SYSTEMATIC REVIEW OF ADVERSE REPRODUCTIVE HEALTH OUTCOMES IN GULF WAR VETERANS

ABSTRACT

We aimed to identify and summarise the findings from studies that have compared the prevalence of adverse reproductive health outcomes in Gulf War veterans to that in a non-Gulf comparison group. Studies published between January 1990 and May 2004 were identified by searching a large number of electronic databases. Reference lists and websites were also searched and key researchers were contacted. Studies that reported the prevalence of any adverse reproductive health outcome both in Gulf War veterans and in a comparison group of non-Gulf veterans were included in the review. 2,401 abstracts were independently reviewed by two authors. Thirteen primary studies that measured cases of spontaneous abortion, stillbirth, birth defects and infant death within the first year after birth fulfilled the inclusion criteria. A meta-analysis of four studies demonstrated a weak but statistically significant association between male veterans' deployment to the Gulf and spontaneous abortion of pregnancies following the war (OR 1.36, 95%CI 1.18-1.57). A meta-analysis of seven studies of male veterans that defined outcome as any birth defect demonstrated statistically significant heterogeneity between studies ($\chi^2 = 74.2$; 6 degrees of freedom; $p < 0.001$). There was little evidence to support the hypothesis that Gulf War veterans are at increased risk of a general range of birth defects being diagnosed in their offspring. However, the current literature cannot exclude a possible increase or reduction in risk associated with Gulf War deployment as there is still considerable uncertainty around the risk estimates, particularly for rare outcomes. Furthermore, given the relatively small number of female veterans included in studies it is difficult to comment accurately on any potential risks in this specific population with any certainty.

INTRODUCTION

Ever since the end of the Gulf War in 1991, veterans of the war have reported concern over a range of health complaints including the possible risk of adverse reproductive health outcomes in their offspring. Anecdotal media reports have suggested an excess of birth defects amongst children of Gulf War veterans.[1] Spontaneous abortion and infertility have also been highlighted as areas of concern.[2]

In 1994 a report by the US General Accounting Office identified 21 reproductive toxicants or teratogens which were present during the period of the Gulf War.[3] Some veterans have also been concerned by their exposure to a combination of multiple agents simultaneously, including pyridostigmine bromide (NAPS tablets) along with a range of vaccinations and other medications. It is possible that some servicewomen could have been exposed to these teratogens during their pregnancy. However, since a far greater proportion of Gulf War veterans were male, a paternal model of exposure might be more relevant whereby the toxic exposures might act either before conception or during pregnancy via affected sperm.

Problems can arise when trying to measure adverse reproductive health outcomes. Measurements of infertility and time taken to conceive tend to rely on self-report which is prone both to random error and more problematic recall bias. Diagnoses of a birth defect such as cleft palate or Down syndrome might be easier to obtain from birth records and hospital data, but these often only contain diagnoses made at or near the time of birth and might underestimate those that take longer to detect or diagnose.

This paper aims to identify and summarise the findings from studies that have assessed adverse reproductive health outcomes in Gulf War veterans together with those in a group who were not deployed to the Gulf (non-Gulf veterans). Our definition of an adverse reproductive health outcome included spontaneous abortion and stillbirths, birth defects, and infant death within the first year after birth.

MATERIALS AND METHODS

Searching

The methods employed in the systematic review have been described in detail in another paper[4] and are summarised here. 5387 studies from the period January 1990 to May 2001 were identified for possible inclusion by searching through databases (EMBASE, Medline, ASSIA, SIGLE, PsycINFO, CancerLit, HealthSTAR, Dissertation Abstracts, Current Contents, Health and Psychosocial Instruments, CINAHL and Biological Abstracts) and websites and by contacting researchers in the field. Studies were eligible for inclusion if they contained data on non-indigenous Gulf War veterans together with an appropriate comparison group which differed in its level of exposure. Those veterans who were not deployed to the Gulf may or may not have been deployed elsewhere on active duty. Therefore rather than referring to them as non-deployed we shall refer to them as non-Gulf veterans throughout this review. Abstracts of 2296 references that remained eligible were examined by two members of the research team.

Studies were then excluded if they measured simulated exposures, if they measured non-health related outcomes or if the subjects were inhabitants of the Persian Gulf rather than deployed military personnel. Studies that examined adverse reproductive outcomes within groups of Gulf veterans that had experienced differential exposures whilst in the Gulf, e.g.

exposure to depleted uranium, were also excluded from this review. All included studies were categorised by health outcome, one of which was adverse reproductive outcomes. Eight studies were identified by this search.

An updated electronic search of the literature from January 2001 to May 2004 was completed which identified a further 538 references. Of databases searched in 2001, CancerLit and HealthStar were now incorporated into Medline whilst Dissertation Abstracts and Health and Psychosocial Instruments were no longer available. Two databases not searched in 2001, the Web of Knowledge Databases and the Science and Social Science Citation Indexes were included in this updated search. Of 538 potentially relevant references, 105 were selected from the abstract (or title if no abstract) as potential research studies with a relevant comparison group. From these, five papers were identified and therefore this review contains thirteen papers that fulfilled our inclusion criteria and which contained data relating to adverse reproductive outcomes both in Gulf War veterans and non-Gulf veterans.

Data extraction

Data relating to the studies' main hypotheses and to methodological quality were extracted independently by two members of the research team onto pre-designed data extraction forms. Information on methodological quality of the individual studies included the statistical power, the potential of selection bias in the sampling of subjects and in the response rate, the potential bias in the measurement of outcomes, and the availability of data on confounders and the controlling for such variables in each study.

Statistical analysis

Meta-analysis statistically combines and analyses data from separate studies with the aim of appraising the evidence objectively, providing a more precise estimate of effect and exploring any heterogeneity between the results of individual studies.[5] A summary odds ratio was calculated with a random effects model using the DerSimonian and Laird method.[6] The estimate of heterogeneity between studies was taken from the inverse variance fixed effect model. All analyses were performed using the "metan" command[7] in Stata Version 9 (Stata Corporation, College Station, TX, USA). We chose to use this approach because of our a priori view that the studies were inherently heterogeneous.

