Weaponised uranium and adverse health outcomes in Iraq: a systematic review

Shelby Surdyk,1 Moustapha Itani,2 Mais Al-Lobaidy,1 Lara A Kahale,3 Aida Farha,4 Omar Dewachi,5 Elie A Akl,6 Rima R Habib 1

ABSTRACT

Background The US military first deployed depleted uranium (DU) weapons in Iraq during the Gulf War in 1990 and in the 2003 invasion of Iraq. Research into the health impacts of DU has been mired in debate and controversy. Research funded by the US government has denied the health risks posed by DU to the Iraqi population, while opponents have claimed that DU is responsible for increased rates of birth defects and cancers in Iraq. Others assert that the public health impacts of DU weapons remain uncertain. This systematic review identified, appraised and synthesised all human observational studies assessing adverse health outcomes associated with DU exposure among the Iraqi population. To our knowledge, no systematic review has been conducted on the topic previously.

Methods We searched 11 electronic databases for human observational studies published between 1990 and 2020 that measured association between exposure to weaponised uranium and health outcomes (including cancer, birth defects, immune system function and mortality) among the Iraqi population. We assessed risk of bias using the Navigation Guide’s risk of bias tool and rated certainty of the evidence using the Grading of Recommendations, Assessment, Development and Evaluations approach (PROSPERO: CRD42018100225).

Results Our searches identified 2601 records, of which 28 met our inclusion criteria. We identified five additional eligible reports from other sources. Two articles reported the results of multiple relevant studies; our final set included 33 articles reporting on 36 eligible studies. Most studies (n=30, 83%) reported a positive association between uranium exposure and adverse health outcomes. However, we found that the reviewed body of evidence suffers from a high risk of bias.

Conclusion The available evidence suggests possible associations between exposure to depleted uranium and adverse health outcomes among the Iraqi population. More primary research and the release of missing data are needed to design meaningful health and policy interventions in Iraq.

INTRODUCTION

Depleted uranium (DU) is a heavy metal possessing chemotoxic and radiotoxic properties. In nature, uranium exists in three isotopic forms, U-238, U-235 and U-234, with mass percentages of 99.284%, 0.711% and 0.005%, respectively. Of those, only U-235 is fissile, that is, will readily undergo a nuclear chain reaction. In order to produce fuel for commercial nuclear energy reactors, the mass percentage of U-235 in natural uranium (0.711%) must be increased through the enrichment process. The US Nuclear Regulatory Commission defines highly enriched...
uranium as uranium with a mass percentage of U-235 of 20% or greater, enriched uranium as uranium with a mass percentage of U-235 of 2%-5% and depleted uranium as uranium with an abundance of U-235 less than 0.7% (by mass). DU is created as a by-product of uranium enrichment and contains fewer U-235 isotopes than natural uranium.

Due to its density, pyrophoricity (ie, the ability to ignite on impact) and availability, the US military deployed DU weapons in 1991 during the first Gulf War in Iraq and again in the 2003 US invasion of Iraq. Recent studies have suggested that the US military may also use slightly enriched uranium (U-235 mass percentage >0.711%, <2%) in conventional weapons in Iraq. Therefore, we use the term weaponised uranium to refer to metallic uranium (of uncertain isotopic composition) that has been introduced to the environment in Iraq via the use of conventional weapons (ie, non-nuclear missiles, bullets and armour).

The possible routes of exposure to weaponised uranium among the Iraqi population are ingestion, inhalation, dermal contact and embedded fragments (eg, shrapnel). On impact with a hard target (eg, an armoured tank), weaponised uranium generates a cloud of uranium oxide particles of varying solubility, a property known as pyrophoricity. Exposure to uranium oxides via inhalation or ingestion has been shown to induce carcinogenic and teratogenic effects in non-human in vivo animal studies. After absorption, uranium oxides are metabolised to uranyl ions that readily bond with other molecules, including proteins and bioligands, and are transported throughout the body via systemic circulation. Most uranium that is absorbed is excreted in urine within 24 hours, but that which remains is stored in the bones, the kidneys and the liver. Uranium is also capable of crossing the blood–brain barrier and the placental barrier. Once absorbed, uranium has been shown to increase the presence of reactive oxygen species, break DNA strands and alter gene expression leading to adverse clinical effects. Chronic exposure scenarios put local populations at greater risk to adverse health effects than veteran populations. According to the WHO, children might be at greater risk of exposure via ingestion due to hand-to-mouth transfer of dust and uranium oxide particles. Other vulnerable groups include scrap metal workers and civilians living near highly contaminated areas such as the Umm Qasr tank storage area in Southern Iraq.

Exactly where, or how much, DU was used in Iraq by US forces since 1991 is still not fully known. The International Atomic Energy Agency (IAEA) estimates that between 170 and 1700 t of DU was deployed in Iraq by the US military since 2003, while other conservative estimates place the total amount of DU used by the US military in Iraq since 1991 at 440 t. The UK has also reported firing 1.9t DU weapons in Iraq since 2003. More than 300 DU-contaminated sites in Iraq have been identified by the Iraqi Ministry of the Environment Radiation Protection Center (RPC). While all sites identified by the RPC were located south of Baghdad, additional research has reported high concentrations of uranium in soil samples from Mosul. In 2014, the Dutch NGO PAX published a set of US DU firing coordinates in the Dutch area of operation in Basrah that were provided at the request of the Dutch Ministry of Defense out of concern for their military personnel serving in the area. Additionally, in 2016, George Washington University and the Dutch NGO PAX released data on DU firing coordinates obtained via a Freedom of Information Act request. The data contain precise locations of 783 out of 1116 airstrikes carried out by DU-armed A-10 Warthogs in Iraq between 20 March and 15 April 2003. While the data do not contain total amounts of DU rounds fired during each strike, they do indicate a far more widespread use of DU during the 2003 Iraq War than had previously been indicated. The target locations are scattered across Iraq, in every governorate.

