

ORIGINAL ARTICLE

Antithyroid drug use during pregnancy and the risk of birth defects in offspring: systematic review and meta-analysis of observational studies with methodological considerations

Daniel R. Morales^{1,2}  | Lionel Fonkwen¹ | Hedvig M. E. Nordeng^{3,4} 

¹Division of Population Health and Genomics, University of Dundee, Dundee, UK

²Department of Public Health, University of Southern Denmark, Denmark

³PharmacoEpidemiology and Drug Safety Research Group, Department of Pharmacy, and PharmaTox Strategic Research Initiative, Faculty of Mathematics and Natural Sciences, University of Oslo, Oslo, Norway

⁴Department of Child Development and Health, Norwegian Institute of Public Health, Oslo, Norway

Correspondence

Dr Daniel Morales, Division of Population Health and Genomics, University of Dundee, Dundee, UK.

Email: d.r.z.morales@dundee.ac.uk; drmorales@health.sdu.dk

Funding information

Wellcome Trust. Grant/Award Number: 214588/Z/18/Z; European Research Council, Grant/Award Number: 639377

Aims: Maternal antithyroid drug (ATD) use during pregnancy has been associated with an increased risk of birth defects in offspring. Uncertainty remains on the size of this risk and how it compares to untreated hyperthyroidism due to methodological limitations of previous studies.

Methods: Systematic review of MEDLINE and EMBASE identifying observational studies examining ATD use during pregnancy and risk of birth defects by 28 August 2020. Data were extracted on study characteristics, effect estimates and comparator groups. Adjusted effect estimates were pooled using a random-effects generic inverse variance method and absolute risk calculated.

Results: Seven cohort studies and 1 case-control study involving 6 212 322 pregnancies and 388 976 birth defects were identified reporting regression effect estimates. Compared to an unexposed population comparison, the association between ATD use during pregnancy and birth defects in offspring was: adjusted risk ratio (aRR) 1.16 95% confidence interval (CI) 1.08–1.25 for propylthiouracil (PTU); aRR 1.28 95%CI 1.06–1.54 for methimazole/carbimazole (MMI/CMZ); aRR 1.51, 95%CI 1.16–1.97 for both MMI/CMZ and PTU; and aRR 1.15 95%CI 1.02–1.29 for untreated hyperthyroidism. The excess risk of *any* and *major* birth defects per 1000, respectively, was: 10.2 and 1.3 for PTU; 17.8 and 2.3 for MMI/CMZ; 32.5 and 4.1 for both MMI/CMZ and PTU; and 9.6 and 1.2 for untreated hyperthyroidism.

Conclusions: When appropriately analysed the risk of birth defects associated with ATD use in pregnancy is attenuated. Although still elevated, the risk of birth defects is smallest with PTU compared to MMI/CMZ and may be similar to that of untreated hyperthyroidism.

KEYWORDS

carbimazole, congenital abnormalities, hyperthyroidism, pregnancy, propylthiouracil

1 | INTRODUCTION

Hyperthyroidism affects between 0.1 and 0.4% of pregnancies and left untreated, may have detrimental maternal and foetal consequences including increasing the risk of preeclampsia, intrauterine growth restriction, preterm birth and maternal heart failure.¹⁻³ The most common cause of hyperthyroidism in young women is Grave's disease, which results from stimulation of the thyroid by thyrotrophin receptor stimulating antibodies. This results in elevated free thyroxine (T4) and/or triiodothyronine (T3) with low thyroid-stimulating hormone (TSH), and typically requires treatment.⁴ In contrast, gestational transient hyperthyroidism resulting from thyroid gland stimulation by the human chorionic gonadotrophin hormone often requires no treatment.

Clinical guidelines recommending treatment of hyperthyroidism during pregnancy parallels that of nonpregnant women with use of antithyroid drugs (ATDs).^{5,6} Commonly prescribed drugs for the management of hyperthyroidism in pregnancy are **methimazole/carbimazole** (MMI/CMZ) and **propylthiouracil** (PTU) that are considered equally effective. Both ATDs can cross the human placenta and has resulted in uncertainty about the risk of birth defects. PTU is the preferred ATD during the first trimester of pregnancy due to rare safety concerns surrounding the possible teratogenic effects of MMI/CMZ.⁷ The reported MMI/CMZ embryofetopathy include aplasia cutis congenita (absence of a portion of skin, often localised on the head), craniofacial birth defects (choanal atresia; facial dysmorphism), defects of the abdominal wall and gastrointestinal tract (exomphalos, oesophageal atresia, omphalomesenteric duct anomaly), and ventricular septal defect.⁸ Consequently, in the first trimester of pregnancy clinical guidelines recommend switching from MMI/CMZ to PTU therapy in women with unplanned pregnancies and then using MMI/CMZ thereafter due to concerns of PTU hepatotoxicity.⁹ The role of PTU during pregnancy has been re-examined following reports of birth defects. These safety concerns recently led to updated warnings from the European Medicines Agency on the use of PTU and MMI/CMZ during pregnancy.^{10,11}

Two meta-analyses were recently published reporting that use of MMI/CMZ and PTU during pregnancy were associated with an increased risk of birth defects.^{12,13} These meta-analyses used the number of cases and the number of people in the sampled population from observational studies to calculate pooled effect estimates for exposure to MMI/CMZ and exposure to PTU during pregnancy. By analysing data from observational studies as if they were trials, such estimates represent crude associations that are likely to be confounded.¹⁴ Heterogeneity between studies may also occur through the use of different study designs and further information from other types of comparator or reference groups may aid decision making. The aim of this study was to examine the association between the use of MMI/CMZ or PTU during pregnancy and the risk of birth defects in offspring, through systematic review and meta-analysis of data using more appropriate methods to support causal inference.

What is already known about this subject

- Maternal antithyroid drug use during pregnancy has been associated with an increased risk of birth defects in offspring.
- Uncertainty remains on the size of this risk and how it compares to untreated hyperthyroidism due to methodological limitations of previous studies, including meta-analyses.