RESULTS

We identified thirteen primary studies which compared the prevalence of adverse reproductive outcomes in both Gulf War veterans and an unexposed comparison group.[8-20] One study used US birth surveillance data as a comparison, all other studies utilised samples of non-Gulf military veterans as a comparison group,.

Table 1 summarises the results and methodology of the two studies that defined an adverse reproductive outcome as adverse semen characteristics.[8-9] Neither Ishoy et al's[8] sample of Danish males nor Maconochie et al's[9] sample of British males exhibited strong evidence for an increased risk of abnormal semen characteristics amongst veterans of the Gulf War.

Table 2 summarises similar information for six studies that investigated the association between Gulf deployment and spontaneous abortion or stillbirth.[8,10-14] The earliest study used a historical comparison which would be affected by population mobility and therefore is more difficult to compare directly with the other studies.[10] Ishoy et al did not provide the total number of pregnancies as a denominator, therefore the quoted prevalence of spontaneous abortions as a proportion of all veterans is likely to be falsely low.[8]

The data from four studies relating to prevalence of spontaneous abortion amongst partners of male veterans was summarised in a meta-analysis (Figure 1).[8,11,12,14] Each individual study reported an association between deployment to the Gulf and an increased odds of spontaneous abortion. The overall pooled estimate gave an OR of 1.36 (95%CI 1.18-1.57). This overall estimate was weighted heavily (48%) by the results of Doyle et al.[14] The degree of statistical heterogeneity between these four studies was non-significant.

Data relating to spontaneous abortion amongst female veterans was provided by three studies, each one involving a relatively small sample of female veterans and therefore a small number of cases.[11,13,14] The sample of Kang et al and Doyle et al did not demonstrate a strong association between female deployment and spontaneous abortion after the war. Araneta et al also reported a relatively weak association between female deployment and spontaneous abortion of pregnancies conceived during the war, but a far stronger association with spontaneous abortion of pregnancies conceived after the war (OR 2.92, 95%CI 1.9-4.6).

All studies that measured prevalence of stillbirth, whether amongst pregnancies of female veterans or amongst pregnancies of partners of male veterans, only reported very small numbers of cases. Overall, these sparse data do not support the hypothesis of an association between Gulf deployment and increased risk of stillbirth in pregnancies.

Table 3 presents data for the ten studies that investigated the association between Gulf deployment and birth defects or infant deaths amongst veterans' offspring.[8,11,12,14,15-20] The majority of these defined outcome as being any birth defect or congenital anomaly (or one out of a range of specific defects). Araneta et al (1997) investigated prevalence of a specific rare syndrome (Goldenhar syndrome).[16] Six of the publications related to American Gulf War veterans, the remaining four publications included data on Canadian, Danish, Australian and British Gulf War veterans respectively. The majority of studies relied on self reported measurements of birth defects, either without independent ascertainment or with relatively limited clinical confirmation.

Figure 2 illustrates the association between deployment to the Gulf War and prevalence of any category of birth defect or congenital anomaly amongst offspring of male veterans. Penman et al's data which compared Gulf War veteran data with a national US comparison were not included in the meta-analysis.[15] Araneta et al's Hawaiian data from 2000 were also not included to avoid duplication, since they also appear in the author's larger study of 2003.[20] The unspecific nature of the outcome measured across all studies might largely explain the statistically significant heterogeneity between these seven studies ($\chi^2 = 74.2$; 6 degrees of freedom; $p < 0.001$). Figure 1 is presented for illustrative purposes only; the summary estimate should be interpreted with great caution due to the significant heterogeneity between studies.

Although these studies have not demonstrated a strong and consistent effect of Gulf Deployment on generalised risk of birth defects, an association might exist with a single specific birth defect. Araneta et al did report an approximate three fold increase in risk of Goldenhar syndrome in offspring of Gulf War veterans compared to non-Gulf veterans, but these data are based on a total of seven cases of the syndrome.[16] Araneta et al (2003) also demonstrated a significantly increased risk of specific birth defects such as aortic valve stenosis and tricuspid valve insufficiency amongst offspring of male veterans.[20] Doyle et al reported a significantly increased risk of specific birth defects of the urinary and musculo-skeletal systems in the offspring of male veterans.[14]

None of the three studies which investigated infant death reported a significant association with Gulf deployment.[8, 11,12]

Instead of quoting relative risks, the absolute risk difference between Gulf War veterans and non-Gulf veterans might be particularly informative when the reproductive outcomes under investigation are relatively rare. For example, Araneta et al estimated that equivalent to nine extra infants with Goldenhar syndrome might arise amongst every 100,000 births to Gulf War veterans compared to 100,000 births amongst non-Gulf veterans.[16] However the 95 percent CI indicates the results are consistent with anything from five extra diagnoses amongst offspring of non-Gulf veterans to twenty four extra diagnoses amongst offspring of Gulf War veterans. Since 34,069 infants were actually born to the total cohort of American Gulf War veterans, the risk difference could range from two extra cases of Goldenhar syndrome amongst non-Gulf veterans to eight extra cases amongst Gulf War veterans.

Since toxic exposures might potentially act as risk factors for adverse reproductive outcomes differently amongst men and women, we utilised data stratified by sex of military parent where possible. Although the calculations were based on far smaller numbers of female veterans than male veterans , each study that reported the odds of any birth defect in the offspring of Gulf War veterans separately by gender showed a greater risk amongst female veterans. However, any differences in risk between male and female veterans were not statistically significant and none of these results provide evidence for an interaction between sex of deployed parent and the odds of birth defects following Gulf deployment.