While many recent reviews attempted to summarise the evidence regarding toxicity of weaponised uranium, only one was conducted systematically—and it focused on US Gulf War veterans. Compared with veterans, the Iraqi population has received scant attention in the DU literature, due partly to the politicisation of DU research. In 2013, Webster PC published a World Report in The Lancet raising concern about a national congenital birth defect (CBD) study carried out by the Iraqi Ministry of Health (IMOH) in 2012. The IMOH study found no proof of unusually high CBD incidence in the country, although Iraqi doctors in DU-impacted zones had been voicing concerns for decades. Webster sharply criticised the IMOH’s methodology, its (lack of) peer-review process, and cited three studies on CBD incidence in Iraq whose findings disagreed with the IMOH report. This controversy speaks to the need for a systematic review (SR) of all the available evidence. Previous (non-systematic) reviews on this topic have lacked comprehensive, transparent search strategies and inclusion criteria, which may have introduced bias into the reviews.

The present study aimed to systematically review the evidence on the associations between weaponised uranium exposure and adverse health impacts among the Iraqi population and to judge the certainty of the evidence. SRs have long been considered the standard approach to synthesising a body of evidence in the field of clinical medicine and are increasingly considered so in the field of environmental health. Their thorough, transparent and reproducible methods make SRs the ideal tool for approaching controversial and politicised questions of environmental exposures.

METHODS
Eligibility search strategy and selection criteria
We defined our research question and eligibility criteria using the Population, Exposure, Comparator, Outcome,
The population of interest consists of individuals of any age residing in Iraq at any time between 1990 and 2020 and of the children of those individuals. We did not limit our exposure inclusion criteria to strictly DU (U-235 content <0.7%), because it has been reported that slightly enriched uranium might also be used in conventional weapons by the US military.

Regarding outcomes, we included studies that reported mortality and morbidity outcomes such as cancer and birth defects. Only human observational studies were eligible for inclusion in this review. In order to overcome possible publication bias, we did not restrict our inclusion criteria based on publication status (eg, peer-reviewed studies).

We developed our search strategy with the assistance of a medical librarian and searched the following 11 electronic databases: MEDLINE, Embase, PubMed, Scopus, Toxline, Iraqi Academic Scientific Journals, WHO Digital Library, United Nations Environment Programme Knowledge Repository, IAEA Scientific and Technical Publications, Google Scholar and ProQuest Dissertations and Theses Global. Complete search strategies for each database are provided in the online supplemental material 1, table S2.

We restricted our initial search to human observational studies published between January 1990 and April 2020, and we used no language restrictions. In addition to database searches, we hand searched reference lists of included studies and forward-tracked citations of included studies in Google Scholar in order to identify additional relevant studies.

Title/abstract and full-text screening was conducted by independent screeners with two screeners per article (SS, MI and MA-L), using the reference management software EndNote V.X8 and following the prespecified eligibility criteria (see online supplemental material 1, ‘Screening form’). Calibration exercises were carried out prior to beginning title/abstract and full-text screening, in order to test the clarity of the screening protocol and to ensure a high level of agreement between the three primary reviewers (SS, MI and MA-L). During the calibration exercises, the reviewers met and reviewed the screening forms. Then, each reviewer independently screened the same subset of records and full texts, and the reviewers compared their results and discussed any disagreements.

Studies deemed eligible for inclusion by any reviewer during the title and abstract screening phase underwent full-text screening. At the full-text screening phase, reviewers again screened full texts independently and in duplicate. However, after screening at the full text phase, the two reviewers assigned to each study met to compare screening decisions. In cases of disagreement, the two reviewers discussed their justifications or invited a third reviewer for input (RRH and LAK) until consensus was achieved.

This study is reported in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (online supplemental material 1, Table S5). The study protocol is available online (PROSPERO, number CRD42018108225, see https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=108225).

There were no patient or public involvement in the study protocol of this SR.

Data analysis

Two reviewers abstracted data from full-text articles in duplicate and independently by using a prespecified data abstraction form designed by the authors in Microsoft Excel (for Office 365) (see online supplemental material 1, ‘Data abstraction items’). A data abstraction calibration exercise was conducted with a small sample of full texts prior to initiating data abstraction. After abstracting data independently, the two reviewers met to compare their results and resolve disagreements. Data abstraction items included summary measures from individual studies, such as difference in means and relative risk.

Risk of bias was assessed using the Navigation Guide’s risk of bias tool (provided in online supplemental material 1, ‘The Navigation Guide instructions for making risk of bias determinations’) that is adapted from the Cochrane Collaboration’s Risk of Bias tool and the Agency for Healthcare Research and Quality’s Methods Guide. The Navigation Guide’s risk of bias tool includes nine domains: recruitment strategy, blinding, confounding, exposure assessment, outcome assessment, incomplete outcome data, selective outcome reporting, conflict of interest and other sources of bias. For each domain, reviewers documented the risk of bias for each individual study as ‘low risk of bias’, ‘probably low risk of bias’, ‘probably high risk of bias’, ‘high risk of bias’ or ‘not applicable’ following prespecified criteria. Complete definitions of each domain and rating criteria are found in the online supplemental material 1 ‘The Navigation Guide instructions for making risk of bias determinations’.

Regarding the domain of confounding, we judged a study to be low risk of bias if it accounted for all five of the following confounders: age of participants, sex of participants, tobacco use, obesity and toxic environmental exposures. For studies that assessed birth defects, we required four additional confounders to be accounted for: maternal folate deficiency, maternal age (at birth), maternal education (at birth) and consanguinity. Two reviewers (SS and MI) rated risk of bias for each included study in duplicate and independently, and then met to compare their results. In cases of disagreement, input from a third reviewer (RRH) was sought.

We analysed data using a narrative synthesis, following guidelines produced by the Center for Reviews and Dissemination at the University of York. As per the Popay et al. guidelines, our synthesis sought to organise findings by health outcome, explore patterns in the results (including direction and magnitude of effects) and
examine relationships in the data. We organised our synthesis by eight health outcomes: CBDs, other birth-related outcomes, cancer, BCL-2 oncogene expression, PTEN gene expression, kidney failure, immune system function and other outcomes.