What this study adds

- Previously conducted meta-analyses have over-estimated the risk of birth defects associated with antithyroid drug use during pregnancy due to bias.
- Propylthiouracil use during pregnancy is associated with a smaller risk of birth defects compared to methimazole/carbimazole and may be similar in size as that observed with untreated hyperthyroidism.

2 | METHODS

A review of MEDLINE and EMBASE was conducted using a prespecified search strategy to identify all observational studies (cohort, case-control and sibling studies) assessing the risk of birth defects associated with the use of PTU or MMI/CMZ during pregnancy, published on or before 28 August 2020. The search strategy is reported in the online supplementary material. Titles and abstracts were screened, and full texts of relevant articles assessed for eligibility by 2 authors. Only English language publications and published data were included as we had no resources to translate articles. The systematic review was registered on the EU Register of Post-Authorisation Studies (EUPAS30990) and reported according to PRISMA (Preferred Reporting Items for Systematic Reviews).¹⁵

2.1 | Inclusion and exclusion criteria

First, studies were included if they reported regression effect estimates, and the type of confounding adjustment and study design were described. Of these, meta-analysis was performed using only those studies reporting adjusted effect estimates for the association between ATDs and birth defects. Second, for estimating absolute risk difference (excess number of birth defects per 1000 pregnancies) studies were included if they reported birth defect rates from ATD exposed *and* from an untreated general population comparison group. Studies reporting birth defect rates among other selected populations of women were excluded.

2.2 | Risk of bias

Methodological quality and risk of bias were evaluated for each study using the ROBINS-I tool.¹⁶ This tool is designed to assess the strengths and weaknesses of nonrandomised studies on the effects of interventions in relation to their risk of bias. The ROBINS-I tool covers 7 distinct domains, through which bias could be introduced: confounding, selection of participants into the study, classification of interventions, deviations due to intended interventions, missing data, measurement of outcome measures and selection of the reported result.

2.3 | Data extraction

We extracted data from eligible studies for the following characteristics: study design; study population; sample size; exposure definition; type of comparator or reference group reported; and the accuracy and completeness of information on confounders. For each comparison, crude and adjusted regression effect estimates (odds ratios, hazard ratios or rate ratios) were identified with corresponding 95% confidence intervals (CIs). The outcome of interest were the risk of birth defects in children following MMI/CMZ exposure, PTU exposure, and both MMI/CMZ and PTU exposure during pregnancy.

2.4 | Comparator groups and alternative study designs

Confounding by indication may occur when the underlying indication is associated with the outcome being studied. Using different comparator or reference groups may help to circumvent or minimise this potential confounding.^{17,18} To examine the effect of different comparator groups on such confounding, regression effect estimates for the following prespecified comparator groups were sought:

- i. maternal exposure during pregnancy vs. all unexposed women (also called the unexposed population comparison group, and referred to as the classical comparison);
- ii. maternal exposure during the prepregnancy period (*discontinuers*) vs. all unexposed women (this exposure may act as a negative control and should theoretically be noncausal);
- iii. maternal exposure during pregnancy vs. an unexposed disease comparison group that may reduce confounding by indication or severity (in this context the direct effect of hyperthyroidism);
- iv. siblings analyses with discordant prenatal medication exposure (this study design accounts for all time-fixed within-family confounding);
- v. paternal exposure during pregnancy vs. all unexposed women (this exposure may act as a negative control and should theoretically be noncausal); and
- vi. an unexposed disease comparison group vs. all unexposed women (testing for confounding by indication).^{17,18}

2.5 | Analysis

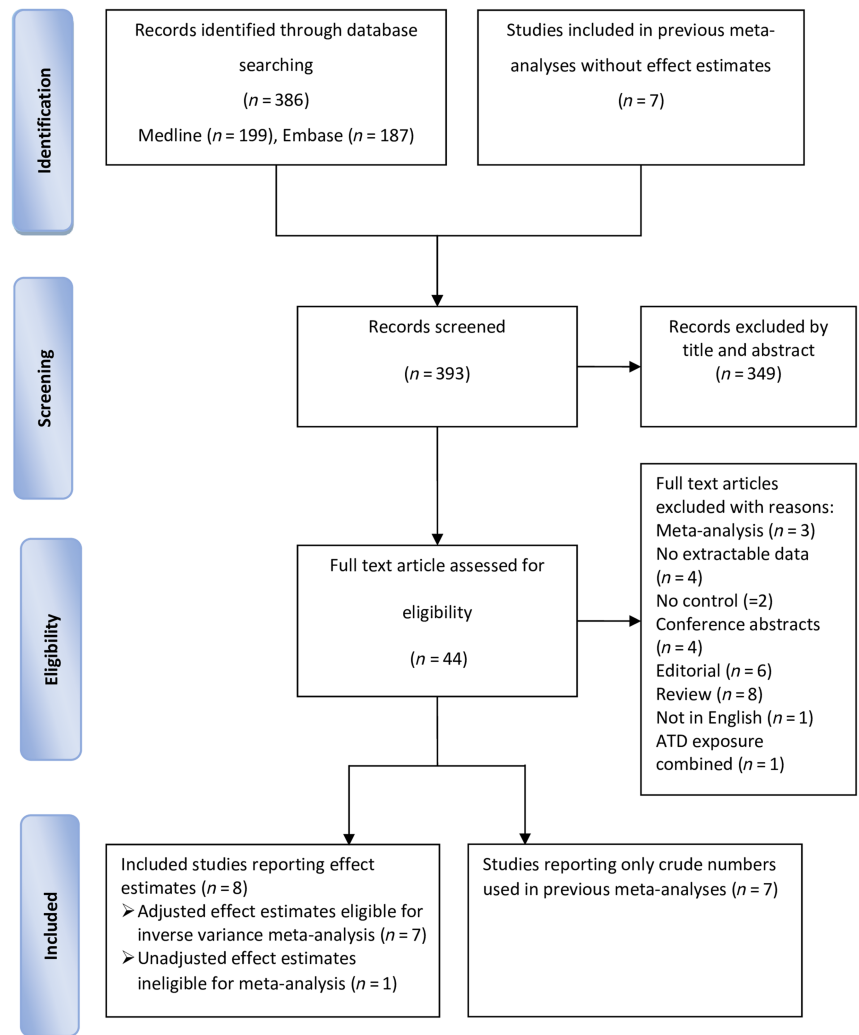
The characteristics of each study and method of confounding adjustment for studies reporting regression effect estimates were first described. The number of cases and total population from each study were then used to replicate the meta-analytical approach used by Li *et al.*, which used a Mantel-Haenzel fixed-effect model.¹⁴ Adjusted effect estimates from all identified studies were then extracted, transformed on the natural log scale and pooled using the generic inverse variance method of analysis. Random-effects models were generated for each type of exposure and comparator group separately. When >1 study used the same data source, the largest study was initially selected with sensitivity analysis substituting this with other studies from the same data source. We also substituted effect estimates for studies that reported data for both *any* birth defect and for a subgroup related to CMZ/MMI. Furthermore, we performed a leave-1-out comparison with 2 studies that reported exposure as *during pregnancy* as opposed to all other studies that specifically measured first trimester exposure. Odds ratios from case-control studies and hazard ratios from cohort studies were combined because they closely approximate each other.^{19,20} For reporting, pooled effect estimates are subsequently referred to as risk ratios (RR) throughout. Publication bias was assessed by testing for funnel-plot asymmetry using the Egger test. Analyses were conducted in Review Manager 5.3 (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014).