DISCUSSION

This systematic review shows some evidence for a weak association between male veterans' deployment to the Gulf War of 1991and increased risk of spontaneous abortion during pregnancy amongst their partners following the war. The data included in the review do not support the hypothesis of increased risk of stillbirth or generalised categories of birth defects amongst offspring of male Gulf War veterans. However, the possibility of an association between deployment to the Gulf and risk of individual specific birth defects cannot be ruled out. Finally, the review does not support the hypothesis that Gulf War veterans exhibited any adverse semen characteristics after their deployment.

No strong evidence for an increased risk of any of the main adverse reproductive outcomes of this review was reported for female veterans. However, given the relatively small number of female veterans in the individual studies and the further reduced subsample of those reporting any adverse reproductive outcomes, it is difficult to comment accurately on any potential risks in this cohort with any certainty.

Limitations of studies investigating birth defects amongst Gulf War veterans' offspring

Sample size.

All of the studies would suffer from low statistical power if investigating the risk of specific rare birth defects. For example, even though Araneta et al utilised a large cohort only a very small number of cases of Goldenhar syndrome had arisen.[16] Defining outcome to include a range of birth defects will increase the number of outcome events, but complicates the inference of any underlying possible biological mechanisms. For example, Penman et al[15] noted that the birth defects that were observed in their study had no known genetic or chromosomal abnormality or teratogen in common.

Sample selection and response bias.
The study by Penman et al suffered by lacking a military comparison group and from a sample selection driven by local media interest in a possible clustering of children's health problems.[15] However the study benefited from a 90 percent response rate. Ishoy et al did not provide clear details of the number of eligible Danish Gulf War veterans or the sample selection, therefore it is difficult to estimate the possible impact of response bias. [8]

Both Cowan et al's study and Araneta et al's study of Goldenhar syndrome used the total cohort of American Gulf War veterans and approximately 50 percent of the service members who had not been deployed to the Gulf as a comparison group, leading to minimal selection bias.[16,17] However, the sample was limited to children born only within military hospitals which was estimated to include approximately 68 percent of all births to active duty military personnel during the study period.[16] This criterion also effectively excluded all children born after their parent had left active duty and all children born to reserve personnel (more than 100,000 individuals). Importantly, the proportion of Gulf War veterans and non-Gulf veterans still on active duty on Sept 30 1993 was similar (55.6 percent male GWV versus 57.4 percent male NGV; 56.8 percent female GWV versus 56.2 percent female NGV),[12] and these slight differences in follow-up are probably not major sources of bias.

The main strength of the Canadian study was its use of the complete cohort of Gulf War veterans and a sample of non-deployed personnel selected from the Department of National Defence's human resources files.[18] However, the response rate amongst Gulf War veterans was much higher than amongst non-Gulf veterans (73 percent versus 60 percent) and it is likely that this could introduce quite serious bias into the quoted results.

Araneta et al's record linkage studies benefited from a robust sample selection by linking birth certificate data for all livebirths in a particular state with data from the Birth Defects Program and the Defense Manpower Data Center.[19,20] This provided information for infants whose parents were separated from the military, were reservists or members of the National Guard, and for infants born in non-military hospitals. Theoretically 100 percent of livebirths to Gulf War veterans and non-Gulf veterans should have been identified thus abolishing selection and response bias. These authors found no evidence of an association between Gulf War deployment and risk of birth defects.

Although not using the total cohort of American Gulf War veterans, the study by Kang et al did utilise a random sampling frame taken from the Department of Defense Manpower Data Center which was stratified by gender and military unit component.[11] Veterans were included irrespective of whether they remained in the military after the Gulf War and irrespective of their active duty, reserve or National Guard status. The study suffered from differential response rate between Gulf and non-Gulf veterans as previously described in the Canadian study. However, the authors did not believe self-selection of those with an adverse pregnancy outcome to be a major source of bias since the reported rate of birth defects in male Gulf veterans' offspring was actually lowest in the first wave of participant response (8.2 percent), increasing by the fourth wave (10.6 percent). Similarly Doyle et al reported that failure to respond in their study was largely unrelated to reproductive factors.[14]

It is possible that a bias can arise in cohort studies if the exposed and unexposed groups are not comparable at baseline. It seems hard to imagine how such a bias might arise when investigating adverse reproductive outcomes, since it would require that the selection of soldiers for deployment would be influenced by factors related to future reproductive health.

Measurement of birth defects.

The record linkage studies by Araneta et al benefited from a study design which probably provided the most accurate measurement of birth defects.[19,20] Their data linkage technique not only eliminated the possibility of observer bias, but information related to diagnoses throughout the first year of life could be included for all infants including those born in non-military hospitals. Araneta noted that including defects recognised at the time of birth only might underestimate the number of cases of a particular diagnosis, especially a diagnosis that is difficult to define. Only including births in military hospitals might further underestimate the prevalence of diagnoses since the high risk pregnancies, even amongst active duty personnel, might have been referred to a non-military hospital.

However the studies were limited because data were only available for livebirths; data on stillbirths and aborted fetuses (spontaneous and provoked terminations) were unavailable. Furthermore the authors noted that reliance on diagnostic data from birth defects surveillance records might often lack information on severity of a condition or on etiological factors.

The majority of studies relied on the veterans' self report of birth defects amongst their children without any ascertainment from birth records.[8,11-14,18-20] Such measurement could have introduced recall bias whereby Gulf War veterans over-report the prevalence of congenital anomalies amongst their children. In the Canadian study the reported prevalence was greater amongst Gulf War veterans even before and during their deployment.[18] Kang et al used pediatric evaluation blind to deployment status based on the veterans' description of birth defects together with newborn data to identify "likely" defects. Although over-reporting of birth defects occurred at a higher rate amongst Gulf veterans (40 percent versus 33 percent), the analyses based on these likely defects produced similar results to those quoted in this review (OR likely defect in male veterans = 1.94, 95 percent CI: 1.37, 2.74).[11]

Confounding and interactions.