We assessed the certainty of evidence for each outcome using the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) approach. The GRADE approach defines the certainty of a body of evidence as the extent to which one can be confident that an estimate of effect is close to the true value for an outcome. The certainty of evidence is assessed for each outcome separately. According to the GRADE approach for assessing the certainty of the evidence for non-randomised studies (NRS), the rating of the certainty of evidence for NRS may be initially rated as high certainty. Next, the rating considers five factors for which the certainty of evidence is rated down, and three factors for which the certainty can be rated up. The five rating-down factors are: (1) risk of bias, (2) indirectness of evidence, (3) inconsistency, (4) imprecision of effect estimates and (5) risk of publication bias. The three rating-up factors are: (1) large effect, (2) dose–response gradient and (3) plausible confounders or other biases increase the certainty in the effect. We judged whether each of these factors (and for one outcome at a time) is of no serious, serious, or very serious concern, except for the risk of bias domain where ‘extremely serious’ may apply. The final assessment of the certainty of a body of evidence is one of four grades: high, moderate, low or very low (online supplemental material 1, table S4).

There was no funding source for this study.

RESULTS
Search results
Our database searches retrieved 2601 unique records. Of those, 235 met our inclusion criteria at the title and abstract screening phase (figure 1 – diagram style adapted from ROSES). The full-text reports for 15 of those records could not be retrieved. The records for all of the unretrievable texts came from our grey literature search and either (A) contained insufficient bibliographic data for tracking the source of the records (n=5) or (B) a request was made for the full text by a medical librarian, but the text was not supplied (n=10). Of the 220 reports that underwent full-text screening, 28 met our inclusion criteria. Reasons for exclusion at the full-text screening

---

**Figure 1** Study selection, diagram style adapted from ROSES.

- **Total records captured** (n = 3,831)
  - Records after duplicates removed (n = 2,601)
  - Records after title-abstract screening (n = 2,356)
  - Reports retrieved at full text (n = 220)
  - Reports after full text screening (n = 28)
- **Academic journal bibliographic databases** (n = 3,536)
  - Medline (n = 355)
  - EMBASE (n = 627)
  - PubMed (n = 317)
  - Scopus (n = 1,546)
  - Toxline (n = 462)
  - Iraqi academic and scientific Journals (n = 229)
- **Grey literature databases** (n = 355)
  - ProQuest dissertations and theses global (n = 52)
  - Google scholar (n = 122)
  - IAEA scientific and technical publications (n = 116)
  - WHO institutional repository for information sharing (n = 43)
  - UNEP knowledge repository (n = 22)
- **Excluded full texts, with reasons** (n = 192)
  - Exposure (n = 41)
  - Outcome (n = 47)
  - Population (n = 26)
  - Study Design (n = 78)
- **Pre-screened reports from other sources** (n = 5)
  - Citation lists of included studies (n = 4)
  - Tracking citations (n = 1)
stage are provided in the online supplemental material 1, ‘Full text exclusion justifications’. By screening the citation lists of included reports and by tracking their citations forward in Google Scholar, we identified five additional records that were not captured in our database searches, bringing our total number of included reports to 33. Of the 33 included reports, two reported results from the same study. Additionally, three articles reported the results of more than one relevant study, and thus the final number of studies included in our evidence synthesis was 36 (online supplemental material 1, tables S5–S12).

Characteristics of included studies
The timeframes of 26 included studies ranged from 1980 to 2017 (10 studies did not report their research timeframe). The most common study design was case–control, which was used in 61% of studies (n=22) (refs 6, 28 (study 2); refs 30, 42–44 (study 2); ref 44 (study 3); ref 45 (study 1); ref 45 (study 2); refs 46–59). Other study designs used were cross-sectional (14%, n=5) (refs 29, 44 (study 1); refs 60–62 ecological time trend (22%, n=8) (ref 28 (study 1); refs 63–69) and ecological geographic comparison (3%, n=1).

Publication dates of included reports ranged from 1999 to 2020, with the majority of articles (82%) published in the year 2010 or later (n=27). Eighteen studies (50%) focused on populations at the city or district level (Baghdad: n=6; Fallujah: n=7; Basrah city: n=3; Nasriyah: n=1; and Ramadi: n=1), 13 studies (36%) focused on populations at the governorate or province level (Babylon governorate: n=1; Baghdad governorate: n=1; Basrah governorate: n=7; Ninawa province: n=1; Karbala governorate: n=1; and Muthanna governorate: n=1) and 5 studies (14%) focused on populations at the regional level (Southern Iraq) (figure 2). One of the included studies was a national household survey. However, that study only used the subset of data collected from Basrah governorate for the analysis of the association between birth defects and warfare contamination.

Methods of exposure assessment
Eleven studies (30%) used place of residence, 9 (25%) used historical controls or time periods, 15 (42%) directly measured uranium concentrations in human biological samples and a single study (3%) used the source of bullets among patients with gunshot wounds to assess uranium exposure (online supplemental material 1, tables S5–S12). Most studies that assessed exposure by place of residence used prior knowledge of locations where military attacks involving the use of weaponised uranium took place, or cited literature that documented those locations. Only one included study conducted primary environmental monitoring to measure uranium concentrations in their study area.

Outcomes
Regarding health outcomes, 28% of included studies assessed CBDs (n=10) (online supplemental material 1, table S5). Other birth-related outcomes among included studies were: anthropometric measurements of newborns (3%, n=1), sex ratio at birth (3%, n=1) and infant mortality rate (IMR) (3%, n=1) (online supplemental material 1, tables S5–S12). By screening the citation lists of included reports and by tracking their citations forward in Google Scholar, we identified five additional records that were not captured in our database searches, bringing our total number of included reports to 33. Of the 33 included reports, two reported results from the same study. Additionally, three articles reported the results of more than one relevant study, and thus the final number of studies included in our evidence synthesis was 36 (online supplemental material 1, tables S5–S12).