To estimate absolute risk differences, the rate of birth defects (*any, major*) from general unexposed population estimates was first extracted from each study and presented as the rate per 1000 live births. The pooled adjusted effect estimates from the meta-analysis were then used to calculate the absolute risk of birth defects (*any, major*) per 1000 live births for each exposure group of interest. The absolute risk difference was then calculated by subtracting the birth defect rate (*any, major*) in the unexposed general population from each exposure group.

3 | RESULTS

The systematic review identified 15 potentially relevant studies. We included 8 studies reporting regression estimates to describe the types of confounding adjustment and study design, and excluded 7 studies without regression estimates (Figure 1). The characteristics of these studies are presented in Table 1 (included studies) and Table S1 (excluded studies).²¹⁻³⁵ Of the included studies reporting effect estimates, 7 were cohort studies (involving 6 195 342 pregnancies and 388 864 birth defects)^{21-26,28} and 1 was a case-control study (involving 16 980 pregnancies and 112 birth defects).²⁷ Two of these studies were conducted using the same overlapping populations from Denmark.^{21,24} Of the comparator groups reported, 5 studies reported adjusted effect estimates for the risk of birth defects associated with use of MMI/CMZ during pregnancy vs. an unexposed population

FIGURE 1 PRISMA (Preferred Reporting Items for Systematic Reviews) flow diagram for study selection process for the meta-analysis



comparison group (the classical comparison), and 1 study compared MMI/CMZ use to an unexposed disease comparison group (Table S2). The corresponding number of studies for PTU were 5 using a classical comparison and 2 using a disease comparison group respectively. For use of both MMI/CMZ and PTU during pregnancy (i.e. switching), 3 studies reported adjusted effect estimates for the classical comparison. No studies were identified that undertook a sibling study design, examined prepregnancy exposure or paternal exposure. However, 1 study used a comparator consisting of either prepregnancy and postpregnancy exposure.

3.1 | Meta-analysis

The meta-analytical approach used by Li *et al.* was first replicated using the number of reported cases and total population with a Mantel-Haenzel fixed-effect model. When analysed in this way, maternal MMI/CMZ exposure and PTU exposure during pregnancy was associated with a significantly increased risk of birth defects in offspring compared to the classical reference group of all unexposed

women (RR 1.20 95%CI 1.02–1.42 for PTU; RR 1.64 95%CI 1.39–1.92 for MMI/CMZ; RR 1.83 95%CI 1.30–2.56 for MMI/CMZ and PTU use, Figure S1).

When only adjusted effect estimates were pooled using the inverse variance method of analysis, maternal use of MMI/CMZ during pregnancy was associated with a smaller but still significantly increased risk of birth defects compared to all unexposed women (random-effects adjusted RR [RaRR] 1.28 95%CI 1.06–1.54, Figure 2). This was similarly the case for use of PTU (RaRR 1.16 95%CI 1.08–1.25) and exposure to both MMI/CMZ and PTU during pregnancy (RaRR 1.51 95%CI 1.16–1.97).

The risk of birth defects among unexposed women with hyperthyroidism compared to unexposed women without hyperthyroidism was also significantly elevated (RaRR 1.15 95%CI 1.02–1.29, Figure 3). When PTU exposure was compared to unexposed women with hyperthyroidism reported in 2 studies, PTU exposure during pregnancy was not associated with an increased risk of birth defects (RaRR 0.72 95%CI 0.47–1.11, Figure 3). In contrast, only 1 study reported this comparison for MMI/CMZ that was significantly elevated (RR 2.28, 95%CI 1.54–3.33).

TABLE 1 Characteristics of the 8 observational studies reporting effect estimates for the risk of birth defects with MMI/CMZ or PTU

