	100.00%

Appendix VII: Pulse Oximetry at Birth for serious CHD (based on 95% cut off) 2000-2022
(Including only studies that did not exclude positive prenatal diagnoses)

		True Pos	False Pos	False Neg	True Neg	Sensitivity	Specificity
Arlettaz	2006	17	7	3	3255	85.0%	99.8%
Bhola	2014	4	11	0	18786	100.0%	99.9%
de Wahl	2005	59	1	7	133	89.4%	99.3%
de Wahl Granelli	2009	19	68	10	39724	65.5%	99.8%
Ewer	2011	18	177	6	19854	75.0%	99.1%
Gomez-Rodriguez	2015	2	12	0	1023	100.0%	98.8%
Jones	2016	2	21	0	10237	100.0%	99.8%
Kawalec	2006	7	13	1	27179	87.5%	100.0%
Klausner	2017	0	4	1	10315	0.0%	100.0%
Kochilas	2013	1	5	0	7543	100.0%	99.9%
Koppel	2003	3	1	2	11275	60.0%	100.0%
Meberg	2008	27	297	8	49676	77.1%	99.4%
Oakley	2015	7	7	1	6314	87.5%	99.9%
Ozalaka	2017	6	1	4	8197	60.0%	100.0%
Richmond	2002	10	54	9	5553	52.6%	99.0%
Ruangritnamchai	2007	3	0	0	1844	100.0%	100.0%
Tautz	2010	9	9	2	3344	81.8%	99.7%
Turska-Kmiec	2012	15	14	4	51665	78.9%	100.0%
Van Nienerk	2016	1	1	1	998	50.0%	99.9%
Zuppa	2014	75	226	9	151	89.3%	40.1%
Badawi	2019	1	27	9	78505	10.0%	100.0%
Saxena	2015	22	6026	4	12957	84.6%	68.3%
Garg	2013	7	42	48	72597	12.7%	99.9%
Lightfoot	2017	0	4	0	720		99.4%
Total		315	7,028	129	441,845	70.95%	98.43%

Appendix VIII: Effects and Costs by Interventions and CHD type.

	Null	US retro studies	POX	US retro & POX	Deep Learning	Deep Learning & POX
Serious CHD (sCHD)						
Viable at 12 weeks	905	905	905	905	905	905
Abortions (a)	14	173	14	173	2778	2778
Miscarried	87	71	87	71	60	60
Stillborn	2	2	2	2	2	2
Live Births	802	659	802	659	560	560
Undiscovered	802	346	233	102	49	16
Prenatal Diagnosis	0	313	0	313	511	511
Postnatal Diagnosis	0	0	569	244	0	33
Diagnosed (% of live births)	0%	47%	71%	85%	91%	97%
Cost per diagnosis (nis)		866,922	348,298	505,494	683,018	660,394
Minor CHD (mCHD)						
Viable at 12 weeks	1561	1,561	1561	1,561	1561	1561
Abortions (a)	24	41	24	41	41	41
Miscarried	48	48	48	48	48	48
Stillborn	5	5	5	5	5	5
Live Births	1485	1,473	1485	1,473	1473	1473
Undiscovered	1485	1,165	1485	1,165	1165	1165
Prenatal Diagnosis	0	308	0	308	308	308
Postnatal Diagnosis	0	0	0	0	0	0
Diagnosed (% of live births)		21%		21%	21%	21%
Cost per diagnosis (nis)		879,171		913,207		

Appendix IX: Additional Cost (million USD at 2022 prices) of DL-US (& POX) vs US (& POX)

(based on US "routine reports": Sensitivity 58.1%, Specificity 100%)

A: Routine	DL-US Specificity						
DL-US Sensitivity	**90%**	**92%**	**94%**	**96%**	**98%**	**99%**	**100%**
80%	40.2	30.4	20.5	10.7	0.9	-4.2	-8.9
84%	37.8	28.0	18.2	8.3	-1.5	-6.5	-11.3
88%	35.4	25.6	15.8	6.0	-3.9	-8.9	-13.7
92%	33.0	23.2	13.4	3.6	-6.3	-11.0	-16.1
96%	30.7	20.8	11.0	1.2	-8.6	-13.4	-18.5
99%	28.9	19.0	9.2	-0.6	-10.4	-15.2	-20.2
Screening Costs : US ($124) DL ($143)							