Characteristics of included studies
The timeframes of 26 included studies ranged from 1980 to 2017 (10 studies did not report their research timeframe). The most common study design was case–control, which was used in 61% of studies (n=22) (refs 6, 28 (study 2); refs 30, 42–44 (study 2); ref 44 (study 3); ref 45 (study 1); ref 45 (study 2); refs 46–59). Other study designs used were cross-sectional (14%, n=5) (refs 29, 44 (study 1); refs 60–62 ecological time trend (22%, n=8) (ref 28 (study 1); refs 63–69) and ecological geographic comparison (3%, n=1).

Publication dates of included reports ranged from 1999 to 2020, with the majority of articles (82%) published in the year 2010 or later (n=27). Eighteen studies (50%) focused on populations at the city or district level (Baghdad: n=6; Fallujah: n=7; Basrah city: n=3; Nasriyah: n=1; and Ramadi: n=1), 13 studies (36%) focused on populations at the governorate or province level (Babylon governorate: n=1; Baghdad governorate: n=1; Basrah governorate: n=7; Ninawa province: n=1; Karbala governorate: n=1; and Muthanna governorate: n=1) and 5 studies (14%) focused on populations at the regional level (Southern Iraq) (figure 2). One of the included studies was a national household survey. However, that study only used the subset of data collected from Basrah governorate for the analysis of the association between birth defects and warfare contamination.

Methods of exposure assessment
Eleven studies (30%) used place of residence, 9 (25%) used historical controls or time periods, 15 (42%) directly measured uranium concentrations in human biological samples and a single study (3%) used the source of bullets among patients with gunshot wounds to assess uranium exposure (online supplemental material 1, tables S5–S12). Most studies that assessed exposure by place of residence used prior knowledge of locations where military attacks involving the use of weaponised uranium took place, or cited literature that documented those locations. Only one included study conducted primary environmental monitoring to measure uranium concentrations in their study area.

Outcomes
Regarding health outcomes, 28% of included studies assessed CBDs (n=10) (online supplemental material 1, table S5). Other birth-related outcomes among included studies were: anthropometric measurements of newborns (3%, n=1), sex ratio at birth (3%, n=1) and infant mortality rate (IMR) (3%, n=1) (online supplemental material 1, tables S5–S12).
material 1, table S6). Additionally, 15 studies assessed cancer (41%) (online supplemental material 1, table S7), and three studies (8%) measured outcomes related to immune system function (online supplemental material 1, table S8). A single study assessed kidney failure (3%) (online supplemental material 1, table S9), a single study assessed the expression of the oncogene BCL-2 (an important prognosis indicator for breast cancer) (3%) (online supplemental material 1, table S10) and a single study assessed PTEN gene expression (3%) (online supplemental material 1, table S11). Two case–control studies (5%) assessed multiple outcomes (they defined cases as patients with cancer, or parents of children born with CBDs, and then grouped both outcomes together in their analysis) (online supplemental material 1, table S12).

**Risk of bias assessment**

By assessing study quality using the Navigation Guide’s risk of bias tool, we concluded that there was generally a high risk of bias across the body of evidence (figure 3A). We found that confounding was the domain with the highest risk of bias, followed by other sources of bias, exposure assessment, blinding and recruitment (figure 3B). The domain with the lowest risk of bias was incomplete outcome data, followed by conflict of interest. Only two

![Figure 3](http://gh.bmj.com/)

Figure 3  Risk of bias assessments (low, probably low, probably high and high risk) for each included study (A) and given as percentages across all included studies (B). Justifications for risk of bias determinations for individual studies are provided in the online supplemental material 1, tables S13–S48.
Congenital birth defects

Ten studies selected CBD as their health outcome of interest. Of those, four studies examined CBD incidence rate, all of which found higher incidence in their exposed population compared with their control population. However, only two of the studies measuring incidence of CBD tested their observed differences for statistical significance. Alborz found that a significantly higher proportion of children with birth defects in Basrah were living in households that reported exposure to warfare contamination, than children without birth defects (p<0.001). Al-Sadoon et al found CBD incidence during the period 1995–1998 (4.57/1000 live births) to be significantly higher than during the period 1991–1994 (2.5/1000 live births) (p<0.01). Although no statistical test of significance was conducted, Al-Sabbak et al (study 1) observed that CBD incidence in 2003 in Basrah (23/1000 live births) was 17 times higher than the CBD incidence in 1994 (1.37/1000 live births). Six CBD studies compared uranium concentrations in biological samples collected from parents who had given birth to children with CBDs (cases) and healthy controls, of which five found mean uranium concentrations to be higher in samples collected from cases compared with controls.

Eight of the 10 studies that selected CBD as their outcome of interest were judged to have low or probably low risk of bias for the domain of conflict of interest. All 10 studies that selected CBD as their outcome of interest were judged to have high risk of bias for the domain of conflict of interest. All three studies on birth-related outcomes other than CBDs found uranium exposure to be associated with adverse health impacts. However, all studies were found to have high risk of bias in at least one domain.

Other birth-related outcomes

Al-Sahlane et al carried out a cross-sectional study at a hospital in Baghdad to measure association between uranium exposure and anthropometric measurements of infants. The study used a CR-39 fission track detector to measure uranium concentration in maternal and umbilical blood samples in 50 mother–infant volunteer pairs. They found that uranium concentrations in both maternal and umbilical cord blood samples were negatively correlated with measurements for body length, birth weight and head circumference. However, the isotopic ratio of uranium detected in maternal and umbilical cord blood samples was not examined and so the study cannot be said to strictly measure uranium derived from weapons used by the US or coalition forces. Although the study collected data on maternal age, it did not account for maternal age or other confounders in its tests for association.