Important confounding factors that could influence adverse reproductive outcomes might include family history of such outcomes, maternal age and health, alcohol and cigarette use during pregnancy, medication during pregnancy and ethnic origin. The Danish and Canadian studies matched Gulf and non-Gulf veterans on sex, age, and either their military duty or profession,[7,13] whilst the analysis of many of the more recent studies accounted for most of the variables listed above.[11-14,20] Even if the limited data on confounders were available, adjustments for confounders would have been somewhat meaningless in the analysis involving only seven cases of Goldenhar syndrome or other single rare birth defects.[16]

Quite different mechanisms might act to associate maternal or paternal exposure to a toxic agent before conception with individual birth defects. Relatively few women compared to men (although more women than had ever been deployed before) were deployed to the Gulf and there are not enough data to evaluate the role of maternal Gulf War exposure. Several studies did investigate the association between Gulf War deployment and risk of birth defect separately for deployed fathers and deployed mothers, but found no suggestion of an interaction via the sex of the deployed parent. Obviously these analyses were based on reduced sizes of sub-samples and reduced statistical power.

CONCLUSION

We identified thirteen primary studies which compared the prevalence of adverse reproductive outcomes in both Gulf War veterans and an unexposed comparison group.[8-20] A weak but statistically significant association between male veterans' deployment to the Gulf and spontaneous abortion of pregnancies following the war was demonstrated. Overall, there was little evidence to suggest that Gulf War veterans and their families are at increased risk of birth defects being diagnosed in their offspring. As always, there is a degree of uncertainty around any estimate of risk, this is especially true for rare outcomes. As a result the current published literature cannot exclude a possible increase or reduction in risk associated with Gulf War deployment. Furthermore, given the relatively small number of female veterans included in studies it is difficult to comment accurately on any potential risks in this specific population with any certainty.

This review did not aim to investigate possible associations between specific environmental exposures during the Gulf War and adverse reproductive outcomes. Although the review included adverse semen characteristics as a measurement of adverse reproductive outcomes we did not summarise data related to other indicators of infertility. However these have proved difficult to measure due to the very sensitive nature of the issue. Finally, any effect of potential teratogens that were experienced during the Gulf War was most likely to be demonstrated in the time period immediately after deployment. It is now unlikely that it will ever be possible to gather enough data of sufficient accuracy to provide clear answers for veterans of the Gulf War.

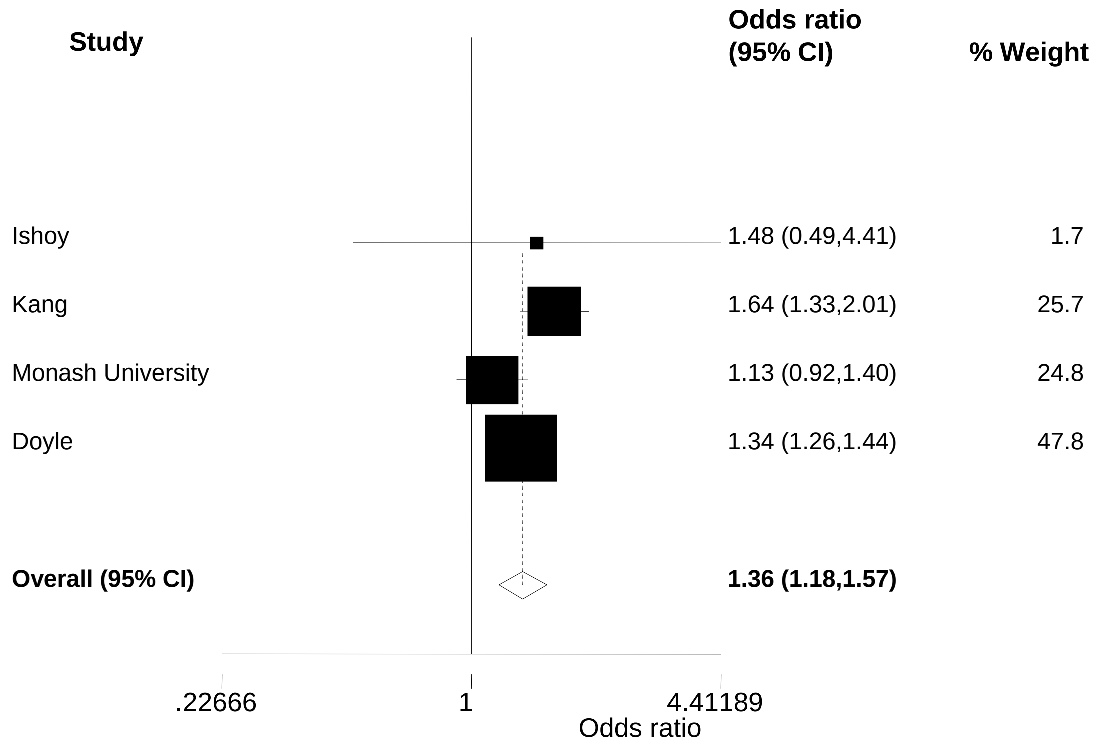
ACKNOWLEDGMENTS

The authors would like to thank Information Services at the School of Medicine, Cardiff University for their assistance in acquiring references for this review. This research was funded by the Ministry of Defence through the Medical Research Council.