Author, publication (Reference)	Type of study design and data source	Country	Study population	Type of birth defect	Timing of exposure	Duration of follow-up in children	ATD exposure (cases/total)	Unexposed women with hyperthyroidism (cases/total)	Unexposed population comparison group (cases/total)
Andersen 2019 ²¹	Cohort study, National registries and birth cohorts	Denmark	Live births	All birth defects and 8 birth defect subgroups diagnosed before age 2 y	6 mo before pregnancy to gestational wk 10	Up to 2 y	PTU (74/889) MMI (151/1574)	Unexposed to ATD (174/951)	Unexposed to ATD (77 791/1 159 181)
Seo 2018 ²²	Cohort study, Claims database	Korea	Live births	All birth defects and 10 birth defect subgroups	During the first trimester	Up to 1 y	PTU (699/9930) MMI (91/1120) MMI/PTU (147/1841)	-	Unexposed to ATD
Andersen 2017 ²³	Cohort study, National registries	Sweden	Live births	All birth defects and 13 birth defect subgroups	6 mo before pregnancy to the end of gestational wk 10	Up to 2 y	PTU (14/218) MMI (11/162) MMI/PTU (4/66)	ATD prior pregnancy and after pregnancy (136/1551)	(170 716/2 872 109) Unexposed to ATD (54 827/682 343)
Andersen 2013 ²⁴	Cohort study, National registries	Denmark	Live births and stillbirths	All birth defects and 8 birth defect subgroups	6 mo before pregnancy to the end of gestational wk 10	Up to 2 y	PTU (45/564) MMI/CMZ (100/1097) MMI/PTU (16/159)	ATD prior pregnancy and after pregnancy (190/3543)	Unexposed to ATD and without hyperthyroidism (45 982/811 730)
Korelitz 2013 ²⁵	Cohort study, Claims database	USA	Live births	Any birth defect	6 mo before pregnancy to the end of pregnancy	Up to 1 y	PTU (66/915) MMI (6/108) MMI/PTU (14/126)	Thyrototoxicosis before/during/after pregnancy (390/5932)	Unexposed to ATD and without hyperthyroidism (37 351/634 858)
Yoshihara 2012 ²⁶	Cohort study, Hospital database	Japan	Live births and perinatal losses of women with Graves' disease	Major birth defects	During the first trimester	1st visit postdelivery	PTU (26/1578) MMI (50/1426)	Graves' disease (40/2065)	-
Chen 2011 ²⁷	Case-control Study, Claims database	Taiwan	Live births of women with hyperthyroidism and age matched controls without	Major birth defects	During pregnancy	Not reported	PTU (5/630) MMI (0/73)	Women with hyperthyroidism (15/2127)	Unexposed to ATD and euthyroid (92/14 150)

TABLE 1 (Continued)

Author, publication (Reference)	Type of study design and data source	Country	Study population	Type of birth defect ^b	Timing of exposure	Duration of follow-up in children	ATD exposure ^c (cases/total)	Unexposed women with hyperthyroidism (cases/total)	Unexposed population comparison group (cases/total)
Rosenfeld 2009 ²⁸	Prospective cohort study by Teratology Information Service	Israel	Live births and elective terminations ^a Controls exposed to nonteratogens	Major birth defects	Gestational wk 4–13	0–6 y	PTU (1/80)	-	Exposed to nonteratogenic drugs (34/1066)

ATD = antithyroid drug; CMZ = carbimazole; MMI = methimazole; PTU = propylthiouracil.

All studies report adjusted effect estimates except Rosenfeld 2009²⁸ that only reports unadjusted effect estimates.

^aElective terminations of pregnancy due to prenatally diagnosed anomalies.

^bBirth defects due to genetic disorders included.

^cAndersen 2013: MMI/CMZ followed by PTU ($n = 149$) or PTU followed by MMI/CMZ ($n = 10$). Andersen 2017: MMI/CMZ followed by PTU ($n = 64$) or PTU followed by MMI/CMZ ($n = 2$). Korelitz 2013. Seo 2018: switching order not specified. Chen 2011: 516/630 and 65/73 received PTU and MMI in the first trimester, respectively.

3.2 | Sensitivity analyses

Two studies were conducted using the same data source from Denmark but covering different time periods.^{21,24} When Andersen *et al.* (2019)²¹ was substituted for Andersen *et al.* (2013),²⁴ pooled effect estimates for MMI/CMZ exposure and PTU exposure were RaRR 1.28 (95%CI 0.96–1.71) and RaRR 1.17 (95%CI 1.09–1.26) respectively. Two studies also reported effect estimates for subgroups of birth defects and exposure to CMZ/MMI and 1 study for PTU (Table S3).^{21,23} When effect estimates for subgroups of birth defects were substituted, pooled effect estimates for exposure to MMI/CMZ was RaRR 1.37 (95%CI 0.98–1.91) and RaRR 1.17 (95%CI 1.09–1.26) for PTU (Figure S2). The results of the leave-1-out analysis for 2 studies measuring ATD exposure during pregnancy rather than specifically as first trimester exposure were similar to the main results (Table S4).

3.3 | Absolute risk

For calculating overall absolute risk differences, we included 4 studies for *any* birth defects and 3 studies for *major* birth defects (Table 2). The overall rate of birth defects in unexposed women was 63.7 per 1000 live births for any birth defect and 8.2 per 1000 for major birth defects (Table 2). The absolute risk difference for any birth defect compared to the unexposed population comparison group per 1000 live births was estimated at: 9.6 in women with unexposed hyperthyroidism; 10.2 in women treated with PTU; 17.8 in women treated with MMI/CMZ; and 32.5 in women treated with both MMI/CMZ and PTU during pregnancy (Table 2). Corresponding numbers for major birth defects per 1000 live births were 1.2, 1.3, 2.3 and 4.1.

3.4 | Assessment of confounding factors and risk of bias

Of the 8 studies, 6 adjusted for maternal age, and 3 for each of the following: the infant's sex; year of birth; parity; and pregnancy type (Table S5). Adjustment for maternal physical history and smoking status occurred in 3 and 2 studies respectively. Adjustment of maternal physical history consisted mainly of metabolic conditions including diabetes and hyperlipidaemia. No studies adjusted for alcohol use during pregnancy and only 1 study adjusted for any other type of medication use during pregnancy, namely antiepileptic use.

For the classical comparison using all unexposed women, the risk of bias varied according to the domain studied and was influenced by variation in the type of confounding adjustment undertaken and uncertainty around exposure and outcome ascertainment bias (Table S6). The Egger test for MMI/CMZ exposed vs. all unexposed women ($P = .020$), PTU exposed vs. all unexposed women ($P = .005$) and PTU and MMI/CMZ exposed vs. all unexposed women ($P = .027$) were statistically significant for funnel plot asymmetry, although the number of included studies evaluating this were limited.