Busby et al (study 1) sought to examine the association between birth sex ratio and radiation exposure in Fallujah using a cross-sectional household survey. They found that the sex ratio of boys to girls for children born during the period 2006–2010 was 0.86, compared with 1.182 for children born between 2001 and 2005, 1.109 for children born between 1996 and 2000 and 1.010 for children born between 1991 and 1995. The birth sex ratio for children born between 2006 and 2010 was found to differ significantly from the expected ratio (p<0.01). However, the study did not account for any confounders, including parental age or consanguinity. Additionally, the study did not explicitly discuss or report in their narrative text
<table>
<thead>
<tr>
<th>Outcome</th>
<th>No. of studies</th>
<th>Study design</th>
<th>Risk of bias</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Publication bias</th>
<th>Factors than increase the certainty</th>
<th>No. of patients and findings</th>
<th>Certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congenital birth defect (CBD)</td>
<td>10</td>
<td>Observational</td>
<td>Very serious*</td>
<td>Not serious</td>
<td>Serious†</td>
<td>Not serious</td>
<td>Serious‡</td>
<td>None</td>
<td>Although most studies are of small sample size, one study has a very large sample size (n=16,746), and its findings are both precise and consistent with the results of the remaining studies.</td>
<td>Very low</td>
</tr>
<tr>
<td>Cancer</td>
<td>15</td>
<td>Observational</td>
<td>Very serious§</td>
<td>Not serious</td>
<td>Serious¶</td>
<td>Not serious</td>
<td>Serious‡</td>
<td>None</td>
<td>The total number of participants included in all the studies was 997. Some studies reported small significant increase in cancer incidence among exposed populations, and other studies (mainly ecological studies) reported 'non-significant results'. The findings of one population-based study of 711 households were both precise and consistent with the results of the remaining studies.</td>
<td>Very low</td>
</tr>
<tr>
<td>BCL-2 oncogene</td>
<td>1</td>
<td>Observational</td>
<td>Very serious**</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Serious††</td>
<td>Serious‡</td>
<td>None</td>
<td>The total number of participants included in the study was 80 (low sample size). The study reported BCL-2 expression was found to be significantly higher among exposed populations.</td>
<td>Very low</td>
</tr>
<tr>
<td>Immune system function</td>
<td>3</td>
<td>Observational</td>
<td>Very serious‡‡</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Serious‡</td>
<td>None</td>
<td>The total number of participants included in all the studies was 365. All studies reported worse immune system function among exposed populations.</td>
<td>Very low</td>
</tr>
<tr>
<td>Kidney failure</td>
<td>1</td>
<td>Observational</td>
<td>Very serious §§</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Serious¶¶</td>
<td>Serious‡</td>
<td>None</td>
<td>The total number of participants included in the study was 26. The study found higher exposure among kidney failure patients than healthy controls.</td>
<td>Very low</td>
</tr>
</tbody>
</table>

Continued
Table 1

<table>
<thead>
<tr>
<th>Outcome</th>
<th>No. of studies</th>
<th>Study design</th>
<th>Risk of bias</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Publication bias</th>
<th>Factors than increase the certainty</th>
<th>No. of patients and findings</th>
<th>Certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTEN gene</td>
<td>1</td>
<td>Observational</td>
<td>Serious‡‡‡</td>
<td>Not serious</td>
<td>Serious†‡</td>
<td>Serious‡‡‡</td>
<td>Serious†</td>
<td>None</td>
<td>The total number of patients included in the study was 43. The study found that PTEN gene expression mean fold change was greater among the exposed group than the unexposed group.</td>
<td>Very low</td>
</tr>
</tbody>
</table>

*All 10 studies were judged to have high risk of bias related to confounding, because none accounted for any confounders known to be associated with birth defects, including consanguinity, maternal age, maternal nutritional status (ie, folate deficiency) or exposure to other environmental teratogens.

†We judged the evidence to have serious indirectness as the studies that assessed CBD incidence did not directly measure uranium exposure. For the case–control studies, which did measure uranium concentrations in human biological samples, it was not known whether the source of uranium exposure came from natural or artificial sources (ie, conventional weapons used by the US or coalition forces in Iraq). One study, carried out in Fallujah, measured the isotopic ratio of uranium in participants’ samples and found that the uranium had an enriched, rather than depleted, isotopic signature.

‡We suspect publication bias given that the sanctions played a role in limiting research publications on the health impacts of weaponised uranium in Iraq.

§All 15 studies were judged to have high risk of bias related to confounding. Five studies were rated high risk of bias in the domain ‘other sources of bias’. One study was potentially subject to recall bias or over-reporting.

¶We judged the evidence to have serious indirectness as the studies that assessed cancer incidence did not directly measure uranium exposure. For the case–control studies that did measure uranium concentrations in human biological samples, it was not known whether the source of uranium exposure came from natural or artificial sources (ie, conventional weapons used by the US or coalition forces in Iraq).

**The study was rated high risk of bias due to lack of adjusting for confounding and reporting of adequate blinding.

†††The total number of patients included in the study was almost 80 (low sample size). The study reported BCL-2 expression was found to be significantly higher among exposed populations.

‡‡‡All three of the studies were rated high risk of bias due to lack of adjusting for confounding and probably high risk of bias due to lack of reporting of adequate blinding.

§§The study was judged to have high risk of bias due to lack of adjusting for confounding and probably high risk of bias due to lack of reporting of adequate blinding.

***The study found higher exposure among kidney failure patients than healthy controls, but the sample size for the study was small (n=28).

†††The study did not directly measure uranium exposure.

‡‡‡The total number of patients included in the study was 43 (low sample size).
the birth sex ratio for age cohorts above 19 years. The next age cohort (20–24 years, children born between 1986 and 1990) has a birth sex ratio even lower than the 0–4 years age cohort (which represents children born between 2006 and 2010); the study data show a birth sex ratio of 776 males per 1000 females in the 20–24 years age cohort (although the ratio was not calculated or reported by the study authors) compared with 860 males per 1000 females in the 0–4 years age cohort (which was calculated and explicitly reported). Hence, the study was also judged to be high risk for the domain of selective reporting.

In the same household survey, Busby et al.44 (study 2) also investigated the IMR in Fallujah between 2006 and 2010.44 They found the IMR in Fallujah over that time period to be 80 per 1000 births, which was four times higher than the IMR reported in Egypt and Jordan (p<0.00001) and nine times higher than the IMR in Kuwait (p value not reported). However, the study did not account for any confounding variables and was rated high risk of bias in the domain of confounding.