B. Routine high Cost	DL-US Specificity						
DL-US Sensitivity	**90%**	**92%**	**94%**	**96%**	**98%**	**99%**	**100%**
80%	60.1	50.0	40.2	30.4	20.5	16.5	11.2
84%	57.7	47.9	37.8	28.0	18.2	14.0	8.7
88%	55.4	45.5	35.7	25.9	15.8	11.5	6.2
92%	53.0	43.2	33.3	23.5	14.3	9.0	4.0
96%	50.6	40.8	31.0	21.1	11.8	6.5	1.6
99%	48.8	39.0	29.2	19.3	9.9	4.7	-0.3
Screening Costs : US ($124) DL-US ($248)							

(based on US "research studies": Sensitivity 79.9%, Specificity 99.95%)

C. Routine High Performance & Costs	DL-US Specificity						
DL-US Sensitivity	**90%**	**92%**	**94%**	**96%**	**98%**	**99%**	**100%**
80%	54.2	44.4	34.5	24.7	14.9	9.8	5.1
84%	51.8	42.0	32.1	22.3	12.5	7.4	2.7
88%	49.4	39.6	29.8	19.9	10.1	5.1	0.3
92%	47.0	37.2	27.4	17.6	7.7	2.7	-2.1
96%	44.6	34.8	25.0	15.2	5.4	0.3	-4.5
99%	42.9	33.0	23.2	13.4	3.6	-1.8	-6.3

Screening Costs : US ($280) DL-US ($301)	
DL-US (& POX) costs less than US (& POX)	

Appendix X: Additional QALYS using DL-US (& POX) vs US (& POX)

(based on US "routine reports" data: sensitivity 58.1%, specificity 100%)

A & B.	DL-US Specificity						
DL-US Sensitivity	**90%**	**92%**	**94%**	**96%**	**98%**	**99%**	**100%**
80%	4	202	400	599	797	896	995
84%	185	344	582	781	979	1,078	1,177
88%	367	566	764	963	1,161	1,240	1,359
92%	549	748	946	1,145	1,343	1,442	1,541
96%	731	930	1,128	1,327	1,525	1,624	1,723
99%	868	1,066	1,265	1,463	1,661	1,761	1,860

(based on US research study data: sensitivity 79.9%, specificity 99.95%)

C.	DL-US Specificity						
DL-US Sensitivity	**90%**	**92%**	**94%**	**96%**	**98%**	**99%**	**100%**
80%	-984	-786	-587	-389	-190	-91	8
84%	-802	-604	-405	-207	-9	91	190
88%	-620	-422	-223	-25	173	273	372
92%	-438	-240	-41	157	355	455	554
96%	-256	-58	141	339	537	637	736
99%	-120	79	277	475	674	773	872

DL-US (&POX) provide fewer QALYS than US (& POX)	

Appendix XI: Meta Analysis of Miscarriage Rates (IUD) for Serious (ie: Critical or Major) CHD.

Ref	Ref year	Time of Diagnosis weeks	Fetal CHD (n)	Aborted (n)	Viable Fetus (n)	Mis-Carriage (n)	Mis-Carriage (%)
Vinals	2008	12.5	3	1	2	2	100.0%
Tudorache	2016	13.5	27	20	7	4	57.1%
Michailiadis	2001	17.2	6	4	2	1	50.0%
Eleftheriades (a)	2012	12.5	20	15	5	2	40.0%
Chen	2008	17.5	8	5	3	1	33.3%
Grande	2012	14.4	38	27	11	3	27.3%
Todros	1997	26.4	10	2	8	1	12.5%
Bull	1999	20.0	738	433	305	38	12.5%
Waern	2021	20.0	112	59	53	6	11.3%
Ozkutlu	2005	27.6	37	14	23	2	8.7%
Syngelaki	2011	22.1	79	37	42	3	7.1%
Jin	2021	24.5	124	94	30	2	6.7%
Xie	2017	21.0	59	41	18	1	5.6%
Qiu	2020	24.5	683	493	190	9	4.7%
Jorgensen	2014	20.0	84	51	33	1	3.0%
Gabriel	2002	14.2	35	31	4	0	0.0%
Luck	1992	19.7	24	6	18	0	0.0%
Vanya	2018	14.2	23	18	5	0	0.0%
Orlandi	2014	13.5	19	13	6	0	0.0%
Weiner	2008	13.6	14	12	2	0	0.0%
Rustico	2000	14.3	11	10	1	0	0.0%
Kenkuis	2018	17.0	9	4	5	0	0.0%
Erenel	2019	12.9	7	3	4	0	0.0%
Colosi	2015	22.0	3	0	3	0	0.0%
McAuliff	2005	15.0	2	0	2	0	0.0%
Vellamkondu	2017	15.8	2	1	1	0	0.0%
TOTAL			2,177	1,394	783	76	**9.7%**