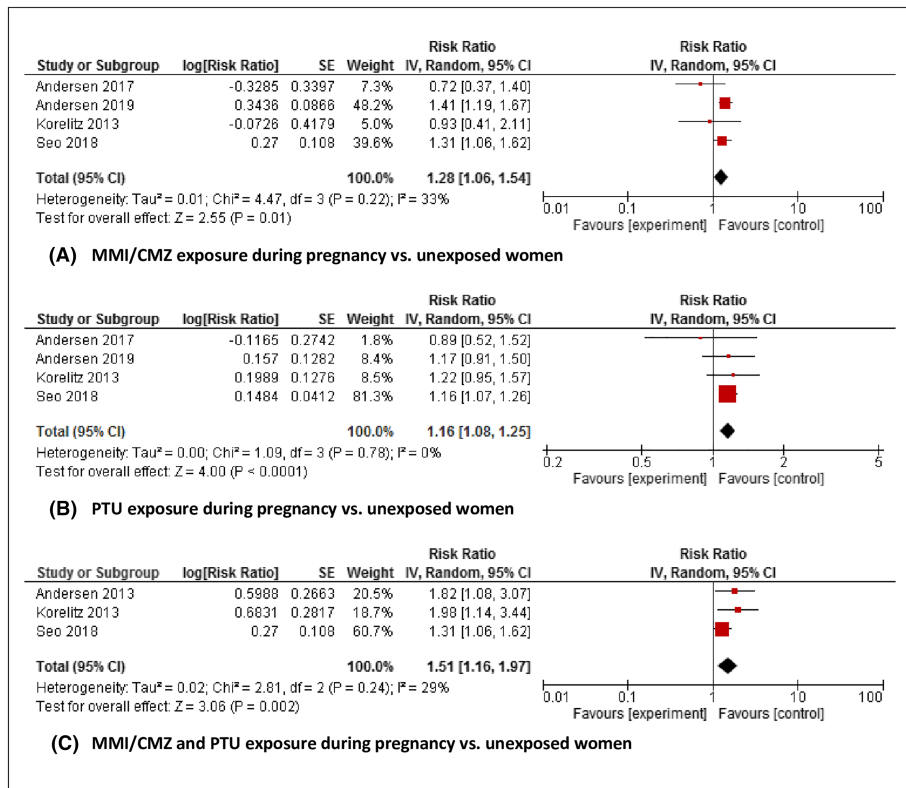


FIGURE 2 Association between maternal antithyroid drug exposure during pregnancy and risk of congenital anomalies in offspring compared to untreated women without hyperthyroidism when analysing adjusted effect estimates. CMZ = carbimazole; MMI = methimazole; PTU = propylthiouracil

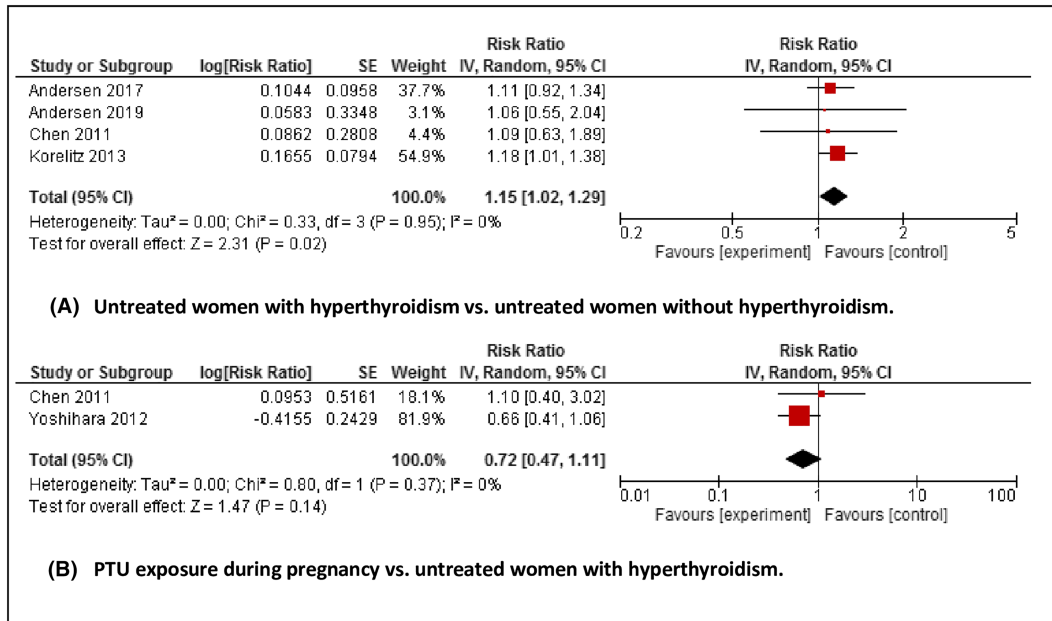


FIGURE 3 Associations between the risk of congenital anomalies in offspring of nonmedicated women with hyperthyroidism compared to all nonmedicated women without hyperthyroidism (A) and propylthiouracil (PTU) exposure during pregnancy compared to all nonmedicated women with hyperthyroidism (B)

4 | DISCUSSION

Meta-analysis is a powerful tool that combines estimates from multiple studies to improve power and precision, whilst allowing questions to be answered that are limited by individual studies. Meta-analysing

data from observational studies examining the risk of ATD exposure during pregnancy as if they were trials may overestimate risk, which attenuates when more appropriate methods are used. However, observed associations between birth defects and exposure to PTU and/or MMI/CMZ were still elevated. Pooled effect estimates for

TABLE 2 Estimated excess number of birth defects per 1000 live births associated with methimazole/carbimazole (MMI/CMZ) exposure, propylthiouracil (PTU) exposure and unexposed hyperthyroidism during pregnancy

Study (reference)	Country	Total number of pregnancies in the unexposed population comparison group	Birth defect rate in unexposed women in the unexposed population comparison group (per 1000 live births)	Estimated excess number of birth defects (per 1000 live births)			
				Unexposed hyperthyroidism	PTU exposure	MMI/CMZ exposure	MMI/CMZ and PTU exposure
<i>Any birth defect</i>							
Andersen 2019 ²¹	Denmark	1 159 181	67.1	10.1	10.7	18.8	34.2
Seo 2018 ²²	Korea	2 872 109	59.4	8.9	9.5	16.8	30.3
Andersen 2017 ²³	Sweden	682 343	80.4	12.1	12.9	22.5	41.0
Andersen 2013 ²⁴	Denmark	811 730	56.7	8.5	9.1	15.9	28.9
Korelitz 2013 ²⁵	USA	634 858	58.8	8.8	9.4	16.5	30.0
	Overall*	5 001 040	63.7	9.6	10.2	17.8	32.5
<i>Major birth defect</i>							
Chen 2011 ²⁷	Taiwan	14 150	6.5	1.0	1.0	1.8	3.3
Lian 2005 ³²	China	22 765	9.3	1.4	1.5	2.6	4.7
Momotani 1984 ³⁵	Japan	350	2.9	0.4	0.5	0.8	1.5
	Overall	40 095	8.2	1.2	1.3	2.3	4.1

*Overall rate for any birth defect calculated using Andersen 2019²¹ and excluding Andersen 2013²⁴ due to the overlapping population.