We were unable to rate the certainty of evidence for studies of ‘other birth-related outcomes’, because the three studies in this category each assessed a different type of birth-related outcome (anthropometric measurements, birth sex ratio and infant mortality).

Cancer
Fifteen studies selected cancer (specific types of cancer or all cancer) as their health outcome of interest, of which seven assessed cancer incidence. Two of the seven studies that assessed cancer incidence found a higher incidence among their exposed populations compared with their controls. In contrast, all seven of the case–control studies that compared uranium concentrations in biological samples collected from patients with cancer (cases) and healthy controls found mean uranium concentrations to be higher among cases, but only two of the four studies tested the difference in means for statistical significance. Like the CBD case–control studies, the cancer case–control studies that measured uranium concentrations in participants’ samples did not assess the source of uranium (natural or artificial).

All 15 studies that selected cancer as their outcome of interest were rated low or probably low risk for conflict of interest. However, all 15 were rated high risk of bias in the domain of confounding. Additionally, 5 of the 15 studies were rated high risk for the domain ‘other sources of bias’.47–49, 58–60 Among the ‘other sources of bias’, two studies did not report their research timeframe (or the year that blood samples obtained from hospitals were collected from cancer patients),47 48 one study examined cancer case counts (as opposed to cancer incidence)69 and two studies had very small sample sizes (n=5 and n=12).49, 58

We judged the certainty of the evidence for the association between uranium exposure and cancer among the Iraqi population to be very low mainly due to the extremely serious risk of bias and serious publication bias (table 1).

Immune system function
All three of the studies that selected outcomes related to immune system function reported finding an association between uranium exposure and adverse health outcomes. Abdul-Wahid46 found that individuals with higher levels of uranium exposure had lower levels of selected lymphocytes compared with individuals with lower levels of exposure (no p values reported from statistical test).46 Humadi and Khalaf52 investigated several outcomes related to immune system function among patients who had been injured by bullets fired by US and coalition forces (a proxy for DU exposure) compared with patients injured by bullets from other sources (controls).52 They found that DU exposure was associated with adverse health outcomes including lower blood haemoglobin concentrations and faster erythrocyte sedimentation rates. The third immune system study found that higher exposure to uranium was associated with higher mitotic index scores.62 However, they did not test the difference for statistical significance.

Although all three immune system studies were consistent in their findings of direction of effect, they were all rated probably high risk of bias in the domain of blinding. For all three studies, blinding of key personnel for either exposure or outcome measurement was not reported, and lack of blinding could have introduced bias. All were rated high risk or probably high risk in the domain of confounding. All three studies were rated low or probably low risk of bias in the domain of conflict of interest.

We judged the certainty of the evidence for the association between uranium exposure and adverse impacts on immune system function among the Iraqi population to be very low mainly due to the extremely serious risk of bias in the domains of blinding and confounding, and serious publication bias (table 1).

Kidney failure
Our search strategy captured two reports of the same case–control study that assessed uranium concentration in blood and urine samples from kidney failure patients (cases) and healthy controls in a hospital in Al-Muthanna governorate.32 49 The study found that uranium concentrations were higher in both urine and blood samples of kidney failure patients than healthy controls, but the differences were not tested for statistical significance. The study was rated probably high risk of bias in the domain of blinding, because no procedures accounting for blinding were reported. It was also rated high risk of bias in the domain of confounding. Although the study collected data on age and sex of participants, these confounders were not controlled for, and neither were any other confounders.

We judged the certainty of the evidence for the association between uranium exposure and kidney failure in patients in the Iraqi population to be very low mainly
due to the extremely serious risk of bias in the domains of blinding and confounding, serious imprecision and serious publication bias (table 1).

**BCL-2 oncogene expression**

One case–control study examined the expression and intensity of the oncogene BCL-2 (an important prognostic factor for breast cancer) in Iraqi patients with breast cancer (n=50) and compared findings with a control population in Italy (n=30). The study found that BCL-2 expression in Iraqi breast cancer tissue samples was significantly higher (p=0.037) than in Italian samples, and among individuals for which BCL-2 was positively expressed, Iraqi participants had higher intensities than Italian participants.

However, the study was rated high risk of bias in the domain of confounding and probably high risk in the domain of blinding. Additionally, the study was rated probably high risk of bias in the domain of outcome assessment, because the authors reported that differences in tissue processing time, tissue quality and timing for embedding tissues in paraffin between cases and controls may have impacted outcome measurements, but data are not available for those variables.

We judged the certainty of the evidence for the association between uranium exposure and BCL-2 expression among the Iraqi population to be very low mainly due to the extremely serious risk of bias in the domains of blinding and confounding, serious imprecision and serious publication bias (table 1).

**PTEN gene expression**

A single study measured the expression of the tumour-suppressor gene, PTEN, in endometrial carcinoma patients living in DU-exposed areas (cases) and unexposed areas (controls) in southern Iraq. The study found that PTEN gene expression mean fold change was greater among the exposed group (0.19±0.185) than the unexposed group (0.003±0.0029), but the difference in means was not statistically significant when tumour grade and cancer stage were controlled for (p=0.286 and p=0.98, respectively).

The study did not conduct environmental monitoring to determine uranium exposure in the study area, nor directly measure uranium concentration in biological samples from study participants. Hence, the study was rated probably high risk of bias in the domain of exposure assessment. The study was also rated probably high risk of bias in the domain of confounding. While the study controlled for the age and sex of participants, it did not control for any other important confounders, such as obesity and tobacco use.

We judged the certainty of the evidence for the association between uranium exposure and PTEN gene expression among the Iraqi population to be very low mainly due to the extremely serious risk of bias in the domains of exposure assessment and confounding, serious imprecision and serious publication bias (table 1).