REFERENCES

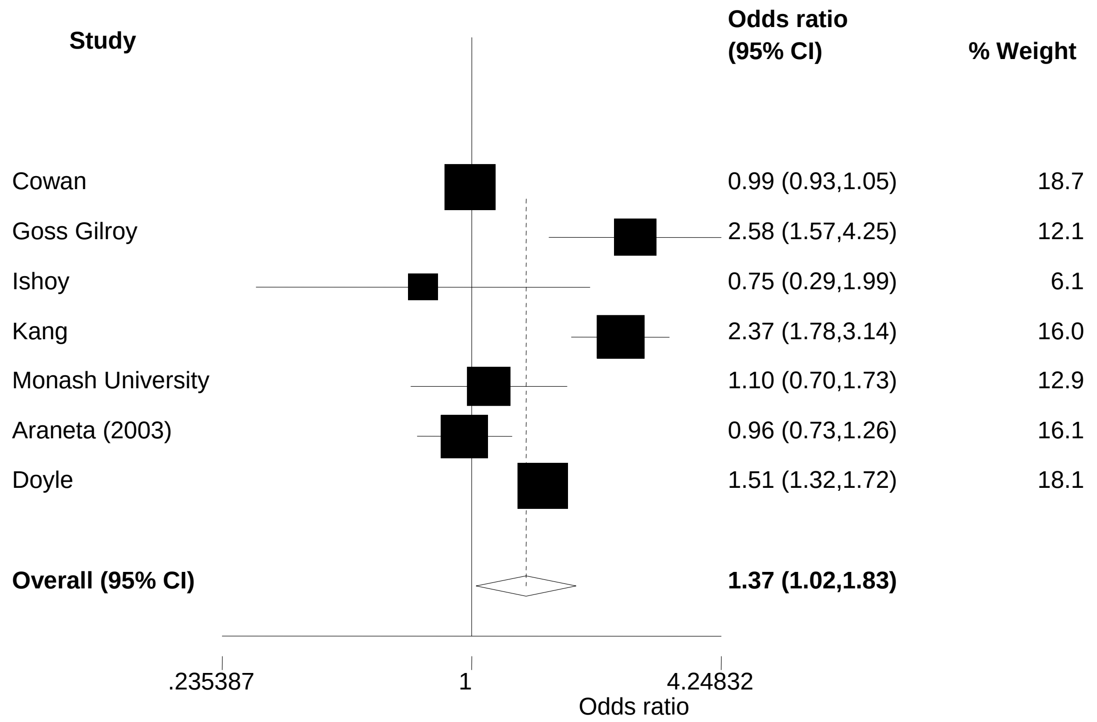
1. Briggs J. The tiny victims of Desert Storm. LIFE 1995:46-62.
2. Calderon RL, Wishik SM, Gray GC, et al. The Persian Gulf war - effects on human reproduction. 123rd Annual Meeting and Exhibition of the American Public Health Association 1995, Session 2198.
3. United Sates General Accounting Office (GAO). Operation Desert Storm: questions on possible exposure to reproductive toxins. Washington, DC. GAO/PEMD 30, 1994.
4. Stimpson NJ, Thomas HV, Weightman AL, et al. Psychiatric disorder in veterans of the Persian Gulf war of 1991: systematic review. Br J Psychiatry 2003; 182:391-403.
5. Egger M, Davey Smith G: Meta-analysis: potentials and promise. BMJ 1997, 315:1371-1374.
6. DerSimonian R, Laird N: Meta-analysis in clinical trials. Control Clin Trials 1986, 17:177-88.
7. Bradburn MJ, Deeks JJ, Altman DG: Sbe24: metan- an alternative meta-analysis command. Stata Technical Bulletin 1998, 44:4-15.
8. Ishoy T, Andersson AM, Suadicani P, et al. Major reproductive health characteristics in male Gulf War veterans: the Danish Gulf War study. Danish Medical Bulletin 2001;48(1):29-32.
9. Maconochie N, Doyle P, Carson C. Infertility among male UK veterans of the 1990-1 Gulf war: reproductive cohort study. BMJ. 2004;329(7459):196-201.
10. Rosa C. Spontaneous abortion rate and the Gulf War mobilization. J US Army Med Dep 1993.
11. Kang H, Magee C, Mahan C, et al. Pregnancy outcomes among U.S. Gulf War veterans: a population-based survey of 30,000 veterans. Ann Epidemiol 2001;11:504-11.
12. Monash University, 2003. Australian Gulf War Veterans’ Health Study Volume 2. http://www.dva.gov.au/media/publicat/2003/gulfwarhs/
13. Araneta MRG, Kamens DR, Zae AC, et al. Conception and pregnancy during the Persian Gulf War: the risk to women veterans. Ann Epidemiol 2004;14:109-116.
14. Doyle P, Maconochie N, Davies G et al. Miscarriage, stillbirth and congenital malformation in the offspring of UK veterans of the first Gulf war. Int J Epidemiol 2004;33:74-86.
15. Penman AD, Currier MM, Tarver RS. No evidence of increase in birth defects and health problems among children born to Persian Gulf War veterans in Mississippi. Military Medicine 1996;161(1):1-6.
16. Araneta MR, Moore CA, Olney RS, et al. Goldenhar syndrome among infants born in military hospitals to Gulf War veterans. Teratology 1997;56:244-51.
17. Cowan DN, DeFraites RF, Gray GC, et al. The risk of birth defects among children of Persian Gulf War veterans. New England Journal of Medicine 1997;336(23):1650-6.
18. Goss Gilroy, Inc. Health study of Canadian forces personnel involved in the 1991 conflict in the Persian Gulf. Vol 1. Ontario, Canada. Department of National Defence, 1998.
19. Araneta MR, Destiche DA, Schlangen KM, et al. Birth defects prevalence among infants of Persian Gulf War veterans born in Hawaii, 1989-1993. Teratology 2000;62(4):195-204.
20. Araneta MRG, Schlangen KM, Edmonds LD et al. Prevalence of birth defects among infants of Gulf War veterans in Arkansas, Arizona, California, Georgia, Hawaii, and Iowa, 1989-1993. Birth Defects Research (Part A) 2003;67:246-260.

FIGURE 1. Association between deployment to the Gulf War and spontaneous abortion during pregnancy among partners of male veterans.



Summary odds ratio calculated by random effects model using DerSimonian and Laird method. The estimate of heterogeneity between studies was taken from the inverse variance fixed effect model (χ^2 =6.2; 3 degrees of freedom; P=0.102).

FIGURE 2. Association between deployment to the Gulf War and birth defects among male veterans' children.



Summary odds ratio calculated by random effects model using DerSimonian and Laird method. The estimate of heterogeneity between studies was taken from the inverse variance fixed effect model (χ^2 =74.2; 6 degrees of freedom; P<0.001).