(a) includes two embro-reductions of twin fetuses

Note: Miscarriage Rate excluding the four largest studies is 9.80%

Appendix XII: Meta Analysis of Miscarriage Rates (IUD) for Minor CHD.

Ref	Ref (year}	Time of Diagnosis (weeks)	Fetal CHD (n)	Aborted (n)	Viable Fetus (n)	Miscarriages (n)	Miscarriage Rate (%)
Zalel	2016	12.7	12	11	1	1	100.0%
Vellamkondu	2017	12.5	3	0	3	1	33.3%
Luck	1992	19.0	5	0	5	1	20.0%
Waern	2021	19.4	7	0	7	1	14.3%
Orlandi	2014	15.5	11	0	11	1	9.1%
Qiu	2020	24.5	452	290	162	13	8.0%
Syngelaki	2011	23.0	28	2	26	2	7.7%
Xie	2017	21.0	51	26	25	1	4.0%
Jin	2021	24.5	646	107	539	5	0.9%
Tudorache	2016	12.2	12	3	9	0	0.0%
Vanya	2018	16.3	8	2	6	0	0.0%
Grande	2012	19.0	7	5	2	0	0.0%
Todros	1997	26.8	7	0	7	0	0.0%
Ozkutlu	2005	27.6	5	0	5	0	0.0%
Chen	2008	15.9	4	2	2	0	0.0%
Weiner	2008	19.9	4	2	2	0	0.0%
Gabriel	2002	15.3	3	2	1	0	0.0%
Kenkuis	2018	16.5	3	0	3	0	0.0%
Bull	1999	20.0	2	0	2	0	0.0%
Eleftheriades	2012	12.5	6	1	5	0	0.0%
Erenel	2019	12.3	2	0	2	0	0.0%
Rustico	2000	14.3	2	0	2	0	0.0%
TOTAL		23.7	1,280	453	827	26	**3.14%**

Appendix XIII: Sensitivity and Specificity of Ultrasound studies (2005-2014) to detect Serious CHD.

Under study conditions:

Ref	Year	True Pos	False Pos	False Neg	True Neg	Sensitivity	Specificity
McAuliffe	2005	0	0	1	273	0.0%	100.0%
Cedergren	2006	0	0	2	2,706	0.0%	100.0%
Srisupundit	2006	2	0	0	595	100.0%	100.0%
Vimpelli	2006	1	0	1	582	50.0%	100.0%
Dane	2007	0	0	1	1,289	0.0%	100.0%
Lombardi	2007	0	0	3	605	0.0%	100.0%
Li	2007	1	0	1	2,226	50.0%	100.0%
Vinals	2008	3	0	2	30	60.0%	100.0%
Chen (control)	2008	0	0	10	3,683	0.0%	100.0%
Chen (study)	2008	5	5	0	3,939	100.0%	99.9%
Oztekin	2009	0	0	2	1,028	0.0%	100.0%
Benasar *	2009	7	0	0	52	100.0%	100.0%
Sinkovskya	2010	4	0	1	95	80.0%	100.0%
Krapp	2011	17	0	2	671	89.5%	100.0%
Volpe	2011	19	5	6	4,415	76.0%	99.89%
Jacobsen	2011	3	0	24	9,297	11.1%	100.0%
Syngelaki	2011	28	0	62	44,769	31.1%	100.0%
Becker	2012	7	0	8	6,529	46.7%	100.0%
Novotna	2012	1	0	11	8,877	8.3%	100.0%
Grande	2012	25	0	20	13,678	55.6%	100.0%
Eleftheriadis	2012	11	0	1	3,743	91.7%	100.0%
Wang	2013	4	0	1	2,817	80.0%	100.0%
Orlandi	2014	16	0	4	4,010	80.0%	100.0%
Total		154	10	163	115,909	48.6%	99.99%

* adjusted by omitting a few observations over 15 weeks gestation
n.a. Not available since 108 false positive results could not be classified as serious or minor CHD.
Notes: All screenings were between 12weeks+0 days and 13 weeks+3 days gestation

Routine conditions

Ref	Year	True Pos	False Pos	False Neg	True Neg	Sensitivity	Specificity
Abu-Rustum	2010	5	1	1	1,355	83.3%	99.9%
Hartge	2011	66	0	10	3,145	86.8%	100.0%
Total		71	1	11	4,500	86.6%	99.98%

Appendix XIV: Sensitivity and Specificity of Ultrasound studies (2005-2014) to detect Minor CHD.