**Other outcomes**

Two included studies (with the same set of participants, published in the same article) used case–control study designs and selected cases based on multiple outcomes—they defined cases as patients with cancer, or parents of children born with CBDS—in Baghdad.45

The stated objective of the study by Al-Hamadany et al45 (study 1) was to compare uranium concentration in blood samples collected from cases and (healthy) controls.45 The study recruited participants from Baghdad in five categories: patients with cancer prior to treatment (n=15), patients with cancer currently receiving treatment (n=15), women who had given birth to children with CBDS (n=15), volunteers who were occupationally exposed to ionising radiation (n=14), residents of areas of Baghdad that were reportedly contaminated with DU (n=15) and healthy volunteers residing in areas of Baghdad not reported to be contaminated with DU (n=14). The study reported the mean uranium concentration measured in blood samples collected from the healthy unexposed group (n=14) and the mean uranium concentration in blood samples from participants from all other groups (n=74).

The study found that mean uranium concentration from the control population (0.11 ppm ±0.009 SE) was significantly lower than the mean uranium concentration in samples from participants from all other recruitment groups (0.21 ppm±0.01 SE) (p<0.05).

The second study by Al-Hamadany et al45 (study 2) assessed outcomes related to immune system function but used the health condition of patients as an indicator for the exposure to uranium.45 Patients with cancer, and mothers of children with birth defects, were defined as exposed, as were individuals occupationally exposed to ionising radiation in a medical setting and individuals residing in areas of Baghdad that were reportedly contaminated with DU. Controls were defined as healthy individuals living in uncontaminated areas of Baghdad. The study found that white cell counts were significantly higher among participants defined as exposed, and haemoglobin concentrations were significantly lower, compared with the group defined as unexposed (p<0.05).

The definition of cases (exposed) and controls (unexposed) in the studies by Al-Hamadany et al45 was problematic and undermined the quality of the study results. We were unable to rate the certainty of evidence for the association between uranium exposure and other outcomes among the Iraqi population, because the definition of cases in these two studies included individuals with a variety of health conditions and a number of exposure scenarios (including environmental uranium exposure and occupational exposure to ionising radiation).

**GRADE summary of findings**

The certainty of the evidence was rated very low across all outcomes (table 1).
DISCUSSION

The investigation of a possible causal link between DU exposure and adverse health outcomes is essential for designing interventions. In this SR, we sought to synthesise the evidence on the association between uranium exposure and adverse health outcomes in Iraq. To our knowledge, this is the first SR on the topic.

Among the 36 studies that met inclusion criteria for this SR, the most common study design was case-control. This study design may have been selected due to the fact that the outcomes of interest among our included studies were either illnesses with long latency periods (such as cancer) or CBDs, which were identified in hospital settings. The most common method of exposure assessment among our included studies was direct measurement of uranium in human biological samples, and the most common technique used was CR-39 fission track detectors. This is probably explained by the fact that uranium is an alpha particle emitter and use of fission track detectors was the most feasible technique for measuring alpha particles.

The findings of our included studies are relatively consistent in the direction of effect; most of the included studies (n=30, 83%) reported a positive association between uranium exposure and adverse health outcomes, including CBDs and cancer. Although the data suggests that there is an association between uranium exposure and adverse health outcomes, the evidence is very uncertain. The very low certainty of the evidence across outcomes is mainly due to the extremely serious risk of bias (especially in the domain of confounding), serious indirectness and imprecision and the very serious publication bias. Until now, the mantra in the literature has been that not enough studies have been conducted to assess whether DU has adversely impacted the health of the Iraqi population.\textsuperscript{22,73} The findings from this SR allow us to make a new argument: the body of evidence does not lack studies, it lacks high quality studies.

Publication bias and missing data

Study publication bias is a well-documented problem in medical research and health sciences.\textsuperscript{74} Typically, studies that find significant results are more likely to be published, leading to bias in the literature, but the problem can also arise when other factors influence the production and publication of research.\textsuperscript{74} By examining the characteristics of our included studies in a historical context, we observed patterns that suggest that publication bias may have occurred. The sanctions imposed on Iraq by the UN Security Council in 1990 after the Iraqi invasion of Kuwait (specifically resolution 661) restricted the flow of physical goods and financial resources into and out of Iraq and effectively acted as an intellectual embargo, incapacitating Iraq’s medical research abilities.\textsuperscript{75,76} The sanctions were modified after the 2003 US invasion of Iraq, and then lifted in 2010. Even though the investigation of health impacts of DU has been a priority for Iraqi scientists since the early 1990s, we found that the majority of the studies included in this SR (82%) were published in 2010 or later, after the sanctions had been lifted. The sanctions may have limited access by Iraqi researchers to specialised equipment needed for detecting and measuring uranium concentration in environmental and human biological samples. All of the included studies in this SR that used specialised equipment, such as inductively coupled plasma mass spectrometry (ICP-MS), or CR-39 fission track detectors, were published in 2011 or later. This suggests that the sanctions played a role in limiting research and publication on the health impacts of weaponised uranium in Iraq. Additionally, the difficulties inflicted by the 20-year intellectual embargo may have impacted the quality of research that has been produced on this topic. A sanction-specific publication bias may have prevented the publication of high-quality studies that found positive results on the association between exposure to DU and adverse public health burdens in Iraq or studies that were authored by Iraqi nationals.\textsuperscript{76,77}

The lack of public access to important data may have also inhibited high-quality research on the public health impacts of DU exposure in Iraq. One source of missing data is the complete set of US firing coordinates for DU weapons in Iraq. Over the last decade, multiple nongovernmental organizations (NGOs) in the USA and abroad have filed Freedom of Information Act requests with the US Department of Defense and Department of State, but only a fraction of the requested data has been publicly released.\textsuperscript{18} Another source of missing data is the IMOH national CBDs study that was carried out in 2012. Summary results from the study were published in 2013, but to date neither a full report has been published, nor has the study data been made accessible to independent researchers for analysis and verification.\textsuperscript{18}

Strengths and limitations

Many previous reviews have attempted to summarise the body of evidence regarding the effects of DU on public health in Iraq. However, the present study is the first to do so systematically. One of the strengths of the SR methodology is the sensitivity and transparency of the search strategy. To our knowledge, this study assembled in a single collection more human observational studies investigating associations between exposure to weaponised uranium and health outcomes among the Iraqi population than any publication has before. Additionally, we assessed the quality of individual studies and rated the certainty of the body of evidence in a transparent and reproducible manner. Even though the risk of bias within our body of evidence was generally high, we found a low risk of bias in the conflict of interest domain. This is a notable strength in our included studies and speaks to both the importance of this topic in Iraqi society and the tenacity of Iraqi academic researchers, who pursued this topic despite technical and political challenges.