Table 1. Studies that have investigated the association between deployment to the Gulf War and adverse semen characteristics

First author	Study design	Sample	Study period	Outcomes	Published results	Ascertainment	Measurement	Confounding
Ishoy 2001	Cross-sectional survey	661 male Danish GWV 215 male Danish NGV	Jan 1997-Jan 1998	Oligospermia (low sperm count)	Prevalence GWV 1.6% NGV 1.6%	Unknown response rate	Hormone measurements have reasonably low detection levels and coefficients of variation	NGVs randomly selected to be matched on gender, age and profession
Maconochie 2004	Retrospective cohort study	732 male British GWV 370 male British GWV All consulted doctor for fertility problem. Arise from total sample 10465 British GWV 7376 British NGV	Aug 1998-March 2001	Oligospermia (low sperm count) Asthenospermia (poor motility) Teratospermia (abnormal morphology) Combined for any indicator of semen quality	Type I infertility (no conceptions) Prevalence any indicator GWV 97/10465 (0.9%) NGV 47/7376 (0.6%) OR 1.16 (95%CI 0.74-1.82) Type II infertility (no livebirths) Prevalence any indicator GWV 125/10465 (1.2%) NGV 54/7376 (0.7%) OR 1.45 (95%CI 0.98-2.14)	Total sample response rate 53% GWV men 42% NGV men Clinical data available for those with fertility problem: 33% GWV 32% NGV Failure to respond largely unrelated to reproductive factors	Relies on self-report of fertility. Highly sensitive issue.	ORs adjusted for age at attempt, age of partner at attempt, year of attempt, n of previous pregnancies (before Gulf War), service and ranks at time of Gulf War

GWV - Gulf War veterans
NGV - non Gulf veterans
OR – odds ratio
95% CI – 95% confidence interval

Table 2. Studies that have investigated the association between deployment to the Gulf War and spontaneous abortion (SAB) or stillbirth

First author	Study design	Sample	Study period	Outcomes	Published results	Sex of veteran	Ascertainment	Measurement	Confounding
Rosa 1993	Ecological study	12000 GWV from Fort Bliss, Texas 684 births + SABs 1990 historical comparison 810 births + SABs	May-Sept 1991	SAB	Prevalence 1991 84/684 (12.28%) Prevalence 1990 99/810 (12.22%)	Unclear whether pregnancies amongst GWV or partners of GWV	Includes some patients not at risk of Gulf exposure	Relatively mobile population affects yearly comparison & denominator	No account of confounding
Ishoy 2001	Cross-sectional survey	661 male Danish GWV 215 male Danish NGV (number of pregnancies unclear)	Jan 1997-Jan 1998	SAB after 1991	SAB prevalence GWV 18/661 (2.7%) NGV 4/215 (1.9%)	Male veterans only	Unknown response rate	Survey relies on self report without ascertainment Prevalence uses n of veterans as denominator	NGVs randomly selected to be matched on gender, age and profession
Kang 2001	Cross-sectional survey	11441 GWV (3397 pregnancies) 9476 NGV (2646 pregnancies) Stratified random sample taken from the DoD Manpower Data Center	Survey period unspecified Index pregnancy was first pregnancy ending after Jun 30 1991, but included beyond Jul 1996	SAB (death < 20 wk gestation) Stillbirth (death >20 wk gestation)	SAB prevalence Male veterans GWV 327/2739 (11.9%) NGV 148/1934 (7.7%) OR 1.62 (95% CI 1.32-1.99) Female veterans GWV 92/632 (14.6%) NGV 77/691 (11.1%) OR 1.35 (95% CI 0.97-1.89) Stillbirth prevalence Male veterans GWV 38/2739 (1.4%) NGV 16/1934 (0.8%) OR 1.65 (95% CI 0.91-2.98) Female veterans GWV 9/632 (1.4%) NGV 7/691 (1.0%) OR 1.26 (95% CI 0.46-3.49)	Results reported separately for male and female GWVs	75% GWV (15000 eligible) 64% NGV (15000 eligible) Non responders were more likely to be younger, unmarried, non-white and enlisted	Survey relies on self report without ascertainment Prevalence uses n of pregnancies as denominator	Analysis accounted for sampling design ORs adjusted for race, age of veteran, year of pregnancy, pregnancy history, smoking history, and military variables

Table 2. (contd)

First author	Study design	Sample	Study period	Outcomes	Published results	Sex of veteran	Ascertainment	Measurement	Confounding
Monash University 2003	Cross-sectional survey	1414 GWV (1448 pregnancies) 1411 NDV (1555 pregnancies) NDVs were random selection from members of Australian Defence Force	July 2000-April 2002. Adverse outcomes occurring in 1991 or later.	Miscarriage Stillbirth	Miscarriage prevalence GWV 204/1448 (14.1%) NGV 197/1555 (12.7%) Stillbirth prevalence GWV 5/1448 (0.4%) NGV 14/1555 (0.9%) Combined OR for both outcomes 1.1 (95% CI 0.8-1.3) P=0.709	Male veterans only	Response rate 78% GWV (1414/1808) 51% NDV (1411/2796)	Survey relies on self report without ascertainment. Data often only partially complete	OR adjusted for age, rank, service type, education, marital status, smoking, alcohol use, multiple pregnancies
Araneta 2004	Record linkage cross-sectional survey	3825 total admissions in 153 military hospitals. 427 conceptions NGV 415 conceptions during war GWV 298 postwar conception GWV	Pregnancy related admissions Aug 1990-May 1992 Questionnaire 1997-1998	SAB Stillbirth (ICD-9-CM discharge diagnostic codes)	SAB prevalence NGV conceptions 39/427 (9.1%) GWV conception during war 48/415 (11.6%) OR 1.44 (95% CI 0.9-2.3) GWV postwar conception 68/298 (22.8%) OR 2.92 (95%CI 1.9-4.6) Stillbirth prevalence NGV conceptions 10/427 (2.3%) GWV conception during war 1/415 (0.2%) GWV postwar conception 2/298 (0.7%)	Female veterans only	Questionnaire response 47%. Participants were more likely to be older, white, married, members of Marine Corps or Air Force, officers, and remained on active duty	Linkage study minimises measurement bias	ORs (comparison to NGV conceptions) adjusted for age, race, education, marital status, branch of service, enlisted rank, parity and reproductive history
Doyle 2004	Retrospective cohort study	Male veterans 16442 GWV pregnancies 11517 NGV pregnancies Female veterans 484 GWV pregnancies 377 NGV pregnancies	Aug 1998-March 2001	Miscarriage Stillbirth	Miscarriage prevalence Male veterans GWV 2903/16442 (18%) NGV 1584/11517 (14%) OR 1.4 (95%CI 1.3-1.5) Female veterans GWV 92/484 (20%) NGV 74/377 (21%) OR 1.0 (95%CI 0.7-1.4)	Results reported separately for male and female GWVs	Total sample response rate 53% GWV men 42% NGV men 72%GWV women 60%NGV women Failure to respond largely unrelated to reproductive factors	Relies on self-report	ORs adjusted for age at attempt, age of partner at attempt, year of attempt, n of previous pregnancies (before Gulf War), service