**Yoshihara 2012,²⁶ Rosenfeld 2009,²⁸ Hawken 2016,²⁹ Lo 2015,³⁰ Clementi 2010,³¹ Gianantonio 2001³³ and Wing 1994³⁴ were excluded from the calculation of overall rate estimates and absolute risk differences as they did not report baseline birth defect rates from an unexposed general population comparison group. Exposure not only restricted to the first trimester (Korelitz 2013, Chen 2011).

women using PTU and unexposed women with hyperthyroidism were similar in size, whilst the highest risk estimates were observed in the groups switching between MMI/CMZ and PTU.

PTU is typically considered the safer option when hyperthyroidism requires treatment during the first trimester of pregnancy. However, safety concerns surrounding PTU exposure and the risk of hepatic toxicity and birth defects have emerged leading to regulatory updates to the product information in 2011 and 2019 respectively.^{9–11} Our findings for PTU exposure during pregnancy are reassuring as they suggest that it does not increase the risk of birth defects beyond that of the underlying maternal disease, but potentially has the advantage of reducing maternal morbidity associated with untreated hyperthyroidism according to previous research.³

The highest risk of birth defects was associated with use of both MMI/CMZ and PTU during pregnancy that corresponds to women who switch from 1 product to the other. The effects of switching can be difficult to interpret with these data as the order of switching was not always specified. However, in the majority of instances, switches appeared to be appropriate as most changed

from MMI/CMZ to PTU during the first trimester of pregnancy. Given current guideline recommendations, this could represent a group more likely to include women with unplanned pregnancies that may differ in terms of lifestyle and early folic acid use for example. Whilst we cannot exclude the possibility that confounding by the severity of hyperthyroidism plays a role we would expect treated women to become euthyroid as their disease is controlled with ATD medication.

Our analysis demonstrates that using adjusted effect estimates is important to appropriately account for measured confounding factors and that risk estimates from meta-analyses based upon crude numbers may be of limited value for clinical and regulatory decision making. We recommend avoiding this method of meta-analysis for nonrandomised studies when adjusted effect estimates from observational studies are available. This is particularly important when such estimates are used to calculate absolute risk. In this regard, individual observational studies should be properly designed and conducted to provide the appropriate primary data, including being adequately powered to reduce the risk of a type-II error.

This study has several potential limitations. First, although accounting for confounding attenuated the observed associations, the included studies were heterogeneous in their approach to confounding adjustment suggesting that residual confounding remains possible. First trimester exposure was specifically evaluated in most but not all studies. However, the leave-1-out analysis produced similar results suggesting that these estimates are relevant to first trimester exposure. Studies were also heterogeneous in the outcome definition with some studies measuring *any* birth defects (also minor) whilst other focused on *major* birth defects. The baseline birth defect rate was as high as 8% in 1 study, probably due to inclusion of neonatal anomalies that according to the EuroCat classification system are *minor anomalies for exclusion*.^{23,36} Heterogeneity may have also resulted from studies using different types of data source. Other features that could have affected the risk estimates include not excluding offspring with birth defects due to genetic disorders. This may introduce bias if women with hyperthyroidism/Graves disease have an increased risk of having children with birth defects of genetic origin.³⁷ Moreover, ascertainment bias related to the outcome is expected as clinicians may scrutinize children of women with hyperthyroidism to a greater extent than children of healthy women, especially after the signals of ATD teratogenicity. Although this could have a larger effect on studies with shorter follow-up to detect outcomes we observed no obvious impact of this with the included studies. For these reasons, it is important to consider using an unexposed disease comparison group as an alternative comparator to better infer causality.^{17,18} Notably no sibling study designs have been reported despite the potential advantage of controlling for shared familial confounding by design.

The majority of included studies assessed only birth defects in live births, which will miss defects due to elective termination, miscarriages or stillbirths. Conditioning on live births may give rise to collider bias if both exposure and outcome are associated with survival.³⁸ Despite the advantages of having an unexposed disease comparison group, this approach may not account for differences in the severity or type of hyperthyroidism. It is likely that transient gestational hyperthyroidism may represent a significant proportion of untreated cases whereas Graves disease will probably be treated. Although we demonstrated comparable risk from PTU exposure and untreated hyperthyroidism, the overall number of studies was small and further observational studies may be useful, including those that explore the relationships between different comparator groups and study designs. We also cannot exclude the possibility of publication bias. Finally, this study examines birth defects as a composite outcome and data may not be generalizable to subtypes of birth defects such as MMI embryofetopathy. However, whilst MMI/CMZ may be associated with a larger risk of birth defects not all studies consistently reported an elevated risk demonstrating that some uncertainty about this association remains. For example, in a large study from Sweden there were no cases of aplasia cutis and choanal or oesophageal atresia that have described as being the MMI embryopathy.²³ The only study that was sufficiently powered to assess this birth defect pattern, however, found 5 additional cases of MMI embryopathy per 10 000 after first trimester exposures to ATDs compared to nonexposed (0.12 vs.

0.07%).²² This was also the only study to report a dose–response relationship and observed larger risk estimates for congenital anomalies with high cumulative doses of CBZ/MMI but not with PTU.