A possible limitation of this SR is that in our inclusion criteria and PECOS statement, we did not account for ‘length of residence’ among our study population. If we had applied an additional participant inclusion criteria
regarding length of residence, such as living in the study location continuously for at least 5 years, it may have reduced our total number of included studies. Another limitation of this SR is that we were only able to judge the certainty of evidence using the GRADE approach for six of the eight outcome categories. Specifically, we were unable to judge the certainty of the evidence for other birth-related outcomes and other (mixed) outcomes, because studies assessed a variety of outcomes within those outcome categories.

**Meta-analysis**
We did not carry out a meta-analysis in this review. The lack of similarity in study outcomes and study designs precluded our ability to conduct a meta-analysis of summary measures among the included studies.

**Implications for future research**
The sparsity of high-quality observational studies on the health impacts of DU among the Iraqi population led to very low certainty in the quality of the evidence. This, however, does not negate the possibility that DU weapons have adversely affected public health in Iraq. A number of experts have expressed concern that the politicisation of DU research might have limited international collaborations and interdisciplinary approaches to the topic, isolating Iraqi scientists over the past decades. Similarly, we suspect that efforts to assess, evaluate and clean-up DU might have been undermined by lack of security, financing or attention to environmental health impacts of war. To date, research on the impact of war in Iraq has mainly focused on body counts and infrastructure damage, and critical data, like the complete set of US DU firing coordinates or data from the IMOH national birth defect survey, are unavailable. Consequently, the ability of Iraqi scientists to conduct research on DU has been limited. More research on this topic is warranted, specifically observational studies with improved methods of exposure assessment, adequate blinding of key personnel and assessment and control of important confounders.

**Improved methods of exposure assessment**
The most common method of exposure assessment among the included studies was the measurement of uranium concentration in human biological samples, using equipment such as ICP-MS and CR-39 fission track detectors. The direct measurement methods used in such biomarker studies are typically reliable and robust. However, only one of the included studies measured the isotopic ratio of uranium in samples, using ICP-MS, a step that is critical to determining the source and potential toxicity of the uranium. Hence, there is a need for more observational studies that assess uranium exposure via the use of biomarkers and that determine the isotopic ratio of detected uranium using methods such as ICP-MS. Additionally, only one of the included studies used environmental monitoring methods to assess uranium exposure. Future observational studies should use environmental monitoring to assess the presence, concentration and isotopic ratio of uranium in water, soil, dust and other environmental media. Recent reports that the US military has used DU weapons in Syria highlight the need for future primary studies to be carried out in countries outside of Iraq as well.

**More SRs and health risk assessments**
In addition to high-quality primary research, we call for SRs on related topics to be conducted. This SR does not represent an assessment of studies that measured uranium or radiological contamination in Iraq without also assessing health outcomes. Although studies that estimated future health risks based on absorbed dose or effective dose calculated from environmental radioactivity measurements in Iraq (as opposed to reporting current health outcomes among participants or target populations) did not meet our inclusion criteria, they still offer insight into possible health hazards faced by the Iraqi population, for example, see Almayahi and Mohammed and Ahmed. An SR of these and similar studies is merited. Furthermore, no SRs on evidence of uranium toxicity from animal studies, or on DU-exposed populations outside of Iraq (such as the Balkans), have been conducted. A large body of primary studies exist for these topics, but they have yet to be systematically synthesised and critically appraised.

In conclusion, more research on DU and other war remnants is imperative for understanding the long-term, environmental health impacts of armed conflict. DU weapons are still in use in other countries, and therefore the study of the health impacts of DU remains essential. We believe that funding for continued research should be provided in tandem with funding for clean-up and personnel training for the removal of debris from known DU-contaminated sites in Iraq. The global health community has been scrutinised for its inadequate response to the 2003 Iraq War. International health organisations (including the WHO) had not been held accountable for their work during and after the war, and unnecessary suffering occurred because of that failure. We are concerned that the global health community may have also failed to adequately support research into the health impacts of DU in Iraq. There is a need for an international regulation to allow and protect research conducted on this sensitive topic going forward. We call on garnering funding, training and collaboration to Iraqi researchers who continue to investigate this unforgotten topic.

**Acknowledgements** We would like to thank the staff at Saab Medical Library (particularly Carla Chalhoub) for their assistance in retrieving full texts. We would also like to thank the following people for their assistance in translation of non-English articles: Shaun Nesheim (Russian), Janina Shirin (French, Spanish and German), Hajar Assaad (French), Asila Atinisik (Turkish), Hady Al-Shamy (Arabic and French) and Dunya Jalloul (Arabic and French). We would like to thank Tamara Lotfi at the Global Evidence Synthesis Initiative for her feedback and support.

**Contributors** SS and RRH conceptualised the study. SS, RRH, LAK, OD and EAA developed the protocol. SS and AF developed the search strategy. SS did the literature search. SS, MI, MA-L, LAK and RRH screened titles, abstracts and full texts. SS and MI abstracted data and assessed risk of bias. LAK and EAA rated the certainty of evidence for each outcome. SS, MI, LAK and RRH interpreted the data.
SS and RRH wrote the first draft. SS, RRH, LAK, and EAA revised the subsequent versions of the manuscript. All authors approved the final manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Map disclaimer The depiction of boundaries on this map does not imply the expression of any opinion whatsoever on the part of BMJ or any member of its group concerning the legal status of any country, territory, jurisdiction or area or of its authorities. This map is provided without any warranty of any kind, either express or implied.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information. Data relevant to this manuscript are included in the supplemental material: these include the full search strategies, the full-text exclusion justification and the risk of bias assessment for each included study.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error or/and omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

Twitter @FHS_AUB

ORCID iD

Rima R Habib http://orcid.org/0000-0001-6280-7238

REFERENCES