		Total eligible sample 52811 GWV 52924 NGV Stratified random sample supplied by the MoD			Stillbirth prevalence Male veterans GWV 77/16442 (0.6%) NGV 60/11517 (0.6%) OR 0.9 (95%CI 0.7-1.3) Female veterans GWV 3/484 (0.4%) NGV 1/377 (0.8%) OR 2.0 (95%CI 0.3-14.9)				and ranks at time of Gulf War. OR for stillbirth also adjusted for multiple pregnancies

GWV - Gulf War veterans
NGV - non Gulf veterans
SAB - spontaneous abortion
OR - odds ratio
95% CI – 95% confidence interval

Table 3. Studies that have investigated the association between deployment to the Gulf War and birth defects or infant deaths amongst veterans' offspring

First author	Study design	Sample	Study period	Outcomes	Main results	Sex of veteran	Ascertainment	Measurement	Confounding
Penman 1996	Cross-sectional survey	254 GWV (55 births) from 2 National Guard units chosen due to local media interest US surveillance comparison	Conception after deployment - May 1994 birth	Major birth defect Minor birth defect	Major birth defect n=3 (5%) (US comparison 2-7%) Minor birth defect n=2 (4%) (US comparison 5-10%)	52 families GWV father n=46 GWV mother n=5 GWV father and mother n=1	90% response (254/282) Data available for stillbirths but n=0	Birth records independently reviewed twice, but no estimate of reliability provided	No account of confounding
Araneta 1997	Cross-sectional survey	579931 GWV (34069 births) 700000 NGV (41345 births)	May 1991 - Sept 1993 livebirths	Goldenhar syndrome	Goldenhar syndrome n=7 GWV prevalence (n=5) 14.7/100,000 births NGV prevalence (n=2) 4.8/100,000 births Risk ratio 3.03 (95% CI 0.63-20.57)	All 7 cases had father in military	No data for still births/abortions births outside military hospitals, reserve or non-active duty parents	Outcome only assessed if defect recorded at birth, blind to deployment status. 2 raters concordant for all cases	Data available on confounders Results are unadjusted
Cowan 1997	Cross-sectional survey	579931 GWV (33998 births) 700000 NGV (41463 births)	Jan 1991 conception - Sept 1993 livebirth	Any birth defect (ICD-9-CM codes 740-759 + neoplasms + hereditary diseases) Severe birth defect (CDC definition)	Prevalence any defect GWV 2534/33998 (7.5%) NGV 3149/41463 (7.6%) OR 0.98 (95% CI 0.93-1.03) Prevalence severe defect GWV 630/33998 (1.9%) NGV 772/41463 (1.9%) OR 1.00 (95% CI 0.90-1.10)	Any birth defect Male veterans GWV 2137/30151 NGV 2339/32638 OR 0.99 (95% CI 0.93-1.05) Female veterans GWV 397/3847 NGV 810/8825 OR 1.12 (95%CI 1.00-1.25)	Data availability as in Araneta [3]	Birth defects as noted at time of birth. No data on reliability or blinding of data	Results are unadjusted; adjusting for marital status, ethnicity & service branch did not alter quoted results
Goss Gilroy 1998	Cross-sectional survey	3113 GWV (894 live+stillbirths) 3439 NGV (825 live+stillbirths)	Survey Jun-Sept 1997 Births after Sept 1991	Congenital anomaly	Prevalence GWV 59/894 (6.6%) NGV 22/825 (2.7%)	No data available	73% GWV (4262 eligible) 60% NGV (5699 eligible)	Survey relies on self report without ascertainment	NGVs matched on sex, age group and regular/ reserve status

Table 3. (contd)