Future studies should evaluate the association between ATD and birth defects using comparable definitions of exposure and outcomes aligned with current recommendations and standards.^{36,39} Estimates should be reported separately for CMZ/MMI and PTU rather than by ATD group that is still evaluated.⁴⁰ Major groups of birth defects should also be investigated separately rather than *any* birth defect, and minor birth defects should preferably be studied as part of malformation patterns. Future studies should also consider incorporating additional types of comparators or reference groups in their designs (such as untreated hyperthyroidism or prepregnancy ATD exposure), ensure data sources have reasonable power to study the birth defects of interest, and consider the use of sibling study designs to address familial confounding. Techniques to quantify residual confounding for example by severity of maternal hyperthyroidism and unmeasured confounders could also be useful.¹⁸

Assuming MMI/CMZ and PTU are equally effective, our findings suggest that PTU use is the safer option in terms of the risk of birth defects in offspring, and that untreated hyperthyroidism may also carry a risk. This suggests that the net benefit of managing hyperthyroidism during the first trimester of pregnancy is in favour of using PTU when clinically indicated. However, such use must also be balanced against other potential risks with PTU such as hepatotoxicity and the limitation of current observational studies.

5 | CONCLUSION

Previous meta-analyses examining the risk of maternal ATD use during pregnancy overestimate the risk of birth defects in offspring. When appropriately analysed this risk is attenuated. Although still elevated, the risk of birth defects is smallest with use of PTU during pregnancy.

ACKNOWLEDGEMENTS

We thank the reviewers of the British Journal of Clinical Pharmacology for the valuable input and thorough review of our manuscript. No funding was obtained to conduct this study. H.N. is supported by the European Research Council Starting Grant “DrugsInPregnancy” grant number 639377. D.M. is supported by a Wellcome Trust Clinical Research Development Fellowship (Grant 214588/Z/18/Z).

COMPETING INTERESTS

The authors have no conflicts of interest in connection with this article. D.M. and H.N. are current independent scientific expert members of the European Medicines Agency Pharmacovigilance Risk Assessment Committee.

CONTRIBUTORS

D.M. and L.F. collected the data and performed the analysis with guidance from H.N. D.M. is the guarantor for the study. All authors

contributed to the study design, interpretation of results, writing the manuscript and approved the final draft.

DATA AVAILABILITY STATEMENT

All data are either contained within the manuscript or can be accessed via the referenced publications.

ORCID

Daniel R. Morales  <https://orcid.org/0000-0002-0063-8069>

Hedvig M. E. Nordeng  <https://orcid.org/0000-0001-6361-2918>

REFERENCES

- Davis LE, Lucas MJ, Hankins GD, Roark ML, Cunningham FG. Thyrotoxicosis complicating pregnancy. *Am J Obstet Gynecol.* 1989;160(1):63-70.
- Millar LK, Wing DA, Leung AS, Koonings PP, Montoro MN, Mestman JH. Low birth weight and preeclampsia in pregnancies complicated by hyperthyroidism. *Obstet Gynecol.* 1994;84(6):946-949.
- Luewan S, Chakkabut P, Tongsong T. Outcomes of pregnancy complicated with hyperthyroidism: a cohort study. *Arch Gynecol Obstet.* 2011;283(2):243-247.
- Lazarus JH. Hyperthyroidism during pregnancy: etiology, diagnosis and management. *Womens Health.* 2005;1(1):97-104.
- Kahaly GJ, Bartalena L, Hegedüs L, Leenhardt L, Poppe K, Pearce SH. 2018 European Thyroid Association Guideline for the Management of Graves' Hyperthyroidism. *Eur Thyroid J.* 2018;7(4):167-186. <https://doi.org/10.1159/000490384> Epub 2018 Jul 25. PMID: 30283735; PMCID: PMC6140607
- Alexander EK, Pearce EN, Brent GA, et al. 2017 Guidelines of the American Thyroid Association for the Diagnosis and Management of Thyroid Disease During Pregnancy and the Postpartum. *Thyroid.* 2017;27(3):315-389.
- Methimazole embryofetopathy. The portal for rare diseases and orphan drugs. https://www.orpha.net/consor/cgi-bin/OC_Exp.php?Lng=GB%26;Expert=1923. Accessed December 13, 2020.
- Seo GH, Kim TH, Chung JH. Antithyroid Drugs and Congenital Malformations: A Nationwide Korean Cohort Study. *Ann Intern Med.* 2018;168(6):405-413.
- U.S. Food and Drug Administration. FDA Drug Safety Communication: New Boxed Warning on severe liver injury with propylthiouracil. 2010. <https://www.fda.gov/drugs/postmarket-drug-safety-information-patients-and-providers/fda-drug-safety-communication-new-boxed-warning-severe-liver-injury-propylthiouracil>. Accessed December 13, 2020.
- PRAC recommendations on signals. November 2018. https://www.ema.europa.eu/en/documents/prac-recommendation/prac-recommendations-signals-adopted-26-29-november-2018-prac-meeting_en.pdf. Accessed December 13, 2020.
- PRAC recommendations on signals. June 2019. https://www.ema.europa.eu/en/documents/prac-recommendation/prac-recommendations-signals-adopted-11-14-june-2019-prac-meeting_en.pdf. Accessed December 13, 2020.
- Song R, Lin H, Chen Y, Zhang X, Feng W. Effects of methimazole and propylthiouracil exposure during pregnancy on the risk of neonatal congenital malformations: A meta-analysis. *PLoS One.* 2017;12(7):e0180108. <https://doi.org/10.1371/journal>
- Li H, Zheng J, Luo J, et al. Congenital anomalies in children exposed to antithyroid drugs in-utero: a meta-analysis of cohort studies. *PLoS One.* 2015;10(5):e0126610.
- Reeves BC, Deeks JJ, Higgins JPT, Shea B, Tugwell P, Wells GA. Chapter 24: Including non-randomized studies on intervention effects. In: Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)*. Cochrane; 2019 Available from www.training.cochrane.org/handbook.
- Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 2009;6(7):e1000097.
- Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ.* 2016;355:i4919.
- Morales DR, Slattery J, Evans S, Kurz X. Antidepressant use during pregnancy and risk of autism spectrum disorder and attention deficit hyperactivity disorder: systematic review of observational studies and ethodological considerations. *BMC Med.* 2018;16(1):6.
- Wood ME, Lapane KL, van Gelder MMHJ, Rai D, Nordeng HME. Making fair comparisons in pregnancy medication safety studies: An overview of advanced methods for confounding control. *Pharmacoepidemiol Drug Saf.* 2018;27(2):140-147.
- D'Agostino RB, Lee ML, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. *Stat Med.* 1990;9(12):1501-1515.
- Lanza A, Ravaut P, Riveros C, Dechartres A. Comparison of Estimates between Cohort and Case-Control Studies in Meta-Analyses of Therapeutic Interventions: A Meta-Epidemiological Study. *PLoS One.* 2016;11(5):e0154877.
- Andersen SL, Knøsgaard L, Olsen J, Vestergaard P, Andersen S. Maternal thyroid function, use of antithyroid drugs in early pregnancy and birth defects. *J Clin Endocrinol Metab.* 2019;04(12):6040-6048.
- Seo GH, Kim TH, Chung JH. Antithyroid Drugs and Congenital Malformations: A Nationwide Korean Cohort Study. *Ann Intern Med.* 2018;168(6):405-413.
- Andersen SL, Lonn S, Vestergaard P, Torring O. Birth defects after use of antithyroid drugs in early pregnancy: a Swedish nationwide study. *Eur J Endocrinol.* 2017;177(4):369-378.
- Andersen SL, Olsen J, Wu CS, Laurberg P. Birth defects after early pregnancy use of antithyroid drugs: a Danish nationwide study. *J Clin Endocrinol Metab.* 2013;98(11):4373-4381.
- Korelitz JJ, McNally DL, Masters MN, Li SX, Xu Y, Rivkees SA. Prevalence of thyrotoxicosis, antithyroid medication use, and complications among pregnant women in the United States. *Thyroid.* 2013;23(6):758-765.
- Yoshihara A, Noh J, Yamaguchi T, et al. Treatment of graves' disease with antithyroid drugs in the first trimester of pregnancy and the prevalence of congenital malformation. *J Clin Endocrinol Metab.* 2012;97(7):2396-2403.
- Chen CH, Xirasagar S, Lin CC, Wang LH, Kou YR, Lin HC. Risk of adverse perinatal outcomes with antithyroid treatment during pregnancy: a nationwide population-based study. *BJOG.* 2011;118(11):1365-1373.
- Rosenfeld H, Ornoy A, Shechtman S, Diav-Citrin O. Pregnancy outcome, thyroid dysfunction and fetal goitre after in utero exposure to propylthiouracil: a controlled cohort study. *Br J Clin Pharmacol.* 2009;68(4):609-617.
- Hawken C, Sarreau M, Bernardin M, et al. Management of Graves' disease during pregnancy in the Poitou-Charentes Region. *Ann Endocrinol (Paris).* 2016;77(5):570-577.
- Lo JC, Rivkees SA, Chandra M, Gonzalez JR, Korelitz JJ, Kuzniewicz MW. Gestational thyrotoxicosis, antithyroid drug use and neonatal outcomes within an integrated healthcare delivery system. *Thyroid.* 2015;25(6):698-705.
- Clementi M, Di Gianantonio E, Cassina M, Leoncini E, Botto LD, Mastroiacovo P. Treatment of hyperthyroidism in pregnancy and birth defects. *J Clin Endocrinol Metab.* 2010;95(11):E337-E341.