Under study conditions

Ref	Year	True Pos	False Pos	False Neg	True Neg	Sensitivity	Specificity
Cedergren	2006	0	0	14	2694	0.0%	100.0%
Vimpelli	2006	0	0	4	580	0.0%	100.0%
Dane	2007	1	0	2	1287	33.3%	100.0%
Li	2007	0	0	3	2225	0.0%	100.0%
Vinals	2008	1	3	0	31	100.0%	91.2%
Chen (control)	2008	2	0	5	3686	28.6%	100.0%
Chen (study)	2008	2	0	5	3942	28.6%	100.0%
Oztekin	2009	0	0	1	1029	0.0%	100.0%
Benasar *	2009	3	2	1	53	75.0%	96.4%
Sinkovskya	2010	1	0	1	98	50.0%	100.0%
Volpe	2011	7	1	6	4431	53.8%	100.0%
Jacobsen	2011	9	0	16	9299	36.0%	100.0%
Syngelaki	2011	1	0	17	44841	5.6%	100.0%
Novotna	2012	0	0	3	8886	0.0%	100.0%
Grande	2012	3	0	77	13643	3.8%	100.0%
Eleftheriadis	2012	2	0	4	3749	33.3%	100.0%
Wang	2013	0	0	2	2820	0.0%	100.0%
Orlandi	2014	5	0	7	4018	41.7%	100.0%
Total		37	6	168	107,312	18.0%	99.994%

* adjusted by omitting a few observations over 15 weeks gestation
Note: All screenings were between 12weeks+0 days and 13 weeks+3 days gestation

Routine conditions

Ref	Year	True Pos	False Pos	False Neg	True Neg	Sensitivity	Specificity
Abu-Rustum	2010	3	0	2	1357	60.0%	100.0%
Hartge	2011	3	0	3	3215	50.0%	100.0%
		6	0	5	4572	54.5%	100.0%

Appendix XV: Survival Rates by CHD Diagnosis.

MINOR CHD	1yr	5yr	10yr	15yr
Atrial Septal Defect	90.9%	90.6%	90.0%	89.7%
Pulmonary Valve Stenosis	94.2%	93.6%	93.6%	93.6%
VSD	94.3%	94.2%	92.4%	92.4%
TOTAL (a)	**92.2%**	**91.9%**	**91.0%**	**90.8%**

SEVERE CHD	1yr	5yr	10yr	15yr
Aortic Valve atresia/stenosis	83.3%	81.7%	81.7%	80.0%
AVSD	77.4%	71.9%	71.7%	70.8%
Coartication of Aorta	82.7%	80.2%	79.7%	79.6%
Common Arterial Trunk	69.8%	65.9%	59.1%	59.1%
Ebstein's Anomaly	72.9%	66.0%	66.0%	64.1%
HLHS	48.0%	43.9%	43.9%	39.9%
Pulmonary Valve Atresia	50.0%	43.7%	41.8%	40.1%
Single Ventricle	66.4%	57.7%	57.7%	50.2%
Tetralogy of Fallot	86.8%	86.7%	85.2%	85.1%
TAPR	67.2%	60.6%	60.6%	60.6%
TGV	81.8%	79.1%	76.0%	75.3%
Tricuspid Atresia	68.8%	63.3%	60.7%	58.4%
TOTAL (b)	**57.3%**	**54.9%**	**54.0%**	**52.8%**

AVSD: Atrioventricular Septal Defect
HLHS: Hypoplastic Left Heart Syndrome
TAPR: Total Anomalous Pulmonary Return
TGV: Transposition of the Great Vessels
Based on pooled estimates from a meta-analysis (main references 59).
(a) weighted by Israeli prevalence data (main references 71)
(b) weighted by Israel national prevalence data (main references 65)