First author	Study design	Sample	Study period	Outcomes	Main results	Sex of veteran	Ascertainment	Measurement	Confounding
Araneta 2000	Record linkage cross-sectional survey	99545 total births in Hawaii GWV 3717 births NGV 13465 births	Jan 1989 - Dec 1993 livebirths + 1yr followup	48 categories of birth defects	Prevalence ≥1 birth defect GWV 77/3717 (2.07%) NGV 290/13465 (2.15%) For GWV, no difference in rate pre/post deployment for each of 48 defects	Female veterans GWV 8 cases/ 165 births NGV 26 cases/ 1100 births OR 2.1 (95%CI 0.95-4.65)	Population sample of all livebirths in Hawaii No data for still births/abortions	Linkage study minimises measurement bias Prevalence uses n of livebirths as denominator	Data available on confounders but results are unadjusted
Ishoy 2001	Cross-sectional survey	661 Danish GWV 215 Danish NGV (n of pregnancies unclear)	Jan 1997-Jan 1998	Congenital disease or malformation Infant death within first year	Congenital anomaly GWV 14/661 (2.1%) NGV 6/215 (2.8%) Death within first year GWV 5/661 (0.8%) NGV 0/215 (0%)	Sample only included men, father must be military parent	Unknown response rate	Survey relies on self report Prevalence uses n of verterans as denominator	NGVs randomly selected to be matched on gender, age and profession
Kang 2001	Cross-sectional survey	11441 GWV (3397 pregnancies) 9476 NGV (2646 pregnancies) Stratified random sample taken from DoD Manpower Data Center	Survey period unspecified Index pregnancy was first pregnancy ending after Jun 30 1991, but included beyond Jul 1996	Any birth defect Infant death within first year	Birth defect prevalence Male veterans GWV 202/2236 (9.0%) NGV 68/1689 (4.0%) OR 2.34 (95%CI 1.76-3.10) Female veterans GWV 41/471 (8.7%) NGV 21/577 (3.6%) OR 2.85 (95%CI 1.62-4.99) Death within first year Male veterans GWV 12/2236 (0.5%) NGV 12/1689 (0.7%) OR 0.76 (95%CI 0.34-1.72) Female veterans GWV 3/471 (0.6%) NGV 6/577 (1.0%) OR 0.80 (95%CI 0.19-3.38)	Results reported separately for male and female GWVs	75% GWV (15000 eligible) 64% NGV (15000 eligible) Non responders were more likely to be younger, unmarried, non-white and enlisted	Survey relies on self report without ascertainment Prevalence uses n of livebirths as denominator	Analysis accounted for sampling design ORs adjusted for race, age of veteran, year of pregnancy, pregnancy history, smoking history, and military variables

Table 3. (contd)

First author	Study design	Sample	Study period	Outcomes	Main results	Sex of veteran	Ascertainment	Measurement	Confounding
Monash University 2003	Cross-sectional survey	1414 GWV (1096 children) 1411 NDV (1145 children) NDVs were random selection from members of Australian Defence Force	July 2000-April 2002. Health of liveborn children born in1992 or later.	Any birth defect Death	Birth defect prevalence GWV 40/1096 (3.6%) NGV 38/1145 (3.3%) OR 1.0 (95%CI 0.6-1.6) Death prevalence GWV 2/1096 (0.2%) NGV 4/1145 (0.3%)	Male veterans only	Response rate 78% GWV (1414/1808) 51% NDV (1411/2796)	Survey relies on self report without ascertainment. Prevalence uses n of livebirths as denominator	OR adjusted for age, rank, service type, education, marital status, smoking, alcohol use, multiple pregnancies
Araneta 2003	Record linkage cross-sectional survey	2,314,908 total births in Arizona, Hawaii, Iowa & selected counties of Arkansas, California and Georgia GWV 11961 births NGV 33052 births	Jan 1989 - Dec 1993 livebirths + 1yr followup	48 categories of birth defects	Prevalence ≥1 birth defect Infants conceived pre-war Male GWV 1.56% Male NGV 1.76% Female GWV 0.7% Female NGV 2.3% Infants conceived post-war Male GWV 1.55% Male NGV 1.61% RR 0.96 (95%CI 0.73-1.26) Female GWV 2.92% Female NGV 1.7% RR 1.68 (95%CI 0.81-3.48) For male GWVs and male and female NDVs no difference seen in rate pre/post deployment. Female GWV post vs pre-war RR 4.24 (95%CI 0.5-33.8) Specific birth defects Male GWV vs NDV Aortic valve stenosis RR 6.0 (95%CI 1.2-31.0) Coarctation of aorta RR 4.0 (95%CI 0.96-16.8) Tricuspid valve insufficiency RR 2.7 (95%CI 1.1-6.6) Female GWV vs NDV Hypospadias in sons RR 6.3 (95%CI 1.5-26.3)	Male veterans GWV 11511 births NGV 29086 births Female veterans GWV 450 births NGV 3966 births	Population sample of all livebirths for the following: Arizona 1989-92 Hawaii 1989-93 Iowa 1989-90 (& 2/3 for 91-93) No data for still births/abortions	Linkage study minimises measurement bias. Birth defects based on self-report. Prevalence uses n of livebirths as denominator	Differences in specific birth defects persisted after adjusting for state, maternal age, paternal age, race, marital status, education, plurality, parity, n of prenatal visits, gestational weight gain, branch of service and rank, prenatal alcohol and smoking exposure, low birth weight, small for gestational age and preeclampsia

First author	Study design	Sample	Study period	Outcomes	Main results	Sex of veteran	Ascertainment	Measurement	Confounding
Doyle 2004	Retrospective cohort study	Male veterans 16442 GWV pregnancies 11517 NGV pregnancies Female veterans 484 GWV pregnancies 377 NGV pregnancies Total eligible sample 52811 GWV 52924 NGV Stratified random sample supplied by the MoD	Aug 1998- March 2001	Congenital malformation	Any malformation prevalence Male veterans GWV n=686 (5.2%) NGV n=342 (3.5%) OR 1.5 (95%CI 1.3-1.7) Female veterans GWV n=19 (5.3%) NGV n=9 (3.2%) Specific malformations Male GWV vs NDV Digestive system RR 1.4 (95%CI 0.9-2.2) Genital malformation RR 1.8 (95%CI 1.0-3.0) Urinary system RR 1.6 (95%CI 1.1-2.3) Musculo-skeletal system RR 1.8 (95%CI 1.4-2.4)	Results reported separately for male and female GWVs	Total sample response rate 53% GWV men 42% NGV men 72%GWV women 60%NGV women Failure to respond largely unrelated to reproductive factors	Relies on self-report. Clinical confirmation for 55% of affected pregnancies. Associations weakened when restricted to clinically confirmed conditions. Prevalence uses n of livebirths, fetal deaths 16+ wks duration and terminations for medical reasons as denominator	ORs adjusted for age at attempt, age of partner at attempt, year of attempt, n of previous pregnancies (before Gulf War), service and ranks at time of Gulf War.

GWV - Gulf War veterans, NGV - non Gulf veterans, OR - odds ratio