32. Lian XL, Bai Y, Xu YH, Dai WX, Guo ZS. Effects of maternal hyperthyroidism and antithyroid drug therapy on congenital malformation of newborn infants. *Chin J Obstet Gynecol*. 2005;40:511-515.
33. Gianantonio ED, Schaefer C, Mastroiacovo PP, Cournot MP, Benedicenti F, Reuvers M. Adverse effects of prenatal methimazole exposure. *Teratology*. 2001;64(5):262-266.
34. Wing DA, Millar LK, Koonings PP, Montoro MN, Mestman JH. A comparison of propylthiouracil versus methimazole in the treatment of hyperthyroidism in pregnancy. *Am J Obstet Gynecol*. 1994;170(1):90-95.
35. Momotani N, Ito K, Hamada N, Ban Y, Nishikawa Y, Mimura T. Maternal hyperthyroidism and congenital malformation in the offspring. *Clin Endocrinol (Oxf)*. 1984;20(6):695-700.
36. EUROCAT. European Surveillance of Congenital Anomalies. Guide 1.3 and reference documents. Instructions for the Registration and Surveillance of Congenital Anomalies. 2005. <https://eu-rd-platform.jrc.ec.europa.eu/sites/default/files/EUROCAT-Guide-1.3.pdf>. Accessed December 13, 2020.
37. Genetics Home Reference. 22q11.2 deletion syndrome. <https://ghr.nlm.nih.gov/condition/22q112-deletion-syndrome>. Accessed December 13, 2020.
38. Liew Z, Olsen J, Cui X, Ritz B, Arah OA. Bias from conditioning on live birth in pregnancy cohorts: an illustration based on neurodevelopment in children after prenatal exposure to organic pollutants. *Int J Epidemiol*. 2015;44(1):345-354.
39. Guideline on good pharmacovigilance practices (GVP). Product- or Population-Specific Considerations III: Pregnant and breastfeeding women. EMA/653036/2019 DRAFT FOR PUBLIC CONSULTATION, 4 December 2019. http://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-good-pharmacovigilance-practices-product-population-specific-considerations-iii_en.pdf. Accessed December 13, 2020.
40. Turunen S, Väärasmäki M, Lahesmaa-Korpinen AM, et al. Maternal hyperthyroidism and pregnancy outcomes: A population-based cohort study. *Clin Endocrinol (Oxf)*. 2020;93(6):721-728. <https://doi.org/10.1111/cen.14282>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Morales DR, Fonkwen L, Nordeng HME. Antithyroid drug use during pregnancy and the risk of birth defects in offspring: systematic review and meta-analysis of observational studies with methodological considerations. *Brit Jnl Clinical Pharma*. 2021;87(10):3890–3900. <https://doi.org/10.1111/bcp.14